Part X

PDE Examples

36

Some Examples of PDE's

Example 36.1 (Traffic Equation). Consider cars travelling on a straight road, i.e. $\mathbb{R}$ and let $u(t,x)$ denote the density of cars on the road at time t and space x and $v(t,x)$ be the velocity of the cars at (t,x). Then for $J = [a,b] \subset \mathbb{R}$, $N_J(t) := \int_a^b u(t,x)dx$ is the number of cars in the set J at time t. We must have

$$\int_a^b \dot{u}(t,x)dx = \dot{N}_J(t) = u(t,a)v(t,a) - u(t,b)v(t,b)$$
$$= -\int_a^b \frac{\partial}{\partial x}\left[u(t,x)v(t,x)\right]dx.$$

Since this holds for all intervals $[a,b]$, we must have

$$\dot{u}(t,x)dx = -\frac{\partial}{\partial x}\left[u(t,x)v(t,x)\right].$$

To make life more interesting, we may imagine that $v(t,x) = -F(u(t,x), u_x(t,x))$, in which case we get an equation of the form

$$\frac{\partial}{\partial t}u = \frac{\partial}{\partial x}G(u,u_x) \text{ where } G(u,u_x) = -u(t,x)F(u(t,x),u_x(t,x)).$$

A simple model might be that there is a constant maximum speed, v_m and maximum density u_m, and the traffic interpolates linearly between 0 (when $u = u_m$) to v_m when $(u = 0)$, i.e. $v = v_m(1 - u/u_m)$ in which case we get

$$\frac{\partial}{\partial t}u = -v_m\frac{\partial}{\partial x}\left(u(1 - u/u_m)\right).$$

Example 36.2 (Burger's Equation). Suppose we have a stream of particles travelling on $\mathbb{R}$, each of which has its own constant velocity and let $u(t,x)$ denote the velocity of the particle at x at time t. Let $x(t)$ denote the trajectory of the particle which is at x_0 at time t_0. We have $C = \dot{x}(t) = u(t,x(t))$. Differentiating this equation in t at $t = t_0$ implies

$$0 = [u_t(t, x(t)) + u_x(t, x(t))\dot{x}(t)]\,|_{t=t_0} = u_t(t_0, x_0) + u_x(t_0, x_0)u(t_0, x_0)$$

which leads to Burger's equation

$$0 = u_t + u\ u_x.$$

Example 36.3 (Minimal surface Equation). (Review Dominated convergence theorem and differentiation under the integral sign.) Let $D \subset \mathbb{R}^2$ be a bounded region with reasonable boundary, $u_0 : \partial D \to \mathbb{R}$ be a given function. We wish to find the function $u : D \to \mathbb{R}$ such that $u = u_0$ on ∂D and the graph of u, $\Gamma(u)$ has least area. Recall that the area of $\Gamma(u)$ is given by

$$A(u) = Area(\Gamma(u)) = \int_D \sqrt{1 + |\nabla u|^2} dx.$$

Assuming u is a minimizer, let $v \in C^1(D)$ such that $v = 0$ on ∂D, then

$$\begin{aligned} 0 &= \frac{d}{ds}|_0 A(u + sv) = \frac{d}{ds}|_0 \int_D \sqrt{1 + |\nabla(u + sv)|^2} dx \\ &= \int_D \frac{d}{ds}|_0 \sqrt{1 + |\nabla(u + sv)|^2} dx \\ &= \int_D \frac{1}{\sqrt{1 + |\nabla u|^2}} \nabla u \cdot \nabla v\ dx \\ &= -\int_D \nabla \cdot \left(\frac{1}{\sqrt{1 + |\nabla u|^2}} \nabla u \right) v\ dx \end{aligned}$$

from which it follows that

$$\nabla \cdot \left(\frac{1}{\sqrt{1 + |\nabla u|^2}} \nabla u \right) = 0.$$

Example 36.4 (Heat or Diffusion Equation). Suppose that $\Omega \subset \mathbb{R}^n$ is a region of space filled with a material, $\rho(x)$ is the density of the material at $x \in \Omega$ and $c(x)$ is the heat capacity. Let $u(t, x)$ denote the temperature at time $t \in [0, \infty)$ at the spatial point $x \in \Omega$. Now suppose that $B \subset \mathbb{R}^n$ is a "little" volume in $\mathbb{R}^n$, ∂B is the boundary of B, and $E_B(t)$ is the heat energy contained in the volume B at time t. Then

$$E_B(t) = \int_B \rho(x)c(x)u(t, x)dx.$$

So on one hand,

$$\dot{E}_B(t) = \int_B \rho(x)c(x)\dot{u}(t, x)dx \tag{36.1}$$

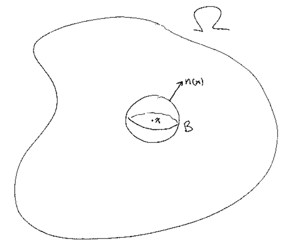


Fig. 36.1. A test volume B in Ω.

while on the other hand,

$$\dot{E}_B(t) = \int_{\partial B} \langle G(x)\nabla u(t, x), n(x)\rangle d\sigma(x), \tag{36.2}$$

where $G(x)$ is a $n \times n$–positive definite matrix representing the conduction properties of the material, $n(x)$ is the outward pointing normal to B at $x \in \partial B$, and $d\sigma$ denotes surface measure on ∂B. (We are using $\langle \cdot, \cdot \rangle$ to denote the standard dot product on $\mathbb{R}^n$.)

In order to see that we have the sign correct in (36.2), suppose that $x \in \partial B$ and $\nabla u(x) \cdot n(x) > 0$, then the temperature for points near x outside of B are hotter than those points near x inside of B and hence contribute to a increase in the heat energy inside of B. (If we get the wrong sign, then the resulting equation will have the property that heat flows from cold to hot!)

Comparing Eqs. (36.1) to (36.2) after an application of the divergence theorem shows that

$$\int_B \rho(x)c(x)\dot{u}(t, x)dx = \int_B \nabla \cdot (G(\cdot)\nabla u(t, \cdot))(x)\, dx. \tag{36.3}$$

Since this holds for all volumes $B \subset \Omega$, we conclude that the temperature functions should satisfy the following partial differential equation.

$$\rho(x)c(x)\dot{u}(t, x) = \nabla \cdot (G(\cdot)\nabla u(t, \cdot))(x)\,. \tag{36.4}$$

or equivalently that

$$\dot{u}(t, x) = \frac{1}{\rho(x)c(x)} \nabla \cdot (G(x)\nabla u(t, x)). \tag{36.5}$$

Setting $g^{ij}(x) := G_{ij}(x)/(\rho(x)c(x))$ and

$$z^j(x) := \sum_{i=1}^n \partial(G_{ij}(x)/(\rho(x)c(x)))/\partial x^i$$

the above equation may be written as:

$$\dot{u}(t,x) = Lu(t,x), \tag{36.6}$$

where

$$(Lf)(x) = \sum_{i,j} g^{ij}(x)\frac{\partial^2}{\partial x^i \partial x^j} f(x) + \sum_j z^j(x)\frac{\partial}{\partial x^j} f(x). \tag{36.7}$$

The operator L is a prototypical example of a second order "elliptic" differential operator.

Example 36.5 (Laplace and Poisson Equations). Laplaces Equation is of the form $Lu = 0$ and solutions may represent the steady state temperature distribution for the heat equation. Equations like $\Delta u = -\rho$ appear in electrostatics for example, where u is the electric potential and ρ is the charge distribution.

Example 36.6 (Shrodinger Equation and Quantum Mechanics).

$$i\frac{\partial}{\partial t}\psi(t,x) = -\frac{\Delta}{2}\psi(t,x) + V(x)\psi(t,x) \text{ with } \|\psi(\cdot,0)\|_2 = 1.$$

Interpretation,

$$\int_A |\psi(t,x)|^2\,dt = \text{ the probability of finding the particle in } A \text{ at time } t.$$

(Notice similarities to the heat equation.)

Example 36.7 (Wave Equation). Suppose that we have a stretched string supported at $x = 0$ and $x = L$ and $y = 0$. Suppose that the string only undergoes vertical motion (pretty bad assumption). Let $u(t,x)$ and $T(t,x)$ denote the height and tension of the string at (t,x), $\rho_0(x)$ denote the density in equilibrium and T_0 be the equilibrium string tension. Let $J = [x, x+\Delta x] \subset [0,L]$,

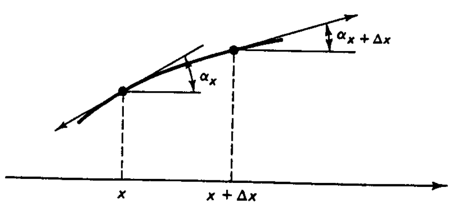


Fig. 36.2. A piece of displace string

then

$$M_J(t) := \int_J u_t(t,x)\rho_0(x)dx$$

is the momentum of the piece of string above J. (Notice that $\rho_0(x)dx$ is the weight of the string above x.) Newton's equations state

$$\frac{dM_J(t)}{dt} = \int_J u_{tt}(t,x)\rho_0(x)dx = \text{Force on String.}$$

Since the string is to only undergo vertical motion we require

$$T(t,x+\Delta x)\cos(\alpha_{x+\Delta x}) - T(t,x)\cos(\alpha_x) = 0$$

for all Δx and therefore that $T(t,x)\cos(\alpha_x) = T_0$, i.e.

$$T(t,x) = \frac{T_0}{\cos(\alpha_x)}.$$

The vertical tension component is given by

$$\begin{aligned} T(t,x+\Delta x)\sin(\alpha_{x+\Delta x}) - T(t,x)\sin(\alpha_x) &= T_0\left[\frac{\sin(\alpha_{x+\Delta x})}{\sin(\alpha_{x+\Delta x})} - \frac{\sin(\alpha_x)}{\cos(\alpha_x)}\right] \\ &= T_0\left[u_x(t,x+\Delta x) - u_x(t,x)\right]. \end{aligned}$$

Finally there may be a component due to gravity and air resistance, say

$$\text{gravity} = -\int_J \rho_0(x)dx \text{ and resistance} = -\int_J k(x)u_t(t,x)dx.$$

So Newton's equations become

$$\begin{aligned} \int_x^{x+\Delta x} u_{tt}(t,x)\rho_0(x)dx &= T_0\left[u_x(t,x+\Delta x) - u_x(t,x)\right] \\ &\quad - \int_x^{x+\Delta x}\rho_0(x)dx - \int_x^{x+\Delta x} k(x)u_t(t,x)dx \end{aligned}$$

and differentiating this in Δx gives

$$u_{tt}(t,x)\rho_0(x) = u_{xx}(t,x) - \rho_0(x) - k(x)u_t(t,x)$$

or equivalently that

$$u_{tt}(t,x) = \frac{1}{\rho_0(x)}u_{xx}(t,x) - 1 - \frac{k(x)}{\rho_0(x)}u_t(t,x). \tag{36.8}$$

Example 36.8 (Maxwell Equations in Free Space).

$$\begin{aligned} \frac{\partial \mathbf{E}}{\partial t} &= \nabla \times \mathbf{B} \\ \frac{\partial \mathbf{B}}{\partial t} &= -\nabla \times \mathbf{E} \\ \nabla \cdot \mathbf{E} &= \nabla \cdot \mathbf{B} = 0. \end{aligned}$$

Notice that

$$\frac{\partial^2 \mathbf{E}}{\partial t^2} = \nabla \times \frac{\partial \mathbf{B}}{\partial t} = -\nabla \times (\nabla \times \mathbf{E}) = \Delta \mathbf{E} - \nabla (\nabla \cdot \mathbf{E}) = \Delta \mathbf{E}$$

and similarly, $\frac{\partial^2 \mathbf{B}}{\partial t^2} = \Delta \mathbf{B}$ so that all the components of the electromagnetic fields satisfy the wave equation.

Example 36.9 (Navier – Stokes). Here $u(t,x)$ denotes the velocity of a fluid ad (t,x), $p(t,x)$ is the pressure. The Navier – Stokes equations state,

$$\frac{\partial u}{\partial t} + \partial_u u = \nu \Delta u - \nabla p + f \text{ with } u(0,x) = u_0(x) \tag{36.9}$$

$$\nabla \cdot u = 0 \text{ (incompressibility)} \tag{36.10}$$

where f are the components of a given external force and u_0 is a given divergence free vector field, ν is the viscosity constant. The Euler equations are found by taking $\nu = 0$. Equation (36.9) is Newton's law of motion again, $F = ma$. See http://www.claymath.org for more information on this Million Dollar problem.

36.1 Some More Geometric Examples

Example 36.10 (Einstein Equations). Einstein's equations from general relativity are

$$\mathrm{Ric}_g - \frac{1}{2} g S_g = T$$

where T is the stress energy tensor.

Example 36.11 (Yamabe Problem). Does there exists a metric $g_1 = u^{4/(n-2)} g_0$ in the conformal class of g_0 so that g_1 has constant scalar curvature. This is equivalent to solving

$$-\gamma \Delta_{g_0} u + S_{g_0} u = k u^{\alpha}$$

where $\gamma = 4\frac{n-1}{n-2}$, $\alpha = \frac{n+2}{n-2}$, k is a constant and S_{g_0} is the scalar curvature of g_0.

Example 36.12 (Ricci Flow). Hamilton introduced the Ricci – flow,

$$\frac{\partial g}{\partial t} = \mathrm{Ric}_g,$$

as another method to create "good" metrics on manifolds. This is a possible solution to the 3 dimensional Poincaré conjecture, again go to the Clay math web site for this problem.

Part XI

First Order Scalar Equations

37

First Order Quasi-Linear Scalar PDE

37.1 Linear Evolution Equations

Consider the first order partial differential equation

$$\partial_t u(t,x) = \sum_{i=1}^{n} a_i(x)\partial_i u(t,x) \text{ with } u(0,x) = f(x) \tag{37.1}$$

where $x \in \mathbb{R}^n$ and $a_i(x)$ are smooth functions on $\mathbb{R}^n$. Let $A(x) = (a_1(x),\dots,a_n(x))$ and for $u \in C^1(\mathbb{R}^n,\mathbb{C})$, let

$$\tilde{A}u(x) := \frac{d}{dt}|_0 u(x+tZ(x)) = \nabla u(x)\cdot A(x) = \sum_{i=1}^{n} a_i(x)\partial_i u(x),$$

i.e. $\tilde{A}(x)$ is the first order differential operator, $\tilde{A}(x) = \sum_{i=1}^{n} a_i(x)\partial_i$. With this notation we may write Eq. (37.1) as

$$\partial_t u = \tilde{A}u \text{ with } u(0,\cdot) = f. \tag{37.2}$$

The following lemma contains the key observation needed to solve Eq. (37.2).

Lemma 37.1. *Let A and $\tilde{A}$ be as above and $f \in C^1(\mathbb{R}^n,\mathbb{R})$, then*

$$\frac{d}{dt} f\circ e^{tA}(x) = \tilde{A}f\circ e^{tA}(x) = \tilde{A}\left(f\circ e^{tA}\right)(x). \tag{37.3}$$

Proof. By definition,

$$\frac{d}{dt}e^{tA}(x) = A(e^{tA}(x))$$

and so by the chain rule

$$\frac{d}{dt} f\circ e^{tA}(x) = \nabla f(e^{tA}(x))\cdot A(e^{tA}(x)) = \tilde{A}f(e^{tA}(x))$$

which proves the first Equality in Eq. (37.3). For the second we will need to use the following two facts: 1) $e^{(t+s)A} = e^{tA}\circ e^{sZ}$ and 2) $e^{tA}(x)$ is smooth in x. Assuming this we find

$$\frac{d}{dt} f\circ e^{tA}(x) = \frac{d}{ds}|_0 f\circ e^{(t+s)A}(x) = \frac{d}{ds}|_0\left[f\circ e^{tA}\circ e^{sZ}(x)\right] = \tilde{A}\left(f\circ e^{tA}\right)(x)$$

which is the second equality in Eq. (37.3). ■

Theorem 37.2. *The function $u \in C^1(\mathcal{D}(A),\mathbb{R})$ defined by*

$$u(t,x) := f(e^{tA}(x)) \tag{37.4}$$

solves Eq. (37.2). Moreover this is the unique function defined on $\mathcal{D}(A)$ which solves Eq. (37.3).

Proof. Suppose that $u \in C^1(\mathcal{D}(A),\mathbb{R})$ solves Eq. (37.2), then

$$\frac{d}{dt}u(t,e^{-tA}(x)) = u_t(t,e^{-tA}(x)) - \tilde{A}u(t,e^{-tA}(x)) = 0$$

and hence

$$u(t,e^{-tA}(x)) = u(0,x) = f(x).$$

Let $(t_0,x_0) \in \mathcal{D}(A)$ and apply the previous computations with $x = e^{tA}(x_0)$ to find $u(t_0,x) = f(e^{tA}(x_0))$. This proves the uniqueness assertion. The verification that u defined in Eq. (37.4) solves Eq. (37.2) is simply the second equality in Eq. (37.3). ■

Notation 37.3 *Let $e^{t\tilde{A}}f(x) = u(t,x)$ where u solves Eq. (37.2), i.e.*

$$e^{t\tilde{A}}f(x) = f(e^{tA}(x)).$$

The differential operator $\tilde{A}: C^1(\mathbb{R}^n,\mathbb{R}) \to \mathbb{C}(\mathbb{R}^n,\mathbb{R})$ is no longer bounded so it is not possible in general to conclude

$$e^{t\tilde{A}}f = \sum_{n=0}^{\infty}\frac{t^n}{n!}\tilde{A}^n f. \tag{37.5}$$

Indeed, to make sense out of the right side of Eq. (37.5) we must know f is infinitely differentiable and that the sum is convergent. This is typically not the case. because if f is only C^1. However there is still some truth to Eq. (37.5). For example if $f \in C^k(\mathbb{R}^n,\mathbb{R})$, then by Taylor's theorem with remainder,

$$e^{t\tilde{A}}f - \sum_{n=0}^{k}\frac{t^n}{n!}\tilde{A}^n f = o(t^k)$$

by which I mean, for any $x \in \mathbb{R}^n$,

$$t^{-k}\left[e^{t\tilde{A}}f(x) - \sum_{n=0}^{k}\frac{t^n}{n!}\tilde{A}^n f(x)\right] \to 0 \text{ as } t \to 0.$$

Example 37.4. Suppose $n = 1$ and $A(x) = 1$, $\tilde{A}(x) = \partial_x$ then $e^{tA}(x) = x + t$ and hence

$$e^{t\partial_x} f(x) = f(x + t).$$

It is interesting to notice that

$$e^{t\partial_x} f(x) = \sum_{n=0}^{\infty} \frac{t^n}{n!} f^{(n)}(x)$$

is simply the Taylor series expansion of $f(x + t)$ centered at x. This series converges to the correct answer (i.e. $f(x+t)$) iff f is "real analytic." For more details see the Cauchy – Kovalevskaya Theorem in Section 39.

Example 37.5. Suppose $n = 1$ and $A(x) = x^2$, $\tilde{A}(x) = x^2\partial_x$ then $e^{tA}(x) = \frac{x}{1-tx}$ on $\mathcal{D}(A) = \{(t, x) : 1 - tx > 0\}$ and hence $e^{t\tilde{A}} f(x) = f(\frac{x}{1-tx}) = u(t, x)$ on $\mathcal{D}(A)$, where

$$u_t = x^2 u_x. \tag{37.6}$$

It may or may not be possible to extend this solution, $u(t, x)$, to a C^1 solution on all $\mathbb{R}^2$. For example if $\lim_{x\to\infty} f(x)$ does not exist, then $\lim_{t\uparrow x} u(t, x)$ does not exist for any $x > 0$ and so u can not be the restriction of C^1 – function on $\mathbb{R}^2$. On the other hand if there are constants $c_\pm$ and $M > 0$ such that $f(x) = c_+$ for $x > M$ and $f(x) = c_-$ for $x < -M$, then we may extend u to all $\mathbb{R}^2$ by defining

$$u(t, x) = \begin{cases} c_+ \text{ if } x > 0 \text{ and } t > 1/x \\ c_- \text{ if } x < 0 \text{ and } t < 1/x. \end{cases}$$

It is interesting to notice that $x(t) = 1/t$ solves $\dot{x}(t) = -x^2(t) = -A(x(t))$, so any solution $u \in C^1(\mathbb{R}^2, \mathbb{R})$ to Eq. (37.6) satisfies $\frac{d}{dt} u(t, 1/t) = 0$, i.e. u must be constant on the curves $x = 1/t$ for $t > 0$ and $x = 1/t$ for $t < 0$. See Example 37.13 below for a more detailed study of Eq. (37.6).

Example 37.6. Suppose $n = 2$.

1. If $A(x, y) = (-y, x)$, i.e. $A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$ then

$$e^{tA} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

and hence

$$e^{t\tilde{A}} f(x, y) = f\left(x \cos t - y \sin t, y \cos t + x \sin t\right).$$

2. If $A(x, y) = (x, y)$, i.e. $A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$ then

$$e^{tA} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} e^t & 0 \\ 0 & e^t \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

and hence

$$e^{t\tilde{A}} f(x, y) = f\left(xe^t, ye^t\right).$$

Theorem 37.7. *Given $A \in C^1(\mathbb{R}^n, \mathbb{R}^n)$ and $h \in C^1(\mathbb{R} \times \mathbb{R}^n, \mathbb{R})$.*

1. *(Duhamel' s Principle) The unique solution $u \in C^1(\mathcal{D}(A), \mathbb{R})$ to*

$$u_t = \tilde{A}u + h \text{ with } u(0, \cdot) = f \tag{37.7}$$

is given by

$$u(t, \cdot) = e^{t\tilde{A}} f + \int_0^t e^{(t-\tau)\tilde{A}} h(\tau, \cdot) d\tau$$

or more explicitly,

$$u(t, x) := f(e^{tA}(x)) + \int_0^t h(\tau, e^{(t-\tau)A}(x)) d\tau. \tag{37.8}$$

2. *The unique solution $u \in C^1(\mathcal{D}(A), \mathbb{R})$ to*

$$u_t = \tilde{A}u + hu \text{ with } u(0, \cdot) = f \tag{37.9}$$

is given by

$$u(t, \cdot) = e^{\int_0^t h(\tau, e^{(t-\tau)A}(x)) d\tau} f(e^{tA}(x)) \tag{37.10}$$

which we abbreviate as

$$e^{t(\tilde{A}+M_h)} f(x) = e^{\int_0^t h(\tau, e^{(t-\tau)A}(x)) d\tau} f(e^{tA}(x)). \tag{37.11}$$

Proof. We will verify the uniqueness assertions, leaving the routine check the Eqs. (37.8) and (37.9) solve the desired PDE's to the reader. Assuming u solves Eq. (37.7), we find

$$\begin{aligned} \frac{d}{dt}\left[e^{-t\tilde{A}} u(t, \cdot)\right](x) &= \frac{d}{dt} u(t, e^{-tA}(x)) = \left(u_t - \tilde{A}u\right)(t, e^{-tA}(x)) \\ &= h(t, e^{-tA}(x)) \end{aligned}$$

and therefore

$$\left[e^{-t\tilde{A}} u(t, \cdot)\right](x) = u(t, e^{-tA}(x)) = f(x) + \int_0^t h(\tau, e^{-\tau A}(x)) d\tau$$

and so replacing x by $e^{tA}(x)$ in this equation implies

$$u(t, x) = f(e^{tA}(x)) + \int_0^t h(\tau, e^{(t-\tau)A}(x)) d\tau.$$

Similarly if u solves Eq. (37.9), we find with $z(t) := \left[e^{-t\tilde{A}}u(t,\cdot)\right](x) = u(t, e^{-tA}(x))$ that

$$\begin{aligned}\dot{z}(t) &= \frac{d}{dt}u(t,e^{-tA}(x)) = \left(u_t - \tilde{A}u\right)(t, e^{-tA}(x)) \\ &= h(t,e^{-tA}(x))u(t,e^{-tA}(x)) = h(t,e^{-tA}(x))z(t).\end{aligned}$$

Solving this equation for $z(t)$ then implies

$$u(t,e^{-tA}(x)) = z(t) = e^{\int_0^t h(\tau,e^{-\tau A}(x))d\tau}z(0) = e^{\int_0^t h(\tau,e^{-\tau A}(x))d\tau}f(x).$$

Replacing x by $e^{tA}(x)$ in this equation implies

$$u(t,x) = e^{\int_0^t h(\tau,e^{(t-\tau)A}(x))d\tau}f(e^{tA}(x)).$$

■

Remark 37.8. It is interesting to observe the key point to getting the simple expression in Eq. (37.11) is the fact that

$$e^{t\tilde{A}}(fg) = (fg)\circ e^{tA} = \left(f\circ e^{tA}\right)\cdot\left(g\circ e^{tA}\right) = e^{t\tilde{A}}f\cdot e^{t\tilde{A}}g.$$

That is to say $e^{t\tilde{A}}$ is an algebra homomorphism on functions. This property does not happen for any other type of differential operator. Indeed, if L is some operator on functions such that $e^{tL}(fg) = e^{tL}f\cdot e^{tL}g$, then differentiating at $t=0$ implies

$$L(fg) = Lf\cdot g + f\cdot Lg,$$

i.e. L satisfies the product rule. One learns in differential geometry that this property implies L must be a vector field.

Let us now use this result to find the solution to the wave equation

$$u_{tt} = u_{xx} \text{ with } u(0,\cdot) = f \text{ and } u_t(0,\cdot) = g. \tag{37.12}$$

To this end, let us notice the $u_{tt} = u_{xx}$ may be written as

$$(\partial_t - \partial_x)(\partial_t + \partial_x)u = 0$$

and therefore noting that

$$(\partial_t + \partial_x)u(t,x)|_{t=0} = g(x) + f'(x)$$

we have

$$(\partial_t + \partial_x)u(t,x) = e^{t\partial_x}(g+f')(x) = (g+f')(x+t).$$

The solution to this equation is then a consequence of Duhamel' s Principle which gives

$$\begin{aligned}u(t,x) &= e^{-t\partial_x}f(x) + \int_0^t e^{-(t-\tau)\partial_x}(g+f')(x+\tau)d\tau \\ &= f(x-t) + \int_0^t (g+f')(x+\tau-(t-\tau))d\tau \\ &= f(x-t) + \int_0^t (g+f')(x+2\tau-t)d\tau \\ &= f(x-t) + \int_0^t g(x+2\tau-t)d\tau + \frac{1}{2}f(x+2\tau-t)|_{\tau=0}^{\tau=t} \\ &= \frac{1}{2}\left[f(x+t)+f(x-t)\right] + \frac{1}{2}\int_{-t}^t g(x+s)ds.\end{aligned}$$

The following theorem summarizes what we have proved.

Theorem 37.9. *If $f \in C^2(\mathbb{R},\mathbb{R})$ and $g \in C^1(\mathbb{R},\mathbb{R})$, then Eq. (37.12) has a unique solution given by*

$$u(t,x) = \frac{1}{2}\left[f(x+t)+f(x-t)\right] + \frac{1}{2}\int_{-t}^t g(x+s)ds. \tag{37.13}$$

Proof. We have already proved uniqueness above. The proof that u defined in Eq. (37.13) solves the wave equation is a routine computation. Perhaps the most instructive way to verify that u solves $u_{tt} = u_{xx}$ is to observe, letting $y = x+s$, that

$$\begin{aligned}\int_{-t}^t g(x+s)ds &= \int_{x-t}^{x+t} g(y)dy = \int_0^{x+t} g(y)dy + \int_{x-t}^0 g(y)dy \\ &= \int_0^{x+t} g(y)dy - \int_0^{x-t} g(y)dy.\end{aligned}$$

From this observation it follows that

$$u(t,x) = F(x+t) + G(x-t)$$

where

$$F(x) = \frac{1}{2}\left(f(x) + \int_0^x g(y)dy\right) \text{ and } G(x) = \frac{1}{2}\left(f(x) - \int_0^x g(y)dy\right).$$

Now clearly F and G are C^2 – functions and

$$(\partial_t - \partial_x)F(x+t) = 0 \text{ and } (\partial_t + \partial_x)G(x-t) = 0$$

so that

$$\left(\partial_t^2 - \partial_x^2\right)u(t,x) = (\partial_t - \partial_x)(\partial_t + \partial_x)(F(x+t) + G(x-t)) = 0.$$

■

Now let us formally apply Exercise 37.45 to the wave equation $u_{tt} = u_{xx}$, in which case we should let $A^2 = -\partial_x^2$, and hence $A = \sqrt{-\partial_x^2}$. Evidently we should take

$$\cos\left(t\sqrt{-\partial_x^2}\right) f(x) = \frac{1}{2}\left[f(x+t) + f(x-t)\right] \text{ and}$$
$$\frac{\sin\left(t\sqrt{-\partial_x^2}\right)}{\sqrt{-\partial_x^2}} g(x) = \frac{1}{2}\int_{-t}^{t} g(x+s)ds = \frac{1}{2}\int_{x-t}^{x+t} g(y)dy$$

Thus with these definitions, we can try to solve the equation

$$u_{tt} = u_{xx} + h \text{ with } u(0,\cdot) = f \text{ and } u_t(0,\cdot) = g \tag{37.14}$$

by a formal application of Exercise 37.43. According to Eq. (37.73) we should have

$$u(t,\cdot) = \cos(tA)f + \frac{\sin(tA)}{A}g + \int_0^t \frac{\sin((t-\tau)A)}{A}h(\tau,\cdot)d\tau,$$

i.e.

$$u(t,x) = \frac{1}{2}\left[f(x+t) + f(x-t)\right] + \frac{1}{2}\int_{-t}^{t} g(x+s)ds + \frac{1}{2}\int_0^t d\tau \int_{x-t+\tau}^{x+t-\tau} dy\, h(\tau,y). \tag{37.15}$$

An alternative way to get to this equation is to rewrite Eq. (37.14) in first order (in time) form by introducing $v = u_t$ to find

$$\frac{\partial}{\partial t}\begin{pmatrix} u \\ v \end{pmatrix} = A\begin{pmatrix} u \\ v \end{pmatrix} + \begin{pmatrix} 0 \\ h \end{pmatrix} \text{ with}$$
$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix} \text{ at } t = 0 \tag{37.16}$$

where

$$A := \begin{pmatrix} 0 & 1 \\ \frac{\partial^2}{\partial x^2} & 0 \end{pmatrix}.$$

A restatement of Theorem 37.9 is simply

$$e^{tA}\begin{pmatrix} f \\ g \end{pmatrix}(x) = \begin{pmatrix} u(t,x) \\ u_t(t,x) \end{pmatrix} = \frac{1}{2}\begin{pmatrix} f(x+t) + f(x-t) + \int_{-t}^{t} g(x+s)ds \\ f'(x+t) - f'(x-t) + g(x+t) + g(x-t) \end{pmatrix}.$$

According to Du hamel's principle the solution to Eq. (37.16) is given by

$$\begin{pmatrix} u(t,\cdot) \\ u_t(t,\cdot) \end{pmatrix} = e^{tA}\begin{pmatrix} f \\ g \end{pmatrix} + \int_0^t e^{(t-\tau)A}\begin{pmatrix} 0 \\ h(\tau,\cdot) \end{pmatrix} d\tau.$$

The first component of the last term is given by

$$\frac{1}{2}\int_0^t \left[\int_{\tau-t}^{t-\tau} h(\tau, x+s)ds\right] d\tau = \frac{1}{2}\int_0^t \left[\int_{x-t+\tau}^{x+t-\tau} h(\tau,y)dy\right] d\tau$$

which reproduces Eq. (37.15).

To check Eq. (37.15), it suffices to assume $f = g = 0$ so that

$$u(t,x) = \frac{1}{2}\int_0^t d\tau \int_{x-t+\tau}^{x+t-\tau} dy\, h(\tau,y).$$

Now

$$\begin{aligned}
u_t &= \frac{1}{2}\int_0^t \left[h(\tau, x+t-\tau) + h(\tau, x-t+\tau)\right] d\tau, \\
u_{tt} &= \frac{1}{2}\int_0^t \left[h_x(\tau, x+t-\tau) - h_x(\tau, x-t+\tau)\right] d\tau + h(t,x) \\
u_x(t,x) &= \frac{1}{2}\int_0^t d\tau \left[h(\tau, x+t-\tau) - h(\tau, x-t+\tau)\right] \text{ and} \\
u_{xx}(t,x) &= \frac{1}{2}\int_0^t d\tau \left[h_x(\tau, x+t-\tau) - h_x(\tau, x-t+\tau)\right]
\end{aligned}$$

so that $u_{tt} - u_{xx} = h$ and $u(0,x) = u_t(0,x) = 0$. We have proved the following theorem.

Theorem 37.10. *If $f \in C^2(\mathbb{R},\mathbb{R})$ and $g \in C^1(\mathbb{R},\mathbb{R})$, and $h \in C(\mathbb{R}^2,\mathbb{R})$ such that h_x exists and $h_x \in C(\mathbb{R}^2,\mathbb{R})$, then Eq. (37.14) has a unique solution $u(t,x)$ given by Eq. (37.14).*

Proof. The only thing left to prove is the uniqueness assertion. For this suppose that v is another solution, then $(u - v)$ solves the wave equation (37.12) with $f = g = 0$ and hence by the uniqueness assertion in Theorem 37.9, $u - v \equiv 0$. ■

37.1.1 A 1-dimensional wave equation with non-constant coefficients

Theorem 37.11. *Let $c(x) > 0$ be a smooth function and $\tilde{C} = c(x)\partial_x$ and $f, g \in C^2(\mathbb{R})$. Then the unique solution to the wave equation*

$$u_{tt} = \tilde{C}^2 u = c u_{xx} + c' u_x \text{ with } u(0,\cdot) = f \text{ and } u_t(0,\cdot) = g \tag{37.17}$$

is

$$u(t,x) = \frac{1}{2}\left[f(e^{-tC}(x)) + f(e^{tC}(x))\right] + \frac{1}{2}\int_{-t}^{t} g(e^{sC}(x))ds. \tag{37.18}$$

defined for $(t,x) \in \mathcal{D}(C) \cap \mathcal{D}(-C)$.

Proof. (Uniqueness) If u is a C^2 – solution of Eq. (37.17), then

$$\left(\partial_t - \tilde{C}\right)\left(\partial_t + \tilde{C}\right) u = 0$$

and

$$\left(\partial_t + \tilde{C}\right) u(t,x)|_{t=0} = g(x) + \tilde{C} f(x).$$

Therefore

$$\left(\partial_t + \tilde{C}\right) u(t,x) = e^{t\tilde{C}} \left(g + f'\right)(x) = \left(g + f'\right)\left(e^{tC}(x)\right)$$

which has solution given by Duhamel' s Principle as

$$\begin{aligned} u(t,x) &= e^{-t\tilde{A}} f(x) + \int_0^t e^{-(t-\tau)\tilde{C}} \left(g + \tilde{C}f\right)\left(e^{\tau C}(x)\right)d\tau \\ &= f(e^{-tC}(x)) + \int_0^t \left(g + \tilde{C}f\right)\left(e^{(2\tau - t)C}(x)\right)d\tau \\ &= f(e^{-tC}(x)) + \frac{1}{2}\int_{-t}^t \left(g + \tilde{C}f\right)\left(e^{sC}(x)\right)ds \\ &= f(e^{-tC}(x)) + \frac{1}{2}\int_{-t}^t g(e^{sC}(x))ds + \frac{1}{2}\int_{-t}^t \frac{d}{ds} f(e^{sC}(x))ds \\ &= \frac{1}{2}\left[f(e^{-tC}(x)) + f(e^{tC}(x))\right] + \frac{1}{2}\int_{-t}^t g(e^{sC}(x))ds. \end{aligned}$$

(Existence.) Let $y = e^{sC}(x)$ so $dy = c(e^{sC}(x))ds = c(y)ds$ in the integral in Eq. (37.18), then

$$\begin{aligned} \int_{-t}^t g(e^{sC}(x))ds &= \int_{e^{-tC}(x)}^{e^{tC}(x)} g(y)\frac{dy}{c(y)} \\ &= \int_0^{e^{tC}(x)} g(y)\frac{dy}{c(y)} + \int_{e^{-tC}(x)}^0 g(y)\frac{dy}{c(y)} \\ &= \int_0^{e^{tC}(x)} g(y)\frac{dy}{c(y)} - \int_0^{e^{-tC}(x)} g(y)\frac{dy}{c(y)}. \end{aligned}$$

From this observation, it follows follows that

$$u(t,x) = F(e^{tC}(x)) + G(e^{-tC}(x))$$

where

$$F(x) = \frac{1}{2}\left(f(x) + \int_0^x g(y)\frac{dy}{c(y)}\right) \text{ and } G(x) = \frac{1}{2}\left(f(x) - \int_0^x g(y)\frac{dy}{c(y)}\right).$$

Now clearly F and G are C^2 – functions and

$$\left(\partial_t - \tilde{C}\right) F(e^{tC}(x)) = 0 \text{ and } \left(\partial_t + \tilde{C}\right) G(e^{-tC}(x)) = 0$$

so that

$$\left(\partial_t^2 - \tilde{C}^2\right) u(t,x) = \left(\partial_t - \tilde{C}\right)\left(\partial_t + \tilde{C}\right)\left[F(e^{tC}(x)) + G(e^{-tC}(x))\right] = 0.$$

■

By Du hamel's principle, we can similarly solve

$$u_{tt} = \tilde{C}^2 u + h \text{ with } u(0,\cdot) = 0 \text{ and } u_t(0,\cdot) = 0. \tag{37.19}$$

Corollary 37.12. *The solution to Eq. (37.19) is*

$$\begin{aligned} u(t,x) &= \frac{1}{2}\int_0^t \begin{pmatrix} \text{Solution to Eq. (37.17)} \\ \text{at time } t-\tau \\ \text{with } f = 0 \text{ and } g = h(\tau,\cdot) \end{pmatrix} d\tau \\ &= \frac{1}{2}\int_0^t d\tau \int_{\tau - t}^{t-\tau} h(\tau, e^{sC}(x))ds. \end{aligned}$$

Proof. This is simply a matter of computing a number of derivatives:

$$\begin{aligned} u_t &= \frac{1}{2}\int_0^t d\tau \left[h(\tau, e^{(t-\tau)C}(x)) + h(\tau, e^{(\tau - t)C}(x))\right] \\ u_{tt} &= h(t,x) + \frac{1}{2}\int_0^t d\tau \left[\tilde{C}h(\tau, e^{(t-\tau)C}(x)) - \tilde{C}h(\tau, e^{(\tau-t)C}(x))\right] \\ \tilde{C}u &= \frac{1}{2}\int_0^t d\tau \int_{\tau-t}^{t-\tau} \tilde{C}h(\tau, e^{sC}(x))ds = \frac{1}{2}\int_0^t d\tau \int_{\tau-t}^{t-\tau} \frac{d}{ds} h(\tau, e^{sC}(x))ds \\ &= \frac{1}{2}\int_0^t d\tau \left[h(\tau, e^{(t-\tau)C}(x)) - h(\tau, e^{(\tau-t)C}(x))\right] \text{ and} \\ \tilde{C}^2 u &= \frac{1}{2}\int_0^t d\tau \left[\tilde{C}h(\tau, e^{(t-\tau)C}(x)) - \tilde{C}h(\tau, e^{(\tau-t)C}(x))\right]. \end{aligned}$$

Subtracting the second and last equation then shows $u_{tt} = \tilde{A}^2 u + h$ and it is clear that $u(0,\cdot) = 0$ and $u_t(0,\cdot) = 0$. ■

37.2 General Linear First Order PDE

In this section we consider the following PDE,

$$\sum_{i=1}^n a_i(x)\partial_i u(x) = c(x)u(x) \tag{37.20}$$

where $a_i(x)$ and $c(x)$ are given functions. As above Eq. (37.20) may be written simply as

$$\tilde{A}u(x) = c(x)u(x). \tag{37.21}$$

The key observation to solving Eq. (37.21) is that the chain rule implies

$$\frac{d}{ds}u(e^{sA}(x)) = \tilde{A}u(e^{sA}(x)), \tag{37.22}$$

which we will write briefly as

$$\frac{d}{ds}u \circ e^{sA} = \tilde{A}u \circ e^{sA}.$$

Combining Eqs. (37.21) and (37.22) implies

$$\frac{d}{ds}u(e^{sA}(x)) = c(e^{sA}(x))u(e^{sA}(x))$$

which then gives

$$u(e^{sA}(x)) = e^{\int_0^s c(e^{\sigma A}(x))d\sigma}u(x). \tag{37.23}$$

Equation (37.22) shows that the values of u solving Eq. (37.21) along any flow line of A, are completely determined by the value of u at any point on this flow line. Hence we can expect to construct solutions to Eq. (37.21) by specifying u arbitrarily on any surface Σ which crosses the flow lines of A transversely, see Figure 37.1 below.

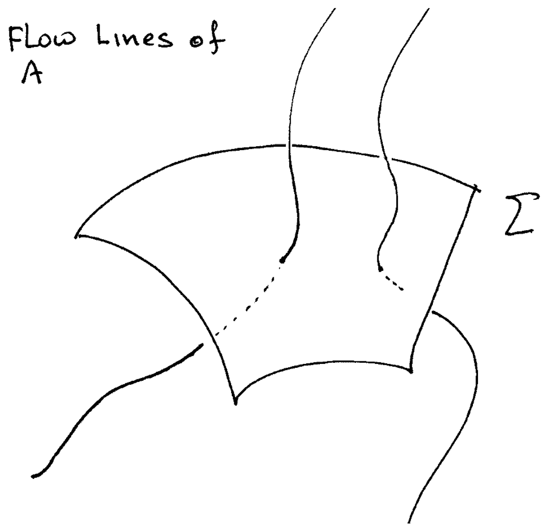


Fig. 37.1. The flow lines of A through a non-characteristic surface Σ.

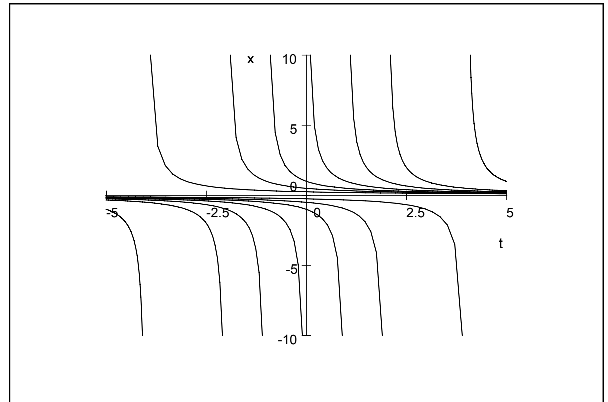
Example 37.13. Let us again consider the PDE in Eq. (37.6) above but now with initial data being given on the line $x = t$, i.e.

$$u_t = x^2 u_x \text{ with } u(\lambda, \lambda) = f(\lambda)$$

for some $f \in C^1(\mathbb{R}, \mathbb{R})$. The characteristic equations are given by

$$t'(s) = 1 \text{ and } x'(s) = -x^2(s) \tag{37.24}$$

and the flow lines of this equations must live on the solution curves to $\frac{dx}{dt} = -x^2$, i.e. on curves of the form $x(t) = \frac{1}{t-C}$ for $C \in \mathbb{R}$ and $x = 0$, see Figure 37.13.



Any solution to $u_t = x^2 u_x$ must be constant on these characteristic curves. Notice the line $x = t$ crosses each characteristic curve exactly once while the line $t = 0$ crosses some but not all of the characteristic curves.

Solving Eqs. (37.24) with $t(0) = \lambda = x(0)$ gives

$$t(s) = s + \lambda \text{ and } x(s) = \frac{\lambda}{1 + s\lambda} \tag{37.25}$$

and hence

$$u(s + \lambda, \frac{\lambda}{1 + s\lambda}) = f(\lambda) \text{ for all } \lambda \text{ and } s > -1/\lambda.$$

(for a plot of some of the integral curves of Eq. (37.24).) Let

$$(t, x) = (s + \lambda, \frac{\lambda}{1 + s\lambda}) \tag{37.26}$$

and solve for λ :

$$x = \frac{\lambda}{1 + (t - \lambda)\lambda} \text{ or } x\lambda^2 - (xt - 1)\lambda - x = 0$$

which gives

$$\lambda = \frac{(xt-1) \pm \sqrt{(xt-1)^2 + 4x^2}}{2x}. \tag{37.27}$$

Now to determine the sign, notice that when $s = 0$ in Eq. (37.26) we have $t = \lambda = x$. So taking $t = x$ in the right side of Eq. (37.27) implies

$$\frac{(x^2-1) \pm \sqrt{(x^2-1)^2 + 4x^2}}{2x} = \frac{(x^2-1) \pm (x^2+1)}{2x} = \begin{cases} x & \text{with } + \\ -2/x & \text{with } - \end{cases}.$$

Therefore, we must take the plus sing in Eq. (37.27) to find

$$\lambda = \frac{(xt-1) + \sqrt{(xt-1)^2 + 4x^2}}{2x}$$

and hence

$$u(t,x) = f\left(\frac{(xt-1) + \sqrt{(xt-1)^2 + 4x^2}}{2x}\right). \tag{37.28}$$

When x is small,

$$\lambda = \frac{(xt-1) + (1-xt)\sqrt{1 + \frac{4x^2}{(xt-1)^2}}}{2x} \cong \frac{(1-xt)\frac{2x^2}{(xt-1)^2}}{2x} = \frac{x}{1-xt}$$

so that

$$u(t,x) \cong f\left(\frac{x}{1-xt}\right) \text{ when } x \text{ is small.}$$

Thus we see that $u(t,0) = f(0)$ and $u(t,x)$ is C^1 if f is C^1. Equation (37.28) sets up a one to one correspondence between solution u to $u_t = x^2 u_x$ and $f \in C^1(\mathbb{R}, \mathbb{R})$.

Example 37.14. To solve

$$xu_x + yu_y = \lambda xyu \text{ with } u = f \text{ on } S^1, \tag{37.29}$$

let $A(x,y) = (x,y) = x\partial_x + y\partial_y$. The equations for $(x(s), y(s)) := e^{sA}(x,y)$ are

$$x'(s) = x(s) \text{ and } y'(s) = y(s)$$

from which we learn

$$e^{sA}(x,y) = e^s(x,y).$$

Then by Eq. (37.23),

$$u(e^s(x,y)) = e^{\lambda \int_0^s e^{2\sigma} xy d\sigma} u(x,y) = e^{\frac{\lambda}{2}(e^{2s}-1)xy} u(x,y).$$

Letting $(x,y) \to e^{-s}(x,y)$ in this equation gives

$$u(x,y) = e^{\frac{\lambda}{2}(1-e^{-2s})xy} u(e^{-s}(x,y))$$

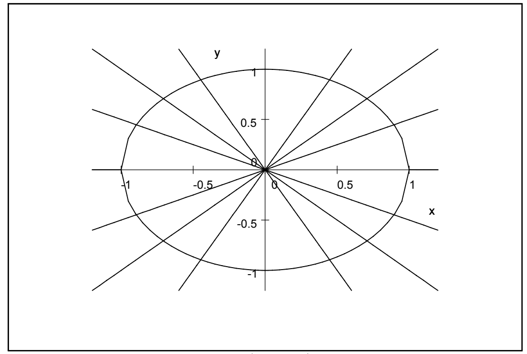
and then choosing s so that

$$1 = \left\|e^{-s}(x,y)\right\|^2 = e^{-2s}(x^2+y^2),$$

i.e. so that $s = \frac{1}{2}\ln\left(x^2+y^2\right)$. We then find

$$u(x,y) = \exp\left(\frac{\lambda}{2}\left(1 - \frac{1}{x^2+y^2}\right)xy\right) f(\frac{(x,y)}{\sqrt{x^2+y^2}}).$$

Notice that this solution always has a singularity at $(x,y) = (0,0)$ unless f is constant.



Characteristic curves for Eq. (37.29) along with the plot of S^1.

Example 37.15. The PDE,

$$e^x u_x + u_y = u \text{ with } u(x,0) = g(x), \tag{37.30}$$

has characteristic curves determined by $x' := e^x$ and $y' := 1$ and along these curves solutions u to Eq. (37.30) satisfy

$$\frac{d}{ds}u(x,y) = u(x,y). \tag{37.31}$$

Solving these "characteristic equations" gives

$$-e^{-x(s)} + e^{-x_0} = \int_0^s e^{-x}x'ds = \int_0^s 1ds = s \tag{37.32}$$

so that

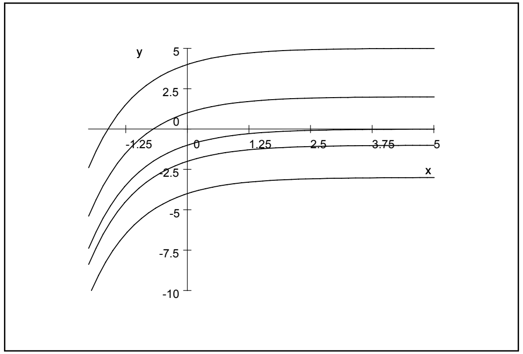
$$x(s) = -\ln(e^{-x_0} - s) \text{ and } y(s) = y_0 + s. \tag{37.33}$$

From Eqs. (37.32) and (37.33) one shows

$$y(s) = y_0 + e^{-x_0} - e^{-x(s)}$$

so the "characteristic curves" are contained in the graphs of the functions

$$y = C - e^{-x} \text{ for some constant } C.$$



Some characteristic curves for Eq. (37.30). Notice that the line $y = 0$ intersects some but not all of the characteristic curves. Therefore Eq. (37.30) does not uniquely determine a function u defined on all of $\mathbb{R}^2$. On the otherhand if the intial condition were $u(0, y) = g(y)$ the method would produce an everywhere defined solution.

Since the initial condition is at $y = 0$, set $y_0 = 0$ in Eq. (37.33) and notice from Eq.(37.31) that

$$u(-\ln(e^{-x_0} - s), s) = u(x(s), y(s)) = e^s u(x_0, 0) = e^s g(x_0). \tag{37.34}$$

Setting $(x, y) = (-\ln(e^{-x_0} - s), s)$ and solving for (x_0, s) implies

$$s = y \text{ and } x_0 = -\ln(e^{-x} + y)$$

and using this in Eq. (37.34) then implies

$$u(x, y) = e^y g\left(-\ln(y + e^{-x})\right).$$

This solution is only defined for $y > -e^{-x}$.

Example 37.16. In this example we will use the method of characteristics to solve the following non-linear equation,

$$x^2 u_x + y^2 \ u_y = u^2 \text{ with } u := 1 \text{ on } y = 2x. \tag{37.35}$$

As usual let (x, y) solve the characteristic equations, $x' = x^2$ and $y' = y^2$ so that

$$(x(s), y(s)) = \left(\frac{x_0}{1 - sx_0}, \frac{y_0}{1 - sy_0}\right).$$

Now let $(x_0, y_0) = (\lambda, 2\lambda)$ be a point on line $y = 2x$ and supposing u solves Eq. (37.35). Then $z(s) = u(x(s), y(s))$ solves

$$z' = \frac{d}{ds} u\left(x, y\right) = x^2 u_x + y^2 \ u_y = u^2(x, y) = z^2$$

with $z(0) = u(\lambda, 2\lambda) = 1$ and hence

$$u\left(\frac{\lambda}{1 - s\lambda}, \frac{2\lambda}{1 - 2s\lambda}\right) = u(x(s), y(s)) = z(s) = \frac{1}{1 - s}. \tag{37.36}$$

Let

$$(x, y) = \left(\frac{\lambda}{1 - s\lambda}, \frac{2\lambda}{1 - 2s\lambda}\right) = \left(\frac{1}{\lambda^{-1} - s}, \frac{1}{\lambda^{-1}/2 - s}\right) \tag{37.37}$$

and solve the resulting equations:

$$\lambda^{-1} - s = x^{-1} \text{ and } \lambda^{-1}/2 - s = y^{-1}$$

for s gives $s = x^{-1} - 2y^{-1}$ and hence

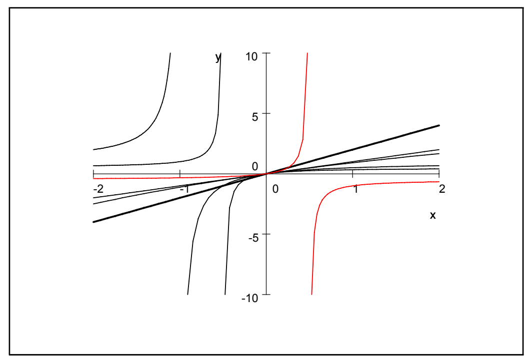
$$1 - s = 1 + 2y^{-1} - x^{-1} = x^{-1} y^{-1} \left(xy + 2x - y\right). \tag{37.38}$$

Combining Eqs. (37.36) – (37.38) gives

$$u(x, y) = \frac{xy}{xy + 2x - y}.$$

Notice that the characteristic curves here lie on the trajectories determined by $\frac{dx}{x^2} = \frac{dy}{y^2}$, i.e. $y^{-1} = x^{-1} + C$ or equivalently

$$y = \frac{x}{1 + Cx}$$



Some characteristic curves

37.3 Quasi-Linear Equations

In this section we consider the following PDE,

$$A(x,z)\cdot\nabla_x u(t,x)=\sum_{i=1}^{n}a_i(x,u(x))\partial_i u(x)=c(x,u(x)) \tag{37.39}$$

where $a_i(x,z)$ and $c(x,z)$ are given functions on $(x,z)\in\mathbb{R}^n\times\mathbb{R}$ and $A(x,z):=(a_1(x,z),\ldots,a_1(x,z))$. Assume u is a solution to Eq. (37.39) and suppose $x(s)$ solves $x'(s)=A(x(s),u(x(s)))$.Then from Eq. (37.39) we find

$$\frac{d}{ds}u(x(s))=\sum_{i=1}^{n}a_i(x(s),u(x(s)))\partial_i u(x(s))=c(x(s),u(x(s))),$$

see Figure 37.2 below. We have proved the following Lemma.

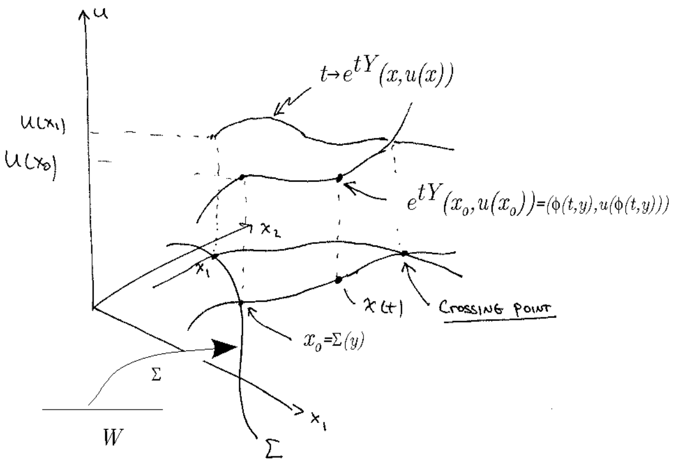


Fig. 37.2. Determining the values of u by solving ODE's. Notice that potential problem though where the projection of characteristics cross in x – space.

Lemma 37.17. *Let $w=(x,z)$, $\pi_1(w)=x$, $\pi_2(w)=z$ and $Y(w)=(A(x,z),c(x,z))$. If u is a solution to Eq. (37.39), then*

$$u(\pi_1\circ e^{sY}(x_0,u(x_0))=\pi_2\circ e^{sY}(x_0,u(x_0)).$$

Let Σ be a surface in $\mathbb{R}^n$ (x– space), i.e. $\Sigma:U\subset_o\mathbb{R}^{n-1}\longrightarrow\mathbb{R}^n$ such that $\Sigma(0)=x_0$ and $D\Sigma(y)$ is injective for all $y\in U$. Now suppose $u_0:\Sigma\to\mathbb{R}$ is given we wish to solve for u such that (37.39) holds and $u=u_0$ on Σ. Let

$$\phi(s,y):=\pi_1\circ e^{sY}\left(\Sigma(y),u_0(\Sigma(y))\right) \tag{37.40}$$

then

$$\frac{\partial\phi}{\partial s}(0,0)=\pi_1\circ Y(x_0,u_0(x_0))=A(x_0,u_0(x_0)) \text{ and}$$
$$D_y\phi(0,0)=D_y\Sigma(0).$$

Assume Σ is **non-characteristic at** x_0, that is $A(x_0,u_0(x_0))\notin\operatorname{Ran}\Sigma'(0)$ where $\Sigma'(0):\mathbb{R}^{n-1}\to\mathbb{R}^n$ is defined by

$$\Sigma'(0)v=\partial_v\Sigma(0)=\frac{d}{ds}|_0\Sigma(sv) \text{ for all } v\in\mathbb{R}^{n-1}.$$

Then $\left(\frac{\partial\phi}{\partial s},\frac{\partial\phi}{\partial y^1},\ldots,\frac{\partial\phi}{\partial y^{n-1}}\right)$ are all linearly independent vectors at $(0,0)\in\mathbb{R}\times\mathbb{R}^{n-1}$. So $\phi:\mathbb{R}\times\mathbb{R}^{n-1}\to\mathbb{R}^n$ has an invertible differential at $(0,0)$ and so the inverse function theorem gives the existence of open neighborhood $0\in W\subset U$ and $0\in J\subset\mathbb{R}$ such that $\phi|_{J\times W}$ is a homeomorphism onto an open set $V:=\phi(J\times W)\subset\mathbb{R}^n$, see Figure 37.3. Because of Lemma 37.17, if

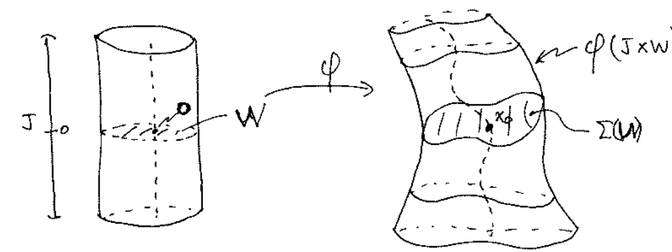


Fig. 37.3. Constructing a neighborhood of the surface Σ near x_0 where we can solve the quasi-linear PDE.

we are going to have a C^1 – solution u to Eq. (37.39) with $u=u_0$ on Σ it would have to satisfy

$$u(x)=\pi_2\circ e^{sY}\left(\Sigma(y),u_0(\Sigma(y))\right) \text{ with } (s,y):=\phi^{-1}(x), \tag{37.41}$$

i.e. $x=\phi(s,y)$.

Proposition 37.18. *The function u in Eq. (37.41) solves Eq. (37.39) on V with $u=u_0$ on Σ.*

Proof. By definition of u in Eq. (37.41) and ϕ in Eq. (37.40),

$$\phi'(s,y) = \pi_1 Y \circ e^{sY}\left(\Sigma(y), u_0(\Sigma(y))\right) = A(\phi(s,y)), u(\phi(s,y))$$

and

$$\frac{d}{ds}u(\phi(s,y)) = \pi_2 Y\left(\phi(s,y), u(\phi(s,y))\right) = c(\phi(s,y), u(\phi(s,y))). \quad (37.42)$$

On the other hand by the chain rule,

$$\begin{aligned}\frac{d}{ds}u(\phi(s,y)) &= \nabla u(\phi(s,u)) \cdot \phi'(s,y) \\ &= \nabla u(\phi(s,u)) \cdot A(\phi(s,y)), u(\phi(s,y)). \quad (37.43)\end{aligned}$$

Comparing Eqs. (37.42) and (37.43) implies

$$\nabla u(\phi(s,y)) \cdot A(\phi(s,y), u(\phi(s,y))) = c(\phi(s,y), u(\phi(s,y))).$$

Since $\phi(J \times W) = V$, u solves Eq. (37.39) on V. Clearly $u(\phi(0,y)) = u_0(\Sigma(y))$ so $u = u_0$ on Σ. ■

Example 37.19 (Conservation Laws). Let $F : \mathbb{R} \to \mathbb{R}$ be a smooth function, we wish to consider the PDE for $u = u(t,x)$,

$$0 = u_t + \partial_x F(u) = u_t + F'(u)u_x \text{ with } u(0,x) = g(x). \quad (37.44)$$

The characteristic equations are given by

$$t'(s) = 1, \quad x'(s) = F'(z(s)) \text{ and } \frac{d}{ds}z(s) = 0. \quad (37.45)$$

The solution to Eqs. (37.45) with $t(0) = 0$, $x(0) = x$ and hence

$$z(0) = u(t(0), x(0)) = u(0,x) = g(x),$$

are given by

$$t(s) = s, \quad z(s) = g(x) \text{ and } x(s) = x + sF'(g(x)).$$

So we conclude that any solution to Eq. (37.44) must satisfy,

$$u(s, x + sF'(g(x)) = g(x).$$

This implies, letting $\psi_s(x) := x + sF'(g(x))$, that

$$u(t,x) = g(\psi_t^{-1}(x)).$$

In order to find ψ_t^{-1} we need to know ψ_t is invertible, i.e. that ψ_t is monotonic in x. This becomes the condition

$$0 < \psi_t'(x) = 1 + tF''(g(x))g'(x).$$

If this holds then we will have a solution.

Example 37.20 (Conservation Laws in Higher Dimensions). Let $F : \mathbb{R} \to \mathbb{R}^n$ be a smooth function, we wish to consider the PDE for $u = u(t,x)$,

$$0 = u_t + \nabla \cdot F(u) = u_t + F'(u) \cdot \nabla u \text{ with } u(0,x) = g(x). \quad (37.46)$$

The characteristic equations are given by

$$t'(s) = 1, \quad x'(s) = F'(z(s)) \text{ and } \frac{d}{ds}z(s) = 0. \quad (37.47)$$

The solution to Eqs. (37.47) with $t(0) = 0$, $x(0) = x$ and hence

$$z(0) = u(t(0), x(0)) = u(0,x) = g(x),$$

are given by

$$t(s) = s, \quad z(s) = g(x) \text{ and } x(s) = x + sF'(g(x)).$$

So we conclude that any solution to Eq. (37.46) must satisfy,

$$u(s, x + sF'(g(x))) = g(x). \quad (37.48)$$

This implies, letting $\psi_s(x) := x + sF'(g(x))$, that

$$u(t,x) = g(\psi_t^{-1}(x)).$$

In order to find ψ_t^{-1} we need to know ψ_t is invertible. Locally by the implicit function theorem it suffices to know,

$$\psi_t'(x)v = v + tF''(g(x))\partial_v g(x) = \left[I + tF''(g(x))\nabla g(x) \cdot\right] v$$

is invertible. Alternatively, let $y = x + sF'(g(x))$, (so $x = y - sF'(g(x))$) in Eq. (37.48) to learn, using Eq. (37.48) which asserts $g(x) = u(s, x + sF'(g(x))) = u(s,y)$,

$$u(s,y) = g\left(y - sF'(g(x))\right) = g\left(y - sF'(u(s,y))\right).$$

This equation describes the solution u implicitly.

Example 37.21 (Burger's Equation). Recall Burger's equation is the PDE,

$$u_t + uu_x = 0 \text{ with } u(0,x) = g(x) \quad (37.49)$$

where g is a given function. Also recall that if we view $u(t,x)$ as a time dependent vector field on $\mathbb{R}$ and let $x(t)$ solve

$$\dot{x}(t) = u(t, x(t)),$$

then

$$\ddot{x}(t) = u_t + u_x \dot{x} = u_t + u_x u = 0.$$

Therefore x has constant acceleration and

$$x(t) = x(0) + \dot{x}(0)t = x(0) + g(x(0))t.$$

This equation contains the same information as the characteristic equations. Indeed, the characteristic equations are

$$t'(s) = 1, \quad x'(s) = z(s) \text{ and } z'(s) = 0. \tag{37.50}$$

Taking initial condition $t(0) = 0$, $x(0) = x_0$ and $z(0) = u(0, x_0) = g(x_0)$ we find

$$t(s) = s, \quad z(s) = g(x_0) \text{ and } x(s) = x_0 + sg(x_0).$$

According to Proposition 37.18, we must have

$$u((s, x_0 + sg(x_0)) = u(s, x(s)) = u(0, x(0)) = g(x_0). \tag{37.51}$$

Letting $\psi_t(x_0) := x_0 + tg(x_0)$, "the" solution to $(t, x) = (s, x_0 + sg(x_0))$ is given by $s = t$ and $x_0 = \psi_t^{-1}(x)$. Therefore, we find from Eq. (37.51) that

$$u(t, x) = g\left(\psi_t^{-1}(x)\right). \tag{37.52}$$

This gives the desired solution provided ψ_t^{-1} is well defined.

Example 37.22 (Burger's Equation Continued). Continuing Example 37.21. Suppose that $g \geq 0$ is an increasing function (i.e. the faster cars start to the right), then ψ_t is strictly increasing and for any $t \geq 0$ and therefore Eq. (37.52) gives a solution for all $t \geq 0$. For a specific example take $g(x) = \max(x, 0)$, then

$$\psi_t(x) = \begin{cases} (1+t)x & \text{if } x \geq 0 \\ x & \text{if } x \leq 0 \end{cases}$$

and therefore,

$$\psi_t^{-1}(x) = \begin{cases} (1+t)^{-1}x & \text{if } x \geq 0 \\ x & \text{if } x \leq 0 \end{cases}$$

$$u(t, x) = g\left(\psi_t^{-1}(x)\right) = \begin{cases} (1+t)^{-1}x & \text{if } x \geq 0 \\ 0 & \text{if } x \leq 0. \end{cases}$$

Notice that $u(t, x) \to 0$ as $t \to \infty$ since all the fast cars move off to the right leaving only slower and slower cars passing $x \in \mathbb{R}$.

Example 37.23. Now suppose $g \geq 0$ and that $g'(x_0) < 0$ at some point $x_0 \in \mathbb{R}$, i.e. there are faster cars to the left of x_0 then there are to the right of x_0, see Figure 37.4. Without loss of generality we may assume that $x_0 = 0$. The projection of a number of characteristics to the (t, x) plane for this velocity profile are given in Figure 37.5 below. Since any C^2 – solution to Eq.(37.49) must be constant on these lines with the value given by the slope, it is impossible to get a C^2 – solution on all of $\mathbb{R}^2$ with this initial condition. Physically, there are collisions taking place which causes the formation of a shock wave in the system.

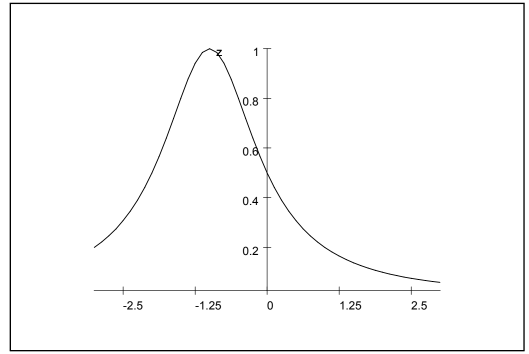


Fig. 37.4. An intial velocity profile where collisions are going to occur. This is the graph of $g(x) = 1/\left(1 + (x+1)^2\right)$.

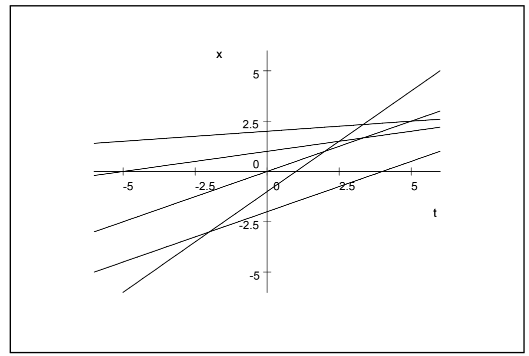


Fig. 37.5. Crossing of projected characteristics for Burger's equation.

37.4 Distribution Solutions for Conservation Laws

Let us again consider the conservation law in Example 37.19 above. We will now restrict our attention to non-negative times. Suppose that u is a C^1 – solution to

$$u_t + (F(u))_x = 0 \text{ with } u(0, x) = g(x) \tag{37.53}$$

and $\phi \in C_c^2([0, \infty) \times \mathbb{R})$. Then by integration by parts,

$$0 = -\int_{\mathbb{R}} dx \int_{t\geq 0} dt(u_t + F(u)_x)\varphi$$
$$= -\int_{\mathbb{R}} [u\varphi]\Big|_{t=0}^{t=\infty} dx + \int_{\mathbb{R}} dx \int_{t\geq 0} dt(u\varphi_t + F(u)\phi_x)$$
$$= \int_{\mathbb{R}} g(x)\varphi(0,x)dx + \int_{\mathbb{R}} dx \int_{t\geq 0} dt(u(t,x)\phi_t(t,x) + F(u(t,x))\phi_x(t,x)).$$

Definition 37.24. *A bounded measurable function $u(t,x)$ is a **distributional solution** to Eq. (37.53) iff*

$$0 = \int_{\mathbb{R}} g(x)\varphi(0,x)dx + \int_{\mathbb{R}} dx \int_{t\geq 0} dt(u(t,x)\phi_t(t,x) + F(u(t,x))\phi_x(t,x))$$

for all test functions $\phi \in C_c^2(D)$ where $D = [0,\infty) \times \mathbb{R}$.

Proposition 37.25. *If u is a distributional solution of Eq. (37.53) and u is C^1 then u is a solution in the classical sense. More generally if $u \in C^1(R)$ where R is an open region contained in $D^0 := (0,\infty) \times \mathbb{R}$ and*

$$\int_{\mathbb{R}} dx \int_{t\geq 0} dt(u(t,x)\phi_t(t,x) + F(u(t,x))\phi_x(t,x)) = 0 \tag{37.54}$$

for all $\phi \in C_c^2(R)$ then $u_t + (F(u))_x := 0$ on R.

Proof. Undo the integration by parts argument to show Eq. (37.54) implies

$$\int_R (u_t + (F(u))_x)\varphi dx dt = 0$$

for all $\phi \in C_c^1(R)$. This then implies $u_t + (F(u))_x = 0$ on R. ■

Theorem 37.26 (Rankine-Hugoniot Condition). *Let R be a region in D^0 and $x = c(t)$ for $t \in [a,b]$ be a C^1 curve in R as pictured below in Figure 37.6.*

Suppose $u \in C^1(R \setminus c([a,b]))$ and u is bounded and has limiting values u^+ and u^- on $x = c(t)$ when approaching from above and below respectively. Then u is a distributional solution of $u_t + (F(u))_x = 0$ in R if and only if

$$u_t + \frac{\partial}{\partial x}F(u) := 0 \text{ on } R \setminus c([a,b]) \tag{37.55}$$

and for all $t \in [a,b]$,

$$\dot{c}(t)\left[u^+(t,c(t)) - u^-(t,c(t))\right] = F(u^+(t,c(t))) - F(u^-(t,c(t)). \tag{37.56}$$

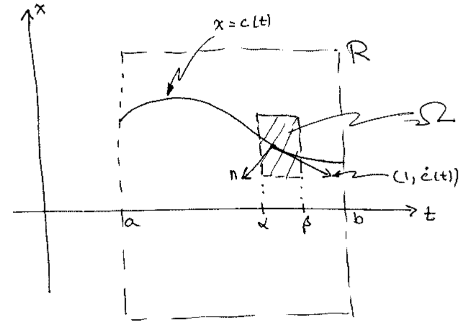


Fig. 37.6. A curve of discontinuities of u.

Proof. The fact that Equation 37.55 holds has already been proved in the previous proposition. For (37.56) let Ω be a region as pictured in Figure 37.6 above and $\phi \in C_c^1(\Omega)$. Then

$$0 = \int_\Omega (u\,\phi_t + F(u)\phi_x)dt\,dx$$
$$= \int_{\Omega_+} (u\phi_t + F(u)\phi_x)dt\,dx + \int_{\Omega_-} (u\phi_t + F(u)\phi_x)dt\,dx \tag{37.57}$$

where

$$\Omega_\pm = \left\{(t,x) \in \Omega : \begin{matrix} x > c(t) \\ x < c(t) \end{matrix}\right\}.$$

Now the outward normal to $\Omega_\pm$ along c is

$$n(t) = \pm\frac{(\dot{c}(t), -1)}{\sqrt{1 + \dot{c}(t)^2}}$$

and the "surface measure" along c is given by $d\sigma(t) = \sqrt{1 + \dot{c}(t)^2}dt$. Therefore

$$n(t)\;d\sigma(t) = \pm(\dot{c}(t), -1)dt$$

where the sign is chosen according to the sign in $\Omega_\pm$. Hence by the divergence theorem,

$$\begin{aligned}
\int_{\Omega_\pm} (u\,\phi_t + F(u)\phi_x)dt\;dx
&= \int_{\Omega_\pm} (u, F(u)) \cdot (\phi_t, \phi_x) dt\;dx \\
&= \int_{\partial\Omega_\pm} \phi(u, F(u)) \cdot n(t)\; d\sigma(t) \\
&= \pm \int_\alpha^\beta \phi(t, c(t))(u_t^\pm(t, c(t))\dot{c}(t) - F(u_t^\pm(t, c(t))))dt.
\end{aligned}$$

Putting these results into Eq. (37.57) gives

$$\begin{aligned}
0 = \int_\alpha^\beta &\{\dot{c}(t)\left[u^+(t, c(t)) - u^-(t, c(t))\right] \\
&- (F(u^+(t, c(t))) - F(u^-(t, c(t)))\}\phi(t, c(t))dt
\end{aligned}$$

for all ϕ which implies

$$\dot{c}(t)\left[u^+(t, c(t)) - u^-(t, c(t))\right] = F(u^+(t, c(t))) - F(u^-(t, c(t)).$$

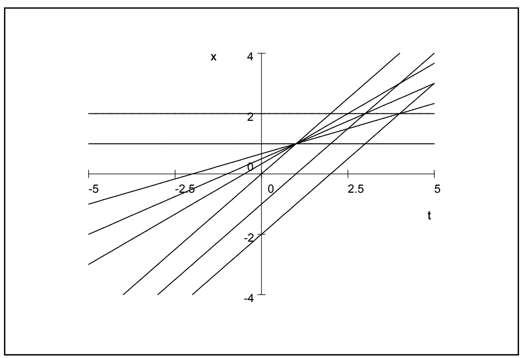
■

Example 37.27. In this example we will find an integral solution to Burger's Equation, $u_t + uu_x = 0$ with initial condition

$$u(0, x) = \begin{cases} 0 & x \geq 1 \\ 1 - x & 0 \leq x \leq 1 \\ 1 & x \leq 0. \end{cases}$$

The characteristics are given from above by

$$\begin{aligned}
x(t) &= (1 - x_0)t + x_0 \quad x_0 \in (0, 1) \\
x(t) &= x_0 + t \text{ if } x_0 \leq 0 \text{ and} \\
x(t) &= x_0 \text{ if } x_0 \geq 1.
\end{aligned}$$



Projected characteristics

For the region bounded determined by $t \leq x \leq 1$ and $t \leq 1$ we have $u(t, x)$ is equal to the slope of the line through (t, x) and $(1, 1)$, i.e.

$$u(t, x) = \frac{x - 1}{t - 1}.$$

Notice that the solution is not well define in the region where characteristics cross, i.e. in the shock zone,

$$R_2 := \{(t, x) : t \geq 1,\; x \geq 1 \text{ and } x \leq t\},$$

see Figure 37.7. Let us now look for a distributional solution of the equation

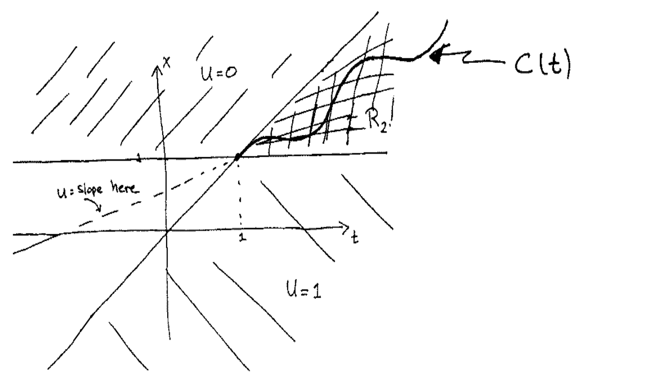


Fig. 37.7. The schock zone and the values of u away from the shock zone.

valid for all (x, t) by looking for a curve $c(t)$ in R_2 such that above $c(t)$, $u = 0$ while below $c(t)$, $u = 1$.

To this end we will employ the Rankine-Hugoniot Condition of Theorem 37.26. To do this observe that Burger's Equation may be written as $u_t + (F(u))_x = 0$ where $F(u) = \frac{u^2}{2}$. So the Jump condition is

$$\dot{c}(u_+ - u_-) = (F(u_+) - F(u_-))$$

that is

$$(0 - 1)\dot{c} = \left(\frac{0^2}{2} - \frac{1^2}{2}\right) = -\frac{1}{2}.$$

Hence $\dot{c}(t) = \frac{1}{2}$ and therefore $c(t) = \frac{1}{2}t + 1$ for $t \geq 0$. So we find a distributional solution given by the values in shown in Figure 37.8.

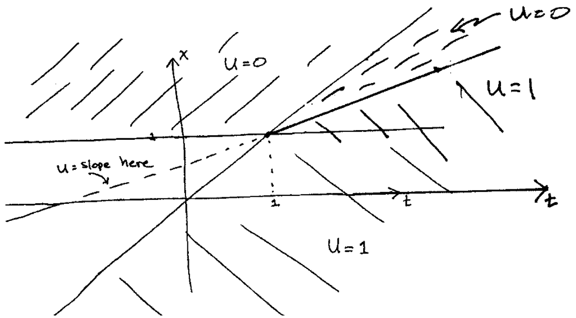


Fig. 37.8. A distributional solution to Burger's equation.

37.5 Exercises

Exercise 37.28. For $A \in L(X)$, let

$$e^{tA} := \sum_{n=0}^{\infty} \frac{t^n}{n!} A^n. \tag{37.58}$$

Show directly that:

1. e^{tA} is convergent in $L(X)$ when equipped with the operator norm.
2. e^{tA} is differentiable in t and that $\frac{d}{dt} e^{tA} = Ae^{tA}$.

Exercise 37.29. Suppose that $A \in L(X)$ and $v \in X$ is an eigenvector of A with eigenvalue λ, i.e. that $Av = \lambda v$. Show $e^{tA}v = e^{t\lambda}v$. Also show that $X = \mathbb{R}^n$ and A is a diagonalizable $n \times n$ matrix with

$$A = SDS^{-1} \text{ with } D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$$

then $e^{tA} = Se^{tD}S^{-1}$ where $e^{tD} = \operatorname{diag}(e^{t\lambda_1}, \ldots, e^{t\lambda_n})$.

Exercise 37.30. Suppose that $A, B \in L(X)$ and $[A, B] := AB - BA = 0$. Show that $e^{(A+B)} = e^A e^B$.

Exercise 37.31. Suppose $A \in C(\mathbb{R}, L(X))$ satisfies $[A(t), A(s)] = 0$ for all $s, t \in \mathbb{R}$. Show

$$y(t) := e^{\left(\int_0^t A(\tau)d\tau\right)} x$$

is the unique solution to $\dot{y}(t) = A(t)y(t)$ with $y(0) = x$.

Exercise 37.32. Compute e^{tA} when

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

and use the result to prove the formula

$$\cos(s + t) = \cos s \cos t - \sin s \sin t.$$

Hint: Sum the series and use $e^{tA}e^{sA} = e^{(t+s)A}$.

Exercise 37.33. Compute e^{tA} when

$$A = \begin{pmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{pmatrix}$$

with $a, b, c \in \mathbb{R}$. Use your result to compute $e^{t(\lambda I + A)}$ where $\lambda \in \mathbb{R}$ and I is the 3×3 identity matrix. **Hint:** Sum the series.

Theorem 37.34. *Suppose that $T_t \in L(X)$ for $t \geq 0$ satisfies*

1. *(Semi-group property.) $T_0 = Id_X$ and $T_t T_s = T_{t+s}$ for all $s, t \geq 0$.*
2. *(Norm Continuity) $t \to T_t$ is continuous at 0, i.e. $\|T_t - I\|_{L(X)} \to 0$ as $t \downarrow 0$.*

Then there exists $A \in L(X)$ such that $T_t = e^{tA}$ where e^{tA} is defined in Eq. (37.58).

Exercise 37.35. Prove Theorem 37.34 using the following outline.

1. First show $t \in [0, \infty) \to T_t \in L(X)$ is continuous.
2. For $\epsilon > 0$, let $S_\epsilon := \frac{1}{\epsilon}\int_0^\epsilon T_\tau d\tau \in L(X)$. Show $S_\epsilon \to I$ as $\epsilon \downarrow 0$ and conclude from this that S_ϵ is invertible when $\epsilon > 0$ is sufficiently small. For the remainder of the proof fix such a small $\epsilon > 0$.
3. Show
$$T_t S_\epsilon = \frac{1}{\epsilon} \int_t^{t+\epsilon} T_\tau d\tau$$
and conclude from this that
$$\lim_{t \downarrow 0} t^{-1} (T_t - I) S_\epsilon = \frac{1}{\epsilon} (T_\epsilon - Id_X).$$
4. Using the fact that S_ϵ is invertible, conclude $A = \lim_{t \downarrow 0} t^{-1}(T_t - I)$ exists in $L(X)$ and that
$$A = \frac{1}{\epsilon} (T_\epsilon - I) S_\epsilon^{-1}.$$
5. Now show using the semigroup property and step 4. that $\frac{d}{dt} T_t = AT_t$ for all $t > 0$.
6. Using step 5, show $\frac{d}{dt} e^{-tA} T_t = 0$ for all $t > 0$ and therefore $e^{-tA}T_t = e^{-0A}T_0 = I$.

Exercise 37.36 (Higher Order ODE). Let X be a Banach space, , $\mathcal{U} \subset_o X^n$ and $f \in C(J \times \mathcal{U}, X)$ be a Locally Lipschitz function in $\mathbf{x} = (x_1, \ldots, x_n)$. Show the n^{th} ordinary differential equation,

$$y^{(n)}(t) = f(t, y(t), \dot{y}(t), \ldots y^{(n-1)}(t))$$
$$\text{with } y^{(k)}(0) = y_0^k \text{ for } k = 0, 1, 2 \ldots, n-1 \tag{37.59}$$

where $(y_0^0, \ldots, y_0^{n-1})$ is given in $\mathcal{U}$, has a unique solution for small $t \in J$. **Hint**: let $\mathbf{y}(t) = \left(y(t), \dot{y}(t), \ldots y^{(n-1)}(t)\right)$ and rewrite Eq. (37.59) as a first order ODE of the form

$$\dot{\mathbf{y}}(t) = Z(t, \mathbf{y}(t)) \text{ with } \mathbf{y}(0) = (y_0^0, \ldots, y_0^{n-1}).$$

Exercise 37.37. Use the results of Exercises 37.33 and 37.36 to solve

$$\ddot{y}(t) - 2\dot{y}(t) + y(t) = 0 \text{ with } y(0) = a \text{ and } \dot{y}(0) = b.$$

Hint: The 2×2 matrix associated to this system, A, has only one eigenvalue 1 and may be written as $A = I + B$ where $B^2 = 0$.

Exercise 37.38. Suppose that $A : \mathbb{R} \to L(X)$ is a continuous function and $U, V : \mathbb{R} \to L(X)$ are the unique solution to the linear differential equations

$$\dot{V}(t) = A(t)V(t) \text{ with } V(0) = I \tag{37.60}$$

and

$$\dot{U}(t) = -U(t)A(t) \text{ with } U(0) = I. \tag{37.61}$$

Prove that $V(t)$ is invertible and that $V^{-1}(t) = U(t)$. **Hint:** 1) show $\frac{d}{dt}[U(t)V(t)] = 0$ (which is sufficient if $\dim(X) < \infty$) and 2) show compute $y(t) := V(t)U(t)$ solves a linear differential ordinary differential equation that has $y \equiv 0$ as an obvious solution. Then use the uniqueness of solutions to ODEs. (The fact that $U(t)$ must be defined as in Eq. (37.61) is the content of Exercise 19.32 in the analysis notes.)

Exercise 37.39 (Duhamel' s Principle I). Suppose that $A : \mathbb{R} \to L(X)$ is a continuous function and $V : \mathbb{R} \to L(X)$ is the unique solution to the linear differential equation in Eq. (37.60). Let $x \in X$ and $h \in C(\mathbb{R}, X)$ be given. Show that the unique solution to the differential equation:

$$\dot{y}(t) = A(t)y(t) + h(t) \text{ with } y(0) = x \tag{37.62}$$

is given by

$$y(t) = V(t)x + V(t) \int_0^t V(\tau)^{-1} h(\tau)\, d\tau. \tag{37.63}$$

Hint: compute $\frac{d}{dt}[V^{-1}(t)y(t)]$ when y solves Eq. (37.62).

Exercise 37.40 (Duhamel' s Principle II). Suppose that $A : \mathbb{R} \to L(X)$ is a continuous function and $V : \mathbb{R} \to L(X)$ is the unique solution to the linear differential equation in Eq. (37.60). Let $W_0 \in L(X)$ and $H \in C(\mathbb{R}, L(X))$ be given. Show that the unique solution to the differential equation:

$$\dot{W}(t) = A(t)W(t) + H(t) \text{ with } W(0) = W_0 \tag{37.64}$$

is given by

$$W(t) = V(t)W_0 + V(t) \int_0^t V(\tau)^{-1} H(\tau)\, d\tau. \tag{37.65}$$

Exercise 37.41 (Non-Homogeneous ODE). Suppose that $U \subset_o X$ is open and $Z : \mathbb{R} \times U \to X$ is a continuous function. Let $J = (a, b)$ be an interval and $t_0 \in J$. Suppose that $y \in C^1(J, U)$ is a solution to the "non-homogeneous" differential equation:

$$\dot{y}(t) = Z(t, y(t)) \text{ with } y(t_o) = x \in U. \tag{37.66}$$

Define $Y \in C^1(J - t_0, \mathbb{R} \times U)$ by $Y(t) := (t + t_0, y(t + t_0))$. Show that Y solves the "homogeneous" differential equation

$$\dot{Y}(t) = \tilde{A}(Y(t)) \text{ with } Y(0) = (t_0, y_0), \tag{37.67}$$

where $\tilde{A}(t, x) := (1, Z(x))$. Conversely, suppose that $Y \in C^1(J - t_0, \mathbb{R} \times U)$ is a solution to Eq. (37.67). Show that $Y(t) = (t + t_0, y(t + t_0))$ for some $y \in C^1(J, U)$ satisfying Eq. (37.66). (In this way the theory of non-homogeneous ode's may be reduced to the theory of homogeneous ode's.)

Exercise 37.42 (Differential Equations with Parameters). Let W be another Banach space, $U \times V \subset_o X \times W$ and $Z \in C(U \times V, X)$ be a locally Lipschitz function on $U \times V$. For each $(x, w) \in U \times V$, let $t \in J_{x,w} \to \phi(t, x, w)$ denote the maximal solution to the ODE

$$\dot{y}(t) = Z(y(t), w) \text{ with } y(0) = x. \tag{37.68}$$

Prove

$$\mathcal{D} := \{(t, x, w) \in \mathbb{R} \times U \times V : t \in J_{x,w}\} \tag{37.69}$$

is open in $\mathbb{R} \times U \times V$ and ϕ and $\dot{\phi}$ are continuous functions on $\mathcal{D}$.

Hint: If $y(t)$ solves the differential equation in (37.68), then $v(t) := (y(t), w)$ solves the differential equation,

$$\dot{v}(t) = \tilde{A}(v(t)) \text{ with } v(0) = (x, w), \tag{37.70}$$

where $\tilde{A}(x, w) := (Z(x, w), 0) \in X \times W$ and let $\psi(t, (x, w)) := v(t)$. Now apply the Theorem 6.21 to the differential equation (37.70).

Exercise 37.43 (Abstract Wave Equation). For $A \in L(X)$ and $t \in \mathbb{R}$, let

$$\cos(tA) := \sum_{n=0}^{\infty} \frac{(-1)^{2n}}{(2n)!} t^{2n} A^{2n} \text{ and}$$
$$\frac{\sin(tA)}{A} := \sum_{n=0}^{\infty} \frac{(-1)^{2n+1}}{(2n+1)!} t^{2n+1} A^{2n}.$$

Show that the unique solution $y \in C^2(\mathbb{R}, X)$ to

$$\ddot{y}(t) + A^2 y(t) = 0 \text{ with } y(0) = y_0 \text{ and } \dot{y}(0) = \dot{y}_0 \in X \tag{37.71}$$

is given by

$$y(t) = \cos(tA) y_0 + \frac{\sin(tA)}{A} \dot{y}_0.$$

Remark 37.44. Exercise 37.43 can be done by direct verification. Alternatively and more instructively, rewrite Eq. (37.71) as a first order ODE using Exercise 37.36. In doing so you will be lead to compute e^{tB} where $B \in L(X \times X)$ is given by

$$B = \begin{pmatrix} 0 & I \\ -A^2 & 0 \end{pmatrix},$$

where we are writing elements of $X \times X$ as column vectors, $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. You should then show

$$e^{tB} = \begin{pmatrix} \cos(tA) & \frac{\sin(tA)}{A} \\ -A\sin(tA) & \cos(tA) \end{pmatrix}$$

where

$$A \sin(tA) := \sum_{n=0}^{\infty} \frac{(-1)^{2n+1}}{(2n+1)!} t^{2n+1} A^{2(n+1)}.$$

Exercise 37.45 (Duhamel's Principle for the Abstract Wave Equation). Continue the notation in Exercise 37.43, but now consider the ODE,

$$\ddot{y}(t) + A^2 y(t) = f(t) \text{ with } y(0) = y_0 \text{ and } \dot{y}(0) = \dot{y}_0 \in X \tag{37.72}$$

where $f \in C(\mathbb{R}, X)$. Show the unique solution to Eq. (37.72) is given by

$$y(t) = \cos(tA) y_0 + \frac{\sin(tA)}{A} \dot{y}_0 + \int_0^t \frac{\sin((t-\tau)A)}{A} f(\tau) d\tau \tag{37.73}$$

Hint: Again this could be proved by direct calculation. However it is more instructive to deduce Eq. (37.73) from Exercise 37.39 and the comments in Remark 37.44.

Exercise 37.46. Number 3 on p. 163 of Evans.

38

Fully nonlinear first order PDE

In this section let $\mathcal{U} \subset_o \mathbb{R}^n$ be an open subset of $\mathbb{R}^n$ and $(x, z, p) \in \bar{\mathcal{U}} \times \mathbb{R}^n \times \mathbb{R} \to F(x, z, p) \in \mathbb{R}$ be a C^2 – function. Actually to simplify notation let us suppose $\mathcal{U} = \mathbb{R}^n$. We are now looking for a solution $u : \mathbb{R}^n \to \mathbb{R}$ to the fully non-linear PDE,

$$F(x, u(x), \nabla u(x)) = 0. \tag{38.1}$$

As above, we "reduce" the problem to solving ODE's. To see how this might be done, suppose u solves (38.1) and $x(s)$ is a curve in $\mathbb{R}^n$ and let

$$z(s) = u(x(s)) \text{ and } p(s) = \nabla u(x(s)).$$

Then

$$z'(s) = \nabla u(x(s)) \cdot x'(s) = p(s) \cdot x'(s) \text{ and} \tag{38.2}$$
$$p'(s) = \partial_{x'(s)} \nabla u(x(s)). \tag{38.3}$$

We would now like to find an equation for $x(s)$ which along with the above system of equations would form and ODE for $(x(s), z(s), p(s))$. The term, $\partial_{x'(s)} \nabla u(x(s))$, which involves two derivative of u is problematic and we would like to replace it by something involving only ∇u and u. In order to get the desired relation, differentiate Eq. (38.1) in x in the direction v to find

$$\begin{aligned} 0 &= F_x \cdot v + F_z \partial_v u + F_p \cdot \partial_v \nabla u = F_x \cdot v + F_z \partial_v u + F_p \cdot \nabla \partial_v u \\ &= F_x \cdot v + F_z \ \nabla u \cdot v + (\partial_{F_p} \nabla u) \cdot v, \end{aligned}$$

wherein we have used the fact that mixed partial derivative commute. This equation is equivalent to

$$\partial_{F_p} \nabla u|_{(x, u(x), \nabla u(x))} = -(F_x + F_z \ \nabla u)|_{(x, u(x), \nabla u(x))}. \tag{38.4}$$

By requiring $x(s)$ to solve $x'(s) = F_p(x(s), z(s), p(s))$, we find, using Eq. (38.4) and Eqs. (38.2) and (38.3) that $(x(s), z(s), p(s))$ solves the **characteristic equations**,

$$\begin{aligned} x'(s) &= F_p(x(s), z(s), p(s)) \\ z'(s) &= p(s) \cdot F_p(x(s), z(s), p(s)) \\ p'(s) &= -F_x(x(s), z(s), p(s)) - F_z(x(s), z(s), p(s)) p(s). \end{aligned}$$

We will in the future simply abbreviate these equations by

$$\begin{aligned} x' &= F_p \\ z' &= p \cdot F_p \\ p' &= -F_x - F_z p. \end{aligned} \tag{38.5}$$

The above considerations have proved the following Lemma.

Lemma 38.1. *Let*

$$A(x, z, p) := (F_p(x, z, p), p \cdot F_p(x, z, p), -F_x(x, z, p) - F_z(x, z, p) p),$$

$$\pi_1(x, z, p) = x \text{ and } \pi_2(x, z, p) = z.$$

If u solves Eq. (38.1) and $x_0 \in U$, then

$$\begin{aligned} e^{sA}(x_0, u(x_0), \nabla u(x_0)) &= (x(s), u(x(s)), \nabla u(x(s))) \text{ and} \\ u(x(s)) &= \pi_2 \circ e^{sA}(x_0, u(x_0), \nabla u(x_0)) \end{aligned} \tag{38.6}$$

where $x(s) = \pi_1 \circ e^{sA}(x_0, u(x_0), \nabla u(x_0))$.

We now want to use Eq. (38.6) to produce solutions to Eq. (38.1). As in the quasi-linear case we will suppose $\Sigma : U \subset_o \mathbb{R}^{n-1} \longrightarrow \mathbb{R}^n$ is a surface, $\Sigma(0) = x_0$, $D\Sigma(y)$ is injective for all $y \in U$ and $u_0 : \Sigma \to \mathbb{R}$ is given. We wish to solve Eq. (38.1) for u with the added condition that $u(\Sigma(y)) = u_0(y)$. In order to make use of Eq. (38.6) to do this, we first need to be able to find $\nabla u(\Sigma(y))$. The idea is to use Eq. (38.1) to determine $\nabla u(\Sigma(y))$ as a function of $\Sigma(y)$ and $u_0(y)$ and for this we will invoke the implicit function theorem. If u is a function such that $u(\Sigma(y)) = u_0(y)$ for y near 0 and $p_0 = \nabla u(x_0)$ then

$$\partial_v u_0(0) = \partial_v u(\Sigma(y))|_{y=0} = \nabla u(x_0) \cdot \Sigma'(0) v = p_0 \cdot \Sigma'(0) v.$$

Notation 38.2 *Let $\nabla_\Sigma u_0(y)$ denote the unique vector in $\mathbb{R}^n$ which is tangential to Σ at $\Sigma(y)$ and such that*

$$\partial_v u_0(y) = \nabla_\Sigma u_0(y) \cdot \Sigma'(0) v \text{ for all } v \in \mathbb{R}^{n-1}.$$

Theorem 38.3. *Let $F : \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ be a C^2 function, $0 \in U \subset_o \mathbb{R}^{n-1}$, $\Sigma : U \subset_o \mathbb{R}^{n-1} \xrightarrow{C^2} \mathbb{R}^n$ be an embedded submanifold, $(x_0, z_0, p_0) \in \Sigma \times \mathbb{R} \times \mathbb{R}^n$ such that $F(x_0, z_0, p_0) = 0$ and $x_0 = \Sigma(0)$, $u_0 : \Sigma \xrightarrow{C^1} \mathbb{R}$ such that $u_0(x_0) = z_0$, $\mathbf{n}(y)$ be a normal vector to Σ at y. Further assume*

1. *$\partial_v u_0(0) = p_0 \cdot \Sigma'(0) v = p_0 \cdot \partial_v \Sigma(0)$ for all $v \in \mathbb{R}^{n-1}$.*

2. $F_p(x_0, y_0, z_0) \cdot \mathbf{n}(0) \neq 0.$

Then there exists a neighborhood $V \subset \mathbb{R}^n$ of x_0 and a C^2-function $u : V \to \mathbb{R}$ such that $u \circ \Sigma = u_0$ near 0 and Eq. (38.1) holds for all $x \in V$.

Proof. Step 1. There exist a neighborhood $U_0 \subset U$ and a function $p_0 : U_0 \to \mathbb{R}^n$ such that

$$p_0(y)^{\tan} = \nabla_\Sigma u_0(y) \text{ and } F(\Sigma(y), u(\Sigma((y)), p_0(y)) = 0 \tag{38.7}$$

for all $y \in U_0$, where $p_0(y)^{\tan}$ is component of $p_0(y)$ tangential to Σ. This is

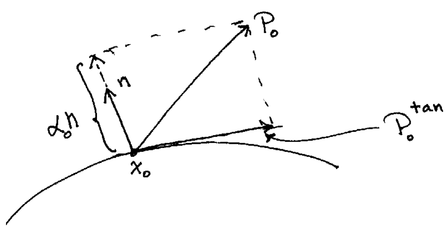


Fig. 38.1. Decomposing p into its normal and tangential components.

an exercise in the implicit function theorem.

Choose $\alpha_0 \in \mathbb{R}$ such that $\nabla_\Sigma u_0(0) + \alpha_0 \mathbf{n}(0) = p_0$ and define

$$f(\alpha, y) := F(\Sigma(y), u_0(y), \nabla_\Sigma u_0(y) + \alpha\, \mathbf{n}(y)).$$

Then

$$\frac{\partial f}{\partial \alpha}(\alpha, 0) = F_p(x_0, z_0, \nabla_\Sigma u_0(0) + \alpha \mathbf{n}(0)) \cdot \mathbf{n}(0) \neq 0,$$

so by the implicit function theorem there exists $0 \in U_0 \subset U$ and $\alpha : U_0 \to \mathbb{R}$ such that $f(\alpha(y), y) = 0$ for all $y \in U_0$. Now define

$$p_0(y) := \nabla_\Sigma u_0(y) + \alpha(y)\mathbf{n}(y) \text{ for } y \in U_0.$$

To simplify notation in the future we will from now on write U for U_0.

Step 2. Suppose (x, z, p) is a solution to (38.5) such that $F(x(0), z(0), p(0)) = 0$ then

$$F(x(s), z(s), p(s)) = 0 \text{ for all } t \in J \tag{38.8}$$

because

$$\begin{aligned}\frac{d}{ds}F(x(s), z(s), p(s)) &= F_x \cdot x' + F_z z' + F_p \cdot p' \\ &= F_x \cdot F_p + F_z(p \cdot F_p) - F_p \cdot (F_x + F_z p) = 0.\end{aligned} \tag{38.9}$$

Step 3. (Notation). For $y \in U$ let

$$(X(s, y), Z(s, y), P(s, y)) = e^{sA}(\Sigma(y), u_0(y), p_0(y)),$$

ie. $X(s, y)$, $Z(s, y)$ and $P(s, y)$ solve the coupled system of O.D.E.'s:

$$\begin{aligned} X' &= F_p \text{ with } X(0, y) = \Sigma(y) \\ Z' &= P \cdot F_p \text{ with } Z(0, y) = u_0(y) \\ P' &= F_x - F_z P \text{ with } P(0, y) = p_0(y). \end{aligned} \tag{38.10}$$

With this notation Eq. (38.8) becomes

$$F(X(s, y), Z(s, y), P(s, y)) = 0 \text{ for all } t \in J. \tag{38.11}$$

Step 4. There exists a neighborhood $0 \in U_0 \subset U$ and $0 \in J \subset \mathbb{R}$ such that $X : J \times U_0 \to \mathbb{R}^n$ is a C^1 diffeomorphism onto an open set $V := X(J \times U_0) \subset \mathbb{R}^n$ with $x_0 \in V$. Indeed, $X(0, y) = \Sigma(y)$ so that $D_y X(0, y)|_{y=0} = \Sigma'(0)$ and hence

$$DX(0, 0)(a, v) = \frac{\partial X}{\partial s}(0, 0)a + \Sigma'(0)v = F_p(x_0, z_0, p_0)a + \Sigma'(0)v.$$

By the assumptions, $F_p(x_0, z_0, p_0) \notin \text{Ran } \Sigma'(0)$ and $\Sigma'(0)$ is injective, it follows that $DX(0, 0)$ is invertible So the assertion is a consequence of the inverse function theorem.

Step 5. Define

$$u(x) := Z(X^{-1}(x)),$$

then u is the desired solution. To prove this first notice that u is uniquely characterized by

$$u(X(s, y)) = Z(s, y) \text{ for all } (s, y) \in J_0 \times U_0.$$

Because of Step 2., to finish the proof it suffices to show $\nabla u(X(s, y)) = P(s, y)$.

Step 6. $\nabla u(X(s, y)) = P(s, y)$. From Eq. (38.10),

$$P \cdot X' = P \cdot F_p = Z' = \frac{d}{ds}u(X) = \nabla u(X) \cdot X' \tag{38.12}$$

which shows

$$[P - \nabla u(X)] \cdot X' = 0.$$

So to finish the proof it suffices to show

$$[P - \nabla u(X)] \cdot \partial_v X = 0$$

for all $v \in \mathbb{R}^{n-1}$ or equivalently that

$$P(s,y) \cdot \partial_v X = \nabla u(x) \cdot \partial_v X = \partial_v u(X) = \partial_v Z \tag{38.13}$$

for all $v \in \mathbb{R}^{n-1}$.

To prove Eq. (38.13), fix a y and let

$$r(s) := P(s,y) \cdot \partial_v X(s,y) - \partial_v Z(s,y).$$

Then using Eq. (38.10),

$$\begin{aligned} r' &= P' \cdot \partial_v X + P \cdot \partial_v X' - \partial_v Z' \\ &= (-F_x - F_z P) \cdot \partial_v X + P \cdot \partial_v F_p - \partial_v (P \cdot F_p) \\ &= (-F_x - F_z P) \cdot \partial_v X - (\partial_v P) \cdot F_p. \end{aligned} \tag{38.14}$$

Further, differentiating Eq. (38.11) in y implies for all $v \in \mathbb{R}^{n-1}$ that

$$F_x \cdot \partial_v X + F_z \partial_v Z + F_p \cdot \partial_v P = 0. \tag{38.15}$$

Adding Eqs. (38.14) and (38.15) then shows

$$r' = -F_z P \cdot \partial_v X + F_z \partial_v Z = -F_z r$$

which implies

$$r(s) = e^{-\int_0^s F_z(X,Z,P)(\sigma,y)d\sigma} r(0).$$

This shows $r \equiv 0$ because $p_0(y)^T = (\nabla_\Sigma u_0)(\Sigma(y))$ and hence

$$\begin{aligned} r(0) &= p_0(y) \cdot \partial_v \Sigma(y) - \partial_v u_0(\Sigma(y)) \\ &= [p_0(y) - \nabla_\Sigma u_0(\Sigma(y))] \cdot \partial_v \Sigma(y) = 0. \end{aligned}$$

■

Example 38.4 (Quasi-Linear Case Revisited). Let us consider the quasi-linear PDE in Eq. (37.39),

$$A(x,z) \cdot \nabla_x u(x) - c(x, u(x)) = 0. \tag{38.16}$$

in light of Theorem 38.3. This may be written in the form of Eq. (38.1) by setting

$$F(x,z,p) = A(x,z) \cdot p - c(x,z).$$

The characteristic equations (38.5) for this F are

$$\begin{aligned} x' &= F_p = A \\ z' &= p \cdot F_p = p \cdot A \\ p' &= -F_x - F_z p = -(A_x \cdot p - c_x) - (A_z \cdot p - c_z)\, p. \end{aligned}$$

Recalling that $p(s) = \nabla u(s,x)$, the z equation above may expressed, by using Eq. (38.16) as

$$z' = p \cdot A = c.$$

Therefore the equations for $(x(s), z(s))$ may be written as

$$x'(s) = A(x,z) \text{ and } z' = c(x,z)$$

and these equations may be solved without regard for the p – equation. This is what makes the quasi-linear PDE simpler than the fully non-linear PDE.

38.1 An Introduction to Hamilton Jacobi Equations

A Hamilton Jacobi Equation is a first order PDE of the form

$$\frac{\partial S}{\partial t}(t,x) + H(x, \nabla_x S(t,x)) = 0 \text{ with } S(0,x) = g(x) \tag{38.17}$$

where $H : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}$ are given functions. In this section we are going to study the connections of this equation to the Euler Lagrange equations of classical mechanics.

38.1.1 Solving the Hamilton Jacobi Equation (38.17) by characteristics

Now let us solve the Hamilton Jacobi Equation (38.17) using the method of characteristics. In order to do this let

$$(p_0, p) = (\frac{\partial S}{\partial t}, \nabla_x S(t,x)) \text{ and } F(t,x,z,p) := p_0 + H(x,p).$$

Then Eq. (38.17) becomes

$$0 = F(t, x, S, \frac{\partial S}{\partial t}, \nabla_x S).$$

Hence the characteristic equations are given by

$$\frac{d}{ds}(t(s), x(s)) = F_{(p_0,p)} = (1, \nabla_p H(x(s), p(s))$$

$$\frac{d}{ds}(p_0, p)(s) = -F_{(t,x)} - F_z(p_0,p) = -F_{(t,x)} = (0, -\nabla_x H(x(s), p(s)))$$

and

$$z'(s) = (p_0, p) \cdot F_{(p_0,p)} = p_0(s) + p(s) \cdot \nabla_p H(x(s), p(s)).$$

Solving the t equation with $t(0) = 0$ gives $t = s$ and so we identify t and s and our equations become

$$\dot{x}(t) = \nabla_p H(x(t), p(t)) \tag{38.18}$$
$$\dot{p}(t) = -\nabla_{\mathbf{x}} H(x(t), p(t)) \tag{38.19}$$

$$\frac{d}{dt}\left[\frac{\partial S}{\partial t}(t, x(t))\right] = \frac{d}{dt}p_0(t) = 0 \text{ and}$$
$$\frac{d}{dt}[S(t, x(t))] = \frac{d}{dt}z(t) = \frac{\partial S}{\partial t}(t, x(t)) + p(t) \cdot \nabla_p H(x(t), p(t))$$
$$= -H(x(t), p(t)) + p(t) \cdot \nabla_p H(x(t), p(t)).$$

Hence we have proved the following proposition.

Proposition 38.5. *If S solves the Hamilton Jacobi Equation Eq. 38.17 and $(x(t), p(t))$ are solutions to the Hamilton Equations (38.18) and (38.19) (see also Eq. (38.29) below) with $p(0) = (\nabla_x g)(x(0))$ then*

$$S(T, x(T)) = g(x(0)) + \int_0^T [p(t) \cdot \nabla_p H(x(t), p(t)) - H(x(t), p(t))]\, dt.$$

In particular if $(T, x) \in \mathbb{R} \times \mathbb{R}^n$ then

$$S(T, x) = g(x(0)) + \int_0^T [p(t) \cdot \nabla_p H(x(t), p(t)) - H(x(t), p(t))]\, dt. \tag{38.20}$$

provided (x, p) is a solution to Hamilton Equations (38.18) and (38.19) satisfying the boundary condition $x(T) = x$ and $p(0) = (\nabla_x g)(x(0))$.

Remark 38.6. Let $X(t, x_0, p_0) = x(t)$ and $P(t, x_0, p_0) = p(t)$ where $(x(t), p(t))$ satisfies Hamilton Equations (38.18) and (38.19) with $(x(0), p(0)) = (x_0, p_0)$ and let $\Psi(t, x) := (t, X(t, x, \nabla g(x))$. Then $\Psi(0, x) = (0, x)$ so

$$\partial_v \Psi(0, 0) = (0, v) \text{ and } \partial_t \Psi(0, 0) = (1, \nabla_p H(x, \nabla g(x)))$$

from which it follows that $\Psi'(0, 0)$ is invertible. Therefore given $a \in \mathbb{R}^n$, the exists $\epsilon > 0$ such that $\Psi^{-1}(t, x)$ is well defined for $|t| < \epsilon$ and $|x - a| < \epsilon$. Writing $\Psi^{-1}(T, x) = (T, x_0(T, x))$ we then have that

$$(x(t), p(t)) := (X(t, x_0(T, x), \nabla g(x_0(T, x)), P(t, x_0, \nabla g(x_0(T, x)))$$

solves Hamilton Equations (38.18) and (38.19) satisfies the boundary condition $x(T) = x$ and $p(0) = (\nabla_x g)(x(0))$.

38.1.2 The connection with the Euler Lagrange Equations

Our next goal is to express the solution $S(T, x)$ in Eq. (38.20) solely in terms of the path $x(t)$. For this we digress a bit to Lagrangian mechanics and the notion of the "classical action."

Definition 38.7. *Let $T > 0$, $L : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$ be a smooth "Lagrangian" and $g : \mathbb{R}^n \to \mathbb{R}$ be a smooth function. The g – **weighted action** $I_T^g(q)$ of a function $q \in C^2([0, T], \mathbb{R}^n)$ is defined to be*

$$I_T^g(q) = g(q(0)) + \int_0^T L(q(t), \dot{q}(t))dt.$$

When $g = 0$ we will simply write I_T for $I_{0,T}$.

We are now going to study the function $S(T, x)$ of "least action,"

$$S(T, x) := \inf \left\{I_T^g(q) : q \in C^2([0, T]) \text{ with } q(T) = x\right\} \tag{38.21}$$
$$= \inf \left\{g(q(0)) + \int_0^T L(q(t), \dot{q}(t))dt : q \in C^2([0, T]) \text{ with } q(T) = x\right\}.$$

The next proposition records the differential of $I_T^g(q)$.

Proposition 38.8. *Let $L \in C^\infty(\mathbb{R}^n \times \mathbb{R}^n, \mathbb{R})$ be a smooth Lagrangian, then for $q \in C^2([0, T], \mathbb{R}^n)$ and $h \in C^1([0, T], \mathbb{R}^n)$*

$$DI_T^g(q)h = [(\nabla g(q) - D_2 L(q, \dot{q})) \cdot h]_{t=0} + [D_2 L(q, \dot{q}) \cdot h]_{t=T}$$
$$+ \int_0^T (D_1 L(q, \dot{q}) - \frac{d}{dt} D_2 L(q, \dot{q}))h\; dt \tag{38.22}$$

Proof. By differentiating past the integral,

$$\partial_h I_T(q) = \frac{d}{ds}|_0 I_T(q + sh) = \int_0^T \frac{d}{ds}|_0 L(q(t) + sh(t), \dot{q}(t) + s\dot{h}(t))dt$$
$$= \int_0^T (D_1 L(q, \dot{q})h + D_2 L(q, \dot{q})\dot{h})dt$$
$$= \int_0^T (D_1 L(q, \dot{q}) - \frac{d}{dt} D_2 L(q, \dot{q}))h\; dt + D_2 L(q, \dot{q})h\Big|_0^T.$$

This completes the proof since $I_T^g(q) = g(q(0)) + I_T(q)$ and $\partial_h [g(q(0))] = \nabla g(q(0)) \cdot h(0)$. ■

Definition 38.9. *A function $q \in C^2([0, T], \mathbb{R}^n)$ is said to solve the **Euler Lagrange equation for** L if q solves*

$$D_1 L(q, \dot{q}) - \frac{d}{dt}[D_2 L(q, \dot{q})] = 0. \tag{38.23}$$

This is equivalently to q satisfying $DI_T(q)h = 0$ for all $h \in C^1([0, T], \mathbb{R}^n)$ which vanish on $\partial[0, T] = \{0, T\}$.

Let us note that the Euler Lagrange equations may be written as:

$$D_1 L(q, \dot{q}) = D_1 D_2 L(q, \dot{q})\dot{q} + D_2^2 L(q, \dot{q})\ddot{q}.$$

Corollary 38.10. *Any minimizer q (or more generally critical point) of $I_T^g(\cdot)$ must satisfy the Euler Lagrange Eq. (38.23) with the boundary conditions*

$$q(T) = x \text{ and } \nabla g(q(0)) = \nabla_{\dot{q}} L(q(0), q(0)) = D_2 L(q(0), q(0)). \tag{38.24}$$

Proof. The corollary is a consequence Proposition 38.8 and the first derivative test which implies $DI_T^g(q)h = 0$ for all $h \in C^1([0,T], \mathbb{R}^n)$ such that $h(T) = 0$. ■

Example 38.11. Let $U \in C^\infty(\mathbb{R}^n, \mathbb{R})$, $m > 0$ and $L(q,v) = \frac{1}{2} m |v|^2 - U(q)$. Then

$$D_1 L(q,v) = -\nabla U(q) \text{ and } D_2 L(q,v) = mv$$

and the Euler Lagrange equations become

$$-\nabla U(q) = \frac{d}{dt}[m\dot{q}] = m\ddot{q}$$

which are Newton's equations of motion for a particle of mass m subject to a force $-\nabla U$. In particular if $U = 0$, then $q(t) = q(0) + t\dot{q}(0)$.

The following assumption on L will be assumed for the rest of this section.

Assumption 1 *We assume $\left[D_2^2 L(q,v)\right]^{-1}$ exists for all $(q,v) \in \mathbb{R}^n \times \mathbb{R}^n$ and $v \to D_2 L(q,v)$ is invertible for all $q \in \mathbb{R}^n$.*

Notation 38.12 *For $q, p \in \mathbb{R}^n$ let*

$$V(q,p) := \left[D_2 L(q,\cdot)\right]^{-1}(p). \tag{38.25}$$

Equivalently, $V(q,p)$ is the unique element of $\mathbb{R}^n$ such that

$$D_2 L(q, V(q,p)) = p. \tag{38.26}$$

Remark 38.13. The function $V : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ is smooth in (q,p). This is a consequence of the implicit function theorem applied to $\Psi(q,v) := (q, D_2 L(q,v))$.

Under Assumption 1, Eq. (38.23) may be written as

$$\ddot{q} = F(q, \dot{q}) \tag{38.27}$$

where

$$F(q,\dot{q}) = D_2^2 L(q,\dot{q})^{-1}\{D_1 L(q,\dot{q}) - D(D_2 L(q,\dot{q})\dot{q}\}.$$

Definition 38.14 (Legendre Transform). *Let $L \in C^\infty(\mathbb{R}^n \times \mathbb{R}^n, \mathbb{R})$ be a function satisfying Assumption 1. The Legendre transform $L^* \in C^\infty(\mathbb{R}^n \times \mathbb{R}^n, \mathbb{R})$ is defined by*

$$L^*(x,p) := p \cdot v - L(x,v) \text{ where } p = \nabla_v L(x,v),$$

i.e.

$$L^*(x,p) = p \cdot V(x,p) - L(x, V(x,p)). \tag{38.28}$$

Proposition 38.15. *Let $H(x,p) := L^*(x,p)$, $q \in C^2([0,T], \mathbb{R}^n)$ and $p(t) := L_v(q(t), \dot{q}(t))$. Then*

1. *$H \in C^\infty(\mathbb{R}^n \times \mathbb{R}^n, \mathbb{R})$ and*
$$H_x(x,p) = -L_x(x, V(x,p)) \text{ and } H_p(x,p) = V(x,p)..$$
2. *H satisfies Assumption 1 and $H^* = L$, i.e. $(L^*)^* = L$.*
3. *The path $q \in C^2([0,T], \mathbb{R}^n)$ solves the Euler Lagrange Eq. (38.23) then $(q(t), p(t))$ satisfies **Hamilton's Equations**:*
$$\begin{aligned}\dot{q}(t) &= H_p(q(t), p(t)) \\ \dot{p}(t) &= -H_x(q(t), p(t)).\end{aligned} \tag{38.29}$$
4. *Conversely if (q,p) solves Hamilton's equations (38.29) then q solves the Euler Lagrange Eq. (38.23) and*
$$\frac{d}{dt} H(q(t), p(t)) = 0. \tag{38.30}$$

Proof. The smoothness of H follows by Remark 38.13.

1. Using Eq. (38.28) and Eq. (38.26)
$$\begin{aligned}H_x(x,p) &= p \cdot V_x(x,p) - L_x(x, V(x,p)) - L_v(x, V(x,p))V_x(x,p) \\ &= p \cdot V_x(x,p) - L_x(x, V(x,p)) - p \cdot V_x(x,p) \\ &= -L_x(x, V(x,p)).\end{aligned}$$
and similarly,
$$\begin{aligned}H_p(x,p) &= V(x,p) + p \cdot V_p(x,p) - L_v(x, V(x,p))V_p(x,p) \\ &= V(x,p) + p \cdot V_p(x,p) - p \cdot V_p(x,p) = V(x,p).\end{aligned}$$
2. Since $H_p(x,p) = V(x,p) = \left[L_v(x,\cdot)\right]^{-1}(p)$ and by Remark 38.13, $p \to V(x,p)$ is smooth with a smooth inverse $L_v(x,\cdot)$, it follows that H satisfies Assumption 1. Letting $p = L_v(x,v)$ in Eq. (38.28) shows
$$\begin{aligned}H(x, L_v(x,v)) &= L_v(x,v) \cdot V(x, L_v(x,v)) - L(x, V(x, L_v(x,v))) \\ &= L_v(x,v) \cdot v - L(x,v)\end{aligned}$$
and using this and the definition of H^* we find
$$\begin{aligned}H^*(x,v) &= v \cdot \left[H_p(x,\cdot)\right]^{-1}(v) - H(x, \left[H_p(x,\cdot)\right]^{-1}(v)) \\ &= v \cdot L_v(x,v) - H(x, L_v(x,v)) = L(x,v).\end{aligned}$$

3. Now suppose that q solves the Euler Lagrange Eq. (38.23) and $p(t) = L_v(q(t), \dot{q}(t))$, then
$$\dot{p} = \frac{d}{dt} L_v(q, \dot{q}) = L_q(q, \dot{q}) = L_q(q, V(q,p)) = -H_q(q,p)$$
and
$$\dot{q} = [L_v(q, \cdot)]^{-1}(p) = V(q,p) = H_p(q,p).$$
4. Conversely if (q,p) solves Eq. (38.29), then
$$\dot{q} = H_p(q,p) = V(q,p).$$
Therefore
$$L_v(q, \dot{q}) = L_v(q, V(q,p)) = p$$
and
$$\frac{d}{dt} L_v(q, \dot{q}) = \dot{p} = -H_q(q,p) = L_q(q, V(q,p)) = L_q(q, \dot{q}).$$
Equation (38.30) is easily verified as well:
$$\begin{aligned}\frac{d}{dt} H(q,p) &= H_q(q,p) \cdot \dot{q} + H_p(q,p) \cdot \dot{p} \\ &= H_q(q,p) \cdot H_p(q,p) - H_p(q,p) \cdot H_q(q,p) = 0.\end{aligned}$$

■

Example 38.16. Letting $L(q,v) = \frac{1}{2} m |v|^2 - U(q)$ as in Example 38.11, L satisfies Assumption 1,
$$V(x,p) = [\nabla_v L(x, \cdot)]^{-1}(p) = p/m$$
$$H(x,p) = L^*(x,p) = p \cdot \frac{p}{m} - L(x, p/m)) = \frac{1}{2m} |p|^2 + U(q)$$
which is the conserved energy for this classical mechanical system. Hamilton's equations for this system are,
$$\dot{q} = p/m \text{ and } \dot{p} = -\nabla U(q).$$

Notation 38.17 *Let $\phi_t(x,v) = q(t)$ where q is the unique maximal solution to Eq. (38.27) (or equivalently 38.23)) with $q(0) = x$ and $\dot{q}(0) = v$.*

Theorem 38.18. *Suppose $L \in C^\infty(\mathbb{R}^n \times \mathbb{R}^n, \mathbb{R})$ satisfies Assumption 1 and let $H = L^*$ denote the Legendre transform of L. Assume there exists an open interval $J \subset \mathbb{R}$ with $0 \in J$ and $U \subset_o \mathbb{R}^n$ such that there exists a smooth function $x_0 : J \times U \to \mathbb{R}^n$ such that*
$$\phi_T(x_0(T,x), V(x_0(T,x), \nabla g(x_0(T,x))) = x. \tag{38.31}$$
Let

$$q_{x,T}(t) := \phi_t(x_0(T,x), V(x_0(T,x), \nabla g(x_0(T,x))) \tag{38.32}$$
so that $q_{x,T}$ solves the Euler Lagrange equations, $q_{x,T}(T) = x$, $q_{x,T}(0) = x_0(T,x)$ and $\dot{q}_{x,T}(0) = V(x_0(T,x), \nabla g(x_0(T,x))$ or equivalently
$$\partial_v L(q_{x,T}(0), \dot{q}_{x,T}(0)) = \nabla g(x_0(T,x)).$$
Then the function
$$S(T,x) := I_T^g(q_{x,T}) = g(q_{x,T}(0)) + \int_0^T L(q_{x,T}(t), \dot{q}_{x,T}(t)) dt. \tag{38.33}$$
solves the Hamilton Jacobi Equation (38.17).

Conjecture 38.19. For general g and L convex in v, the function
$$S(t,x) = \inf_{q \in C^2([0,t], \mathbb{R}^n)} \{g(q(0)) + \int_0^t L(q(\tau), \dot{q}(\tau)) d\tau : \ q(t) = x\}$$
is a distributional solution to the Hamilton Jacobi Equation Eq. 38.17. See Evans to learn more about this conjecture.

Proof. We will give two proofs of this Theorem.

First Proof. One need only observe that the theorem is a consequence of Definition 38.14 and Proposition 38.15 and 38.5.

Second Direct Proof. By the fundamental theorem of calculus and differentiating past the integral,
$$\begin{aligned}\frac{\partial S(T,x)}{\partial T} &= \nabla g(x_0(T,x)) \cdot \frac{\partial}{\partial T} x_0(T,x) + L(q_{x,T}(T), \dot{q}_{x,T}(T)) \\ &\quad + \int_0^T \frac{\partial}{\partial T} L(q_{x,T}(t), \dot{q}_{x,T}(t))) dt \\ &= \nabla g(x_0(T,x)) \cdot \frac{\partial}{\partial T} x_0(T,x) + L(q_{x,T}(T), \dot{q}_{x,T}(T)) \\ &\quad + DI_T(q_{x,T}) \left[\frac{\partial}{\partial T} q_{x,T} \right] \\ &= L(q_{x,T}(T), \dot{q}_{x,T}(T)) + DI_T^g(q_{x,T}) \left[\frac{\partial}{\partial T} q_{x,T} \right].\end{aligned} \tag{38.34}$$

Using Proposition 38.8 and the fact that $q_{x,T}$ satisfies the Euler Lagrange equations and the boundary conditions in Corollary 38.10 we find
$$DI_T^g(q_{x,T}) \left[\frac{\partial}{\partial T} q_{x,T} \right] = \left(D_2 L(q_{x,T}(t), \dot{q}_{x,T}(t)) \frac{\partial}{\partial T} q_{x,T}(t) \right) \Big|_{t=T}. \tag{38.35}$$

Furthermore differentiating the identity, $q_{x,T}(T) = x$, in T implies
$$0 = \frac{d}{dT} x = \frac{d}{dT} q_{x,T}(T) = \dot{q}_{x,T}(T) + \frac{d}{dT} q_{x,T}(t)|_{t=T} \tag{38.36}$$

Combining Eqs. (38.34) – (38.36) gives

$$\frac{\partial S(T,x)}{\partial T} = L(x, \dot{q}_{x,T}(T))) - D_2L(x, \dot{q}_{x,T}(T))\dot{q}_{x,T}(T). \tag{38.37}$$

Similarly for $v \in \mathbb{R}^n$,

$$\begin{aligned}\partial_v S(T,x) &= \partial_v I_T^g(q_{x,T}) = DI_T^g((q_{x,T}))\,[\partial_v q_{x,T}] \\ &= D_2L(q_{x,T}(T), \dot{q}_{x,T}(T))\partial_v q_{x,T}(T) = D_2L(x, \dot{q}_{x,T}(T))v\end{aligned}$$

wherein the last equality we have use $q_{x,T}(T) = x$. This last equation is equivalent to

$$D_2L(x, \dot{q}_{x,T}(T)) = \nabla_x S(T,x)$$

from which it follows that

$$\dot{q}_{x,T}(T) = V(x, \nabla_x S(T,x)). \tag{38.38}$$

Combining Eqs. (38.37) and (38.38) and the definition of H, shows

$$\begin{aligned}\frac{\partial S(T,x)}{\partial T} &= L(x, V(x, \nabla_x S(T,x))) - D_2L(x, \dot{q}_{x,T}(T))V(x, \nabla_x S(T,x)) \\ &= -H(x, \nabla_x S(T,x)).\end{aligned}$$

■

Remark 38.20. The hypothesis of Theorem 38.18 may always be satisfied locally, for let $\psi : \mathbb{R}\times\mathbb{R}^n \to \mathbb{R}\times\mathbb{R}^n$ be given by $\psi(t,y) := (t, \phi_t(y, V(y, \nabla g(y)))$. Then $\psi(0,y) := (0,y)$ and so

$$\dot{\psi}(0,y) = (1,*) \text{ and } \psi_y(0,y) = id_{\mathbb{R}^n}$$

from which it follows that $\psi'(0,y)^{-1}$ exists for all $y \in \mathbb{R}^n$. So the inverse function theorem guarantees for each $a \in \mathbb{R}^n$ that there exists an open interval $J \subset \mathbb{R}$ with $0 \in J$ and $a \in U \subset_o \mathbb{R}^n$ and a smooth function $x_0 : J \times U \to \mathbb{R}^n$ such that

$$\psi(T, x_0(T,x)) = (T, x_0(T,x)) \text{ for } T \in J \text{ and } x \in U,$$

i.e.

$$\phi_T(x_0(T,x), V(x_0(T,x), \nabla g(x_0(T,x))) = x.$$

38.2 Geometric meaning of the Legendre Transform

Let V be a finite dimensional real vector space and $f : V \to \mathbb{R}$ be a strictly convex function. Then the function $f^* : V^* \to \mathbb{R}$ defined by

$$f^*(\alpha) = \sup_{v\in V} (\alpha(v) - f(v)) \tag{38.39}$$

is called the Legendre transform of f. Now suppose the supremum on the right side of Eq. (38.39) is obtained at a point $v \in V$, see Figure 38.2 below. Eq. (38.39) may be rewritten as $f^*(\alpha) \ge \alpha(\cdot) - f(\cdot)$ with equality at v or equivalently that

$$-f^*(\alpha) + \alpha(\cdot) \le f(\cdot) \text{ with equality at some point } v \in V.$$

Geometrically, the graph of $\alpha \in V^*$ defines a hyperplane which if translate by $-f^*(\alpha)$ just touches the graph of f at one point, say v, see Figure 38.2. At the point of contact, v, α and f must have the same tangent plane and

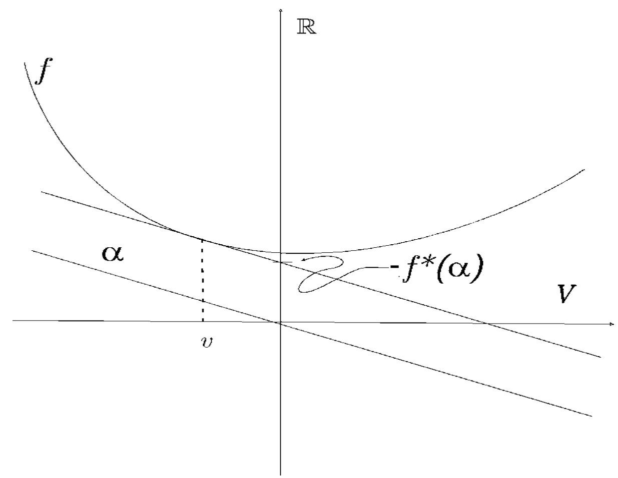


Fig. 38.2. Legendre Transform of f.

since α is linear this means that $f'(v) = \alpha$. Therefore the Legendre transform $f^* : V^* \to \mathbb{R}$ of f may be given explicitly by

$$f^*(\alpha) = \alpha(v) - f(v) \text{ with } v \text{ such that } f'(v) = \alpha.$$

39

Cauchy – Kovalevskaya Theorem

As a warm up we will start with the corresponding result for ordinary differential equations.

Theorem 39.1 (ODE Version of Cauchy – Kovalevskaya, I.). *Suppose $a > 0$ and $f : (-a, a) \to \mathbb{R}$ is real analytic near 0 and $u(t)$ is the unique solution to the ODE*

$$\dot{u}(t) = f(u(t)) \text{ with } u(0) = 0. \tag{39.1}$$

Then u is also real analytic near 0.

We will give four proofs. However it is the last proof that the reader should focus on for understanding the PDE version of Theorem 39.1.

Proof. (First Proof.)If $f(0) = 0$, then $u(t) = 0$ for all t is the unique solution to Eq. (39.1) which is clearly analytic. So we may now assume that $f(0) \neq 0$. Let $G(z) := \int_0^z \frac{1}{f(u)} du$, another real analytic function near 0. Then as usual we have

$$\frac{d}{dt} G(u(t)) = \frac{1}{f(u(t))} \dot{u}(t) = 1$$

and hence $G(u(t)) = t$. We then have $u(t) = G^{-1}(t)$ which is real analytic near $t = 0$ since $G'(0) = \frac{1}{f(0)} \neq 0$. ■

Proof. (Second Proof.) For $z \in \mathbb{C}$ let $u_z(t)$ denote the solution to the ODE

$$\dot{u}_z(t) = z f(u_z(t)) \text{ with } u_z(0) = 0. \tag{39.2}$$

Notice that if $u(t)$ is analytic, then $t \to u(tz)$ satisfies the same equation as u_z. Since $G(z, u) = zf(u)$ is holomorphic in z and u, it follows that u_z in Eq. (39.2) depends holomorphically on z as can be seen by showing $\bar{\partial}_z u_z = 0$, i.e. showing $z \to u_z$ satisfies the Cauchy Riemann equations. Therefore if $\epsilon > 0$ is chosen small enough such that Eq. (39.2) has a solution for $|t| < \epsilon$ and $|z| < 2$, then

$$u(t) = u_1(t) = \sum_{n=0}^{\infty} \frac{1^n}{n!} \partial_z^n u_z(t)|_{z=0}. \tag{39.3}$$

Now when $z \in \mathbb{R}$, $u_z(t) = u(tz)$ and therefore

$$\partial_z^n u_z(t)|_{z=0} = \partial_z^n u(tz)|_{z=0} = u^{(n)}(0) t^n.$$

Putting this back in Eq. (39.3) shows

$$u(t) = \sum_{n=0}^{\infty} \frac{1}{n!} u^{(n)}(0) t^n$$

which shows $u(t)$ is analytic for t near 0. ■

Proof. (Third Proof.) Go back to the original proof of existence of solutions, but now replace t by $z \in \mathbb{C}$ and $\int_0^t f(u(\tau)) d\tau$ by $\int_0^z f(u(\xi)) d\xi = \int_0^1 f(u(tz)) z dt$. Then the usual Picard iterates proof work in the class of holomorphic functions to give a holomorphic function $u(z)$ solving Eq. (39.1). ■

Proof. (Fourth Proof: Method of Majorants) Suppose for the moment we have an analytic solution to Eq. (39.1). Then by repeatedly differentiating Eq. (39.1) we learn

$$\begin{aligned}
\ddot{u}(t) &= f'(u(t))\dot{u}(t) = f'(u(t)) f(u(t)) \\
u^{(3)}(t) &= f''(u(t)) f^2(u(t)) + [f'(u(t))]^2 f(u(t)) \\
&\vdots \\
u^{(n)}(t) &= p_n \left(f(u(t)), \dots, f^{(n-1)}(u(t)) \right)
\end{aligned}$$

where p_n is a polynomial in n variables with all non-negative integer coefficients. The first few polynomials are $p_1(x) = x$, $p_2(x, y) = xy$, $p_3(x, y, z) = x^2 z + xy^2$. Notice that these polynomials are universal, i.e. are independent of the function f and

$$\begin{aligned}
\left| u^{(n)}(0) \right| &= \left| p_n \left(f(0), \dots, f^{(n-1)}(0) \right) \right| \\
&\le p_n \left(|f(0)|, \dots, \left| f^{(n-1)}(0) \right| \right) \le p_n \left(g(0), \dots, g^{(n-1)}(0) \right)
\end{aligned}$$

where g is any analytic function such that $\left| f^{(k)}(0) \right| \le g^{(k)}(0)$ for all $k \in \mathbb{Z}_+$. (We will abbreviate this last condition as $f \ll g$.) Now suppose that $v(t)$ is a solution to

$$\dot{v}(t) = g(v(t)) \text{ with } v(0) = 0, \tag{39.4}$$

then we know from above that

$$v^{(n)}(0) = p_n \left(g(0), \dots, g^{(n-1)}(0) \right) \ge \left| u^{(n)}(0) \right| \text{ for all } n.$$

Hence if knew that v were analytic with radius of convergence larger that some $\rho > 0$, then by comparison we would find

$$\sum_{n=0}^{\infty} \frac{1}{n!} \left| u^{(n)}(0) \right| \rho^n \le \sum_{n=0}^{\infty} \frac{1}{n!} v^{(n)}(0) \rho^n < \infty$$

and this would show

$$u(t) := \sum_{n=0}^{\infty} \frac{1}{n!} p_n \left(f(0), \dots, f^{(n-1)}(0) \right) t^n$$

is a well defined analytic function for $|t| < \rho$.

I now claim that $u(t)$ solves Eq. (39.1). Indeed, both sides of Eq. (39.1) are analytic in t, so it suffices to show the derivatives of each side of Eq. (39.1) agree at $t = 0$. For example $\dot{u}(0) = f(0)$, $\ddot{u}(0) = \frac{d}{dt}|_0 f(u(t))$, etc. However this is the case by the very definition of $u^{(n)}(0)$ for all n.

So to finish the proof, it suffices to find an analytic function g such that $\left| f^{(k)}(0) \right| \le g^{(k)}(0)$ for all $k \in \mathbb{Z}_+$ and for which we know the solution to Eq. (39.4) is analytic about $t = 0$. To this end, suppose that the power series expansion for $f(t)$ at $t = 0$ has radius of convergence larger than $r > 0$, then $\sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(0) r^n$ is convergent and in particular,

$$C := \max_n \left| \frac{1}{n!} f^{(n)}(0) r^n \right| < \infty$$

from which we conclude

$$\max_n \left| \frac{1}{n!} f^{(n)}(0) \right| \le C r^{-n}.$$

Let

$$g(u) := \sum_{n=0}^{\infty} C r^{-n} u^n = C \frac{1}{1 - u/r} = C \frac{r}{r-u}.$$

Then clearly $f \ll g$. To conclude the proof, we will explicitly solve Eq. (39.4) with this function $g(t)$,

$$\dot{v}(t) = C \frac{r}{r - v(t)} \text{ with } v(0) = 0.$$

By the usual separation of variables methods we find $rv(t) - \frac{1}{2} v^2(t) = Crt$, i.e.

$$2Crt - 2rv(t) + v^2(t) = 0$$

which has solutions, $v(t) = r \pm \sqrt{r^2 - 2Crt}$. We must take the negative sign to get the correct initial condition, so that

$$v(t) = r - \sqrt{r^2 - 2Crt} = r - r\sqrt{1 - 2Ct/r} \tag{39.5}$$

which is real analytic for $|t| < \rho := r/C$. ■

Let us now Jazz up this theorem to that case of a system of ordinary differential equations. For this we will need the following lemma.

Lemma 39.2. *Suppose $h : (-a,a)^d \to \mathbb{R}^d$ is real analytic near $0 \in (-a,a)^d$, then*

$$h \ll \frac{Cr}{r - z_1 - \dots - z_d}$$

for some constants C and r.

Proof. By definition, there exists $\rho > 0$ such that

$$h(z) = \sum_\alpha h_\alpha z^\alpha \text{ for } |z| < \rho$$

where $h_\alpha = \frac{1}{\alpha!} \partial^\alpha h(0)$. Taking $z = r(1,1,\dots,1)$ with $r < \rho$ implies there exists $C < \infty$ such that $|h_\alpha| r^{|\alpha|} \le C$ for all α, i.e.

$$|h_\alpha| \le C r^{-|\alpha|} \le C \frac{|\alpha|!}{\alpha!} r^{-|\alpha|}.$$

This completes the proof since

$$\begin{aligned} \sum_\alpha C \frac{|\alpha|!}{\alpha!} r^{-|\alpha|} z^\alpha &= C \sum_{n=0}^{\infty} \sum_{|\alpha|=n} \frac{|\alpha|!}{\alpha!} \left(\frac{z}{r}\right)^\alpha = C \sum_{n=0}^{\infty} \left(\frac{z_1 + \dots + z_d}{r} \right)^n \\ &= C \frac{1}{1 - \left(\frac{z_1 + \dots + z_d}{r} \right)} = \frac{Cr}{r - z_1 - \dots - z_d} \end{aligned}$$

all of which is valid provided $|z| := |z_1| + \dots + |z_d| < r$. ■

Theorem 39.3 (ODE Version of Cauchy – Kovalevskaya, II.). *Suppose $a > 0$ and $f : (-a,a)^d \to \mathbb{R}^d$ be real analytic near $0 \in (-a,a)^d$ and $u(t)$ is the unique solution to the ODE*

$$\dot{u}(t) = f(u(t)) \text{ with } u(0) = 0. \tag{39.6}$$

Then u is also real analytic near 0.

Proof. All but the first proof of Theorem 39.1 may be adapted to the cover this case. The only proof which perhaps needs a little more comment is the fourth proof. By Lemma 39.2, we can find $C, r > 0$ such that

$$f_j(z) \ll g_j(z) := \frac{Cr}{r - z_1 - \dots - z_d}$$

for all j. Let $v(t)$ denote the solution to the ODE,

$$\dot{v}(t) = g(v(t)) = \frac{Cr}{r - v_1(t) - \dots - v_d(t)} (1,1,\dots,1) \tag{39.7}$$

with $v(0) = 0$. By symmetry, $v_j(t) = v_1(t) =: w(t)$ for each j so Eq. (39.7) implies

$$\dot{w}(t) = \frac{Cr}{r - dw(t)} = \frac{C(r/d)}{(r/d) - w(t)} \text{ with } w(0) = 0.$$

We have already solved this equation (see Eq. (39.5) with r replaced by r/d) to find

$$w(t) = r/d - \sqrt{r^2/d^2 - 2Crt/d} = r/d\left(1 - \sqrt{1 - 2Cdt/r}\right). \tag{39.8}$$

Thus $v(t) = w(t)(1, 1, \ldots, 1)$ is a real analytic function which is convergent for $|t| < r/(2Cd)$.

Now suppose that u is a real analytic solution to Eq. (39.6). Then by repeatedly differentiating Eq. (39.6) we learn

$$\ddot{u}_j(t) = \partial_i f_j(u(t))\dot{u}_i(t) = \partial_i f_j(u(t)) f_i(u(t))$$
$$u_j^{(3)}(t) = \partial_k \partial_i f_j(u(t))\dot{u}_k(t)\dot{u}_i(t) + \partial_i f_j(u(t))\ddot{u}_i(t)$$
$$\vdots$$
$$u_j^{(n)}(t) = p_n\left(\{\partial^\alpha f_j(u(t))\}_{|\alpha|<n}, \left\{u_i^{(k)}(t)\right\}_{k<n, 1\le i\le d}\right) \tag{39.9}$$

where p_n is a polynomial with all non-negative integer coefficients. We now define $u_j^{(n)}(0)$ inductively so that

$$u_j^{(n)}(0) = p_n\left(\{\partial^\alpha f_j(u(0))\}_{|\alpha|<n}, \left\{u_i^{(k)}(0)\right\}_{k<n, 1\le i\le d}\right)$$

for all n and j and we will attempt to define

$$u(t) = \sum_{n=0}^{\infty} \frac{1}{n!} u^{(n)}(0) t^n. \tag{39.10}$$

To see this sum is convergent we make use of the fact that the polynomials p_n are universal i.e. are independent of the function f_j) and have non-negative coefficients so that by induction

$$\left|u_j^{(n)}(0)\right| \le p_n\left(\{|\partial^\alpha f_j(u(0))|\}_{|\alpha|<n}, \left\{\left|u_i^{(k)}(0)\right|\right\}_{k<n, 1\le i\le d}\right)$$
$$\le p_n\left(\{\partial^\alpha g_j(u(0))\}_{|\alpha|<n}, \left\{v_i^{(k)}(0)\right\}_{k<n, 1\le i\le d}\right) = v_j^{(n)}(0).$$

Notice the when $n = 0$ that $|u_j(0)| = 0 = v_j(0)$.[1] Thus we have shown $u \ll v$ and so by comparison the sum in Eq. (39.10) is convergent for t near 0. As before $u(t)$ solves Eq. (39.6) since both functions $\dot{u}(t)$ and $f(u(t))$ are analytic functions of t which have common values for all derivatives in t at $t = 0$. ■

[1] The argument shows that $v_j^{(n)}(0) \ge 0$ for all n. This is also easily seen directly by induction using Eq. (39.9) with f replaced by g and the fact that $\partial^\alpha g_j(0) \ge 0$ for all α.

39.1 PDE Cauchy Kovalevskaya Theorem

In this section we will consider the following general quasi-linear system of partial differential equations

$$\sum_{|\alpha|=k} a_\alpha(x, J^{k-1}u)\partial_x^\alpha u(x) + c(x, J^{k-1}u) = 0 \tag{39.11}$$

where

$$J^l u(x) = (u(x), Du(x), D^2u(x), \ldots, D^l u(x))$$

is the "l – jet" of u. Here $u : \mathbb{R}^n \to \mathbb{R}^m$ and $a_\alpha(J^{k-1}u, x)$ is an $m \times m$ matrix. As usual we will want to give boundary data on some hypersurface $\Sigma \subset \mathbb{R}^n$. Let ν denote a smooth vector field along Σ such that $v(x) \notin T_x\Sigma$ ($T_x\Sigma$ is the tangent space to Σ at x) for $x \in \Sigma$. For example we might take $\nu(x)$ to be orthogonal to $T_x\Sigma$ for all $x \in \Sigma$. To hope to get a unique solution to Eq. (39.11) we will further assume there are smooth functions g_l on Σ for $l = 0, \ldots, k-1$ and we will require

$$D^l u(x)(v(x), \ldots, v(x)) = g_l(x) \text{ for } x \in \Sigma \text{ and } l = 0, \ldots, k-1. \tag{39.12}$$

Proposition 39.4. *Given a smooth function u on a neighborhood of Σ satisfying Eq. (39.12), we may calculate $D^l u(x)$ for $x \in \Sigma$ and $l < k$ in terms of the functions g_l and there tangential derivatives.*

Proof. Let us begin by choosing a coordinate system y on $\mathbb{R}^n$ such that $\Sigma \cap D(y) = \{y_n = 0\}$ and let us extend ν to a neighborhood of Σ by requiring $\frac{\partial \nu}{\partial y_n} = 0$. To complete the proof, we are going to show by induction on k that we may compute

$$\left(\frac{\partial}{\partial y}\right)^\alpha u(x) \text{ for all } x \in \Sigma \text{ and } |\alpha| < k$$

from Eq. (39.12).

The claim is clear when $k = 1$, since $u = g_0$ on Σ. Now suppose that $k = 2$ and let $\nu_i = \nu_i(y_1, \ldots, y_{n-1})$ such that

$$\nu = \sum_{i=1}^n \nu_i \frac{\partial}{\partial y_i} \text{ in a neighborhood of } \Sigma.$$

Then

$$g_1 = (Du)\,\nu = \nu u = \sum_{i=1}^n \nu_i \frac{\partial u}{\partial y_i} = \sum_{i<n} \nu_i \frac{\partial g_0}{\partial y_i} + \nu_n \frac{\partial u}{\partial y_n}.$$

Since ν is not tangential to $\Sigma = \{y_n = 0\}$, it follows that $\nu_n \ne 0$ and hence

$$\frac{\partial u}{\partial y_n} = \frac{1}{\nu_n}\left(g_1 - \sum_{i<n} \nu_i \frac{\partial g_0}{\partial y_i}\right) \text{ on } \Sigma. \tag{39.13}$$

For $k=3$, first observe from the equality $u=g_0$ on Σ and Eq. (39.13) we may compute all derivatives of u of the form $\frac{\partial^\alpha u}{\partial y^\alpha}$ on Σ provided $\alpha_n \le 1$. From Eq. (39.12) for $l=2$, we have

$$\begin{aligned} g_2 &= \left(D^2 u\right)(v,v) = v^2 u + \text{l.o.ts.} \\ &= \sum \nu_j \frac{\partial}{\partial y_j}\left(\nu_i \frac{\partial u}{\partial y_i}\right) + \text{l.o.ts.} = \nu_n^2 \frac{\partial^2 u}{\partial y_n^2} + \text{l.o.ts.} \end{aligned}$$

where l.o.ts. denotes terms involving $\frac{\partial^\alpha u}{\partial y^\alpha}$ with $\alpha_n \le 1$. From this result, it follows that we may compute $\frac{\partial^2 u}{\partial y_n^2}$ in terms of derivatives of g_0, g_1 and g_2. The reader is asked to finish the full inductive argument of the proof. ■

Remark 39.5. The above argument shows that from Eq. (39.12) we may compute $\frac{\partial^\alpha u}{\partial y^\alpha}$ for any α such that $\alpha_n < k$.

To study Eq. (39.11) in more detail, let us rewrite Eq. (39.11) in the y – coordinates. Using the product and the chain rule repeatedly Eq. (39.11) may be written as

$$\sum_{|\alpha|=k} b_\alpha(y, J^{k-1}u)\partial_y^\alpha u(y) + c(y, J^{k-1}u) = 0 \tag{39.14}$$

where

$$J^l u(y) = (u(y), Du(y), D^2u(y), \ldots, D^l u(y)).$$

We will be especially concerned with the $b_{(0,0,\ldots,0,k)}$ coefficient which can be determined as follows:

$$\begin{aligned} \sum_{|\alpha|=k} a_\alpha \left(\frac{\partial}{\partial x}\right)^\alpha &= \sum_{|\alpha|=k} a_\alpha \left(\sum_{j=1}^n \frac{\partial y_j}{\partial x}\frac{\partial}{\partial y_j}\right)^\alpha = \sum_{|\alpha|=k} a_\alpha \left(\frac{\partial y_n}{\partial x}\frac{\partial}{\partial y_n}\right)^\alpha + \text{l.o.ts.} \\ &= \sum_{|\alpha|=k} a_\alpha \left(\frac{\partial y_n}{\partial x}\right)^\alpha \left(\frac{\partial}{\partial y_n}\right)^k + \text{l.o.ts.} \end{aligned}$$

where l.o.ts. now denotes terms involving $\frac{\partial^\alpha u}{\partial y^\alpha}$ with $\alpha_n < k$. From this equation we learn that

$$b_{(0,0,\ldots,0,k)}(y, J^{k-1}u) = \sum_{|\alpha|=k} a_\alpha \left(\frac{\partial y_n}{\partial x}\right)^\alpha = \sum_{|\alpha|=k} a_\alpha \left(dy_n\left(\frac{\partial}{\partial x}\right)\right)^\alpha.$$

Definition 39.6. *We will say that boundary data $(\Sigma, g_0, \ldots, g_{k-1})$ is non-characteristic for Eq. (39.11) at $x\in\Sigma$ if*

$$b_{(0,0,\ldots,0,k)}(y, J^{k-1}u) = \sum_{|\alpha|=k} a_\alpha(x, J^{k-1}u(x))\left(dy_n\left(\frac{\partial}{\partial x}\right)\right)^\alpha$$

is invertible at x.

Notice that this condition is independent of the choice of coordinate system y. To see this, for $\xi \in (\mathbb{R}^n)^*$ let

$$\sigma(\xi) = \sum_{|\alpha|=k} a_\alpha(x, J^{k-1}u(x))\left(\xi\left(\frac{\partial}{\partial x}\right)\right)^\alpha$$

which is k – linear form on $(\mathbb{R}^n)^*$. This form is coordinate independent since if f is a smooth function such that $f(x)=0$ and $df_x=\xi$, then

$$\sigma(\xi) = \frac{1}{k!}\sum_{|\alpha|=k} a_\alpha(x, J^{k-1}u(x))\left(\frac{\partial}{\partial x}\right)^\alpha f^k|_x.$$

Noting that

$$b_{(0,0,\ldots,0,k)}(y, J^{k-1}u) = \sigma(dy_n)$$

our non-characteristic condition becomes, $\sigma(dy_n)$ is invertible. Finally dy_n is the unique element ξ of $(\mathbb{R}^n)^*\setminus\{0\}$ up to scaling such that $\xi|_{T_x\Sigma}\equiv 0$. So the non-characteristic condition may be written invariantly as $\sigma(\xi)$ is invertible for all (or any) $\xi\in(\mathbb{R}^n)^*\setminus\{0\}$ such that $\xi|_{T_x\Sigma}\equiv 0$.

Assuming the given boundary data is non-characteristic, Eq. (39.11) may be put into "standard form,"

$$\sum_{|\alpha|=k} b_\alpha(y, J^{k-1}u)\partial_y^\alpha u(y) + c(y, J^{k-1}u) = 0 \tag{39.15}$$

with

$$\frac{\partial^l u}{\partial y_n^l} = g_l \text{ on } y_n = 0 \text{ for } l < k$$

where $b_{(0,0,\ldots,0,k)}(y, J^{k-1}u) = Id$ - matrix and

$$J^l u(y) = (u(y), Du(y), D^2u(y), \ldots, D^l u(y)).$$

By adding new dependent variables and possible a new independent variable for y_n one may reduce the problem to solving the system in Eq. (39.20) below. The resulting theorem may be stated as follows.

Theorem 39.7 (Cauchy Kovalevskaya). *Suppose all the coefficients in Eq. (39.11) are real analytic and the boundary data in Eq. (39.12) are also real analytic and non-characteristic near some point $a\in\Sigma$. Then there is a unique real analytic solution to Eqs. (39.11) and (39.12). (The boundary data in Eq. (39.12) is said to be real analytic if there exists coordinates y as above which are real analytic and the functions ν and g_l for $l=0,\ldots,k-1$ are real analytic functions in the y – coordinate system.)*

Example 39.8. Suppose a, b, C, r are positive constants. We wish to show the solution to the quasi-linear PDE

$$w_t = \frac{Cr}{r-y-aw}\left[bw_y+1\right] \text{ with } w(0,y)=0 \tag{39.16}$$

is real analytic near $(t,y)=(0,0)$. To do this we will solve the equation using the method of characteristics. Let $g(y,z) := \frac{Cr}{r-y-az}$, then the characteristic equations are

$$\begin{aligned} t' &= 0 \text{ with } t(0)=0 \\ y' &= -bg(y,z) \text{ with } y(0)=y_0 \text{ and} \\ z' &= g(y,z) \text{ with } z(0)=0. \end{aligned}$$

From these equations we see that we may identify t with s and that $y+bz=y_0$. Thus $z(t)=w(t,y(t))$ satisfies

$$\begin{aligned} \dot{z} = g(y_0-bz,z) &= \frac{Cr}{r-y_0+bz-az} \\ &= \frac{Cr}{r-y_0+(b-a)z} \text{ with } z(0)=0. \end{aligned}$$

Integrating this equation gives

$$\begin{aligned} Crt &= \int_0^t (r-y_0+(b-a)z(\tau))\dot{z}(\tau)d\tau = (r-y_0)z - \frac{1}{2}(a-b)z^2 \\ &= (r-y-bz)z - \frac{1}{2}(a-b)z^2 = (r-y)z - \frac{1}{2}(a+b)z^2, \end{aligned}$$

i.e.

$$\frac{1}{2}(a+b)z^2 - (r-y)z + Crt = 0.$$

The quadratic formula gives

$$w(t,y) = \frac{1}{a+b}\left[(r-y)\pm\sqrt{(r-y)^2-2(a+b)Crt}\right]$$

and using $w(0,y)=0$ we conclude

$$w(t,y) = \frac{1}{a+b}\left[(r-y)-\sqrt{(r-y)^2-2(a+b)Crt}\right]. \tag{39.17}$$

Notice the w is real analytic for (t,y) near $(0,0)$.

In general we could use the method of characteristics and ODE properties (as in Example 39.8) to show

$$u_t = a(x,u)u_x + b(x,u) \text{ with } u(0,x)=g(x)$$

has local real analytic solutions if a,b and g are real analytic. The method would also work for the fully non-linear case as well. However, the method of characteristics fails for systems while the method we will present here works in this generality.

Exercise 39.9. Verify w in Eq. (39.17) solves Eq. (39.16).

Solution 39.10 (39.9). Let $\rho := \sqrt{(r-y)^2-2(a+b)Crt}$, then

$$\begin{aligned} w(t,y) &= \frac{1}{a+b}\left[r-y-\rho\right] = \frac{r-y}{a+b} - \frac{1}{a+b}\rho, \\ w_t &= Cr/\rho,\ \rho = r-y-(a+b)w \text{ and} \end{aligned}$$

$$bw_y+1 = \frac{b}{a+b}\left[-1+(r-y)/\rho\right]+1 = \frac{1}{a+b}\left[a+b(r-y)/\rho\right].$$

Hence

$$\begin{aligned} \frac{bw_y+1}{w_t} &= \frac{1}{(a+b)Cr}\left[\rho a + b(r-y)\right] \\ &= \frac{1}{(a+b)Cr}\left[(r-y-(a+b)w)a + b(r-y)\right] \\ &= \frac{1}{Cr}\left[r-y-aw\right] \end{aligned}$$

as desired.

Example 39.11. Now let us solve for

$$v(t,x) = \left(v^1,\dots,v^m\right)(t,x_1,\dots,x_n)$$

where v satisfies

$$v_t^j = \frac{Cr}{r-x_1-\dots-x_n-\sum_{k=1}^m v^k}\left[1+\sum_{i=1}^n\sum_{k=1}^m \partial_i v^k\right] \text{ with } v(0,x)=0.$$

By symmetry, $v^j = v^1 =: w(t,y)$ for all j where $y = x_1+\dots+x_n$. Since $\partial_i v^j = w_y$, the above equations all may be written as

$$w_t = \frac{Cr}{r-y-mw}\left[mnw_y+1\right] \text{ with } w(0,y)=0.$$

Therefore from Example 39.8 with $a=m$ and $b=mn$, we find

$$w(t,y) = \frac{1}{m(n+1)}\left[(r-y)-\sqrt{(r-y)^2-2m(n+1)Crt}\right]. \tag{39.18}$$

and hence that

$$v(t,x) = w(t,x_1+\dots+x_n)(1,1,1,\dots,1)\in\mathbb{R}^m. \tag{39.19}$$

39.2 Proof of Theorem 39.7

As is outlined in Evans, Theorem 39.7 may be reduced to the following theorem.

Theorem 39.12. *Let* $(t,x,z) = (t,x_1,\ldots,x_n,z_1,\ldots,z_m) \in \mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^m$ *and assume* $(t,x,z) \to B_j(t,x,z) \in \{m\times m - \text{matrices}\}$ *(for* $j=1,\ldots,n$*) and* $(t,x,z) \to c(t,x,z) \in \mathbb{R}^m$ *are real analytic functions near* $(0,0,0) \in \mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^m$ *and* $x \to f(x) \in \mathbb{R}^m$ *is real analytic near* $0 \in \mathbb{R}^n$*. Then there exists, in a neighborhood of* $(t,x) = (0,0) \in \mathbb{R}\times\mathbb{R}^n$*, a unique real analytic solution* $u(t,x) \in \mathbb{R}^m$ *to the quasi-linear system*

$$u_t(t,x) = \sum_{j=1}^n B_j(t,x,u(t,x))\partial_j u(t,x) + c(t,x,u(t,x)) \text{ with } u(0,x) = f(x). \tag{39.20}$$

Proof. (Sketch.)

Step 0. By replacing $u(t,x)$ by $u(t,x) - f(x)$, we may assume $f \equiv 0$. By letting $u^{m+1}(t,x) = t$ if necessary, we may assume B_j and c do not depend on t. With these reductions we are left to solve

$$u_t(t,x) = \sum_{j=1}^n B_j(x,u(t,x))\partial_j u(t,x) + c(x,u(t,x)) \text{ with } u(0,x) = 0. \tag{39.21}$$

Step 1. Let

$$g(x,z) := \frac{Cr}{r - x_1 - \cdots - x_n - z_1 - \cdots - z_m}$$

where C and r are positive constants such that

$$(B_j)_{kl} \ll g \text{ and } c_k \ll g$$

for all k,l,j. For this choice of C and r, let v denote the solution constructed in Example 39.11 above.

Step 2. By repeatedly differentiating Eq. (39.20), show that if u solves Eq. (39.20) then $\partial_x^\alpha\partial_t^k u^j(0,0)$ is a **universal** polynomial in the derivatives $\{\partial_t^l\partial_x^\alpha\}_{\alpha,l<k}$ of the entries of B_j and c and u with all coefficients being non-negative. Use this fact and induction to conclude

$$\left|\partial_x^\alpha\partial_t^k u^j(0,0)\right| \le \partial_x^\alpha\partial_t^k v^j(0,0) \text{ for all } \alpha, k \text{ and } l.$$

Step 3. Use the computation in Step 2. to define $\partial_x^\alpha\partial_t^k u^j(0,0)$ for all α and k and then defined

$$u(t,x) := \sum_{\alpha,k} \frac{\partial_x^\alpha\partial_t^k u(0,0)}{\alpha!k!} t^k x^\alpha. \tag{39.22}$$

Because of step 2. and Example 39.11, this series is convergent for (t,x) sufficiently close to zero.

Step 4. The function u defined in Step 3. solves Eq. (39.20) because both

$$u_t(t,x) \text{ and } \sum_{j=1}^n B_j(x,u(t,x))\partial_j u(t,x) + c(x,u(t,x))$$

are both real analytic functions in (t,x) each having, by construction, the same derivatives at $(0,0)$. ■

39.3 Examples

Corollary 39.13 (Isothermal Coordinates). *Suppose that we are given a metric* $ds^2 = Edx^2 + 2Fdxdy + Gdy^2$ *on* $\mathbb{R}^2$ *such that* G/E *and* F/E *are real analytic near* $(0,0)$*. Then there exists a complex function* u *and a positive function* ρ *such that* $Du(0,0)$ *is invertible and* $ds^2 = \rho\,|du|^2$ *where* $du = u_x dx + u_y dy$*.*

Proof. Working out $|du|^2$ gives

$$|du|^2 = |u_x|^2\,dx^2 + 2\operatorname{Re}(u_x\bar{u}_y)dxdy + |u_y|^2\,dy^2.$$

Writing $u_y = \lambda u_x$, the previous equation becomes

$$|du|^2 = |u_x|^2\left(dx^2 + 2\operatorname{Re}(\lambda)dxdy + |\lambda|^2\,dy^2\right).$$

Hence we must have

$$E = \rho\,|u_x|^2,\; F = \rho\,|u_x|^2\operatorname{Re}\lambda \text{ and } G = \rho\,|u_x|^2\,|\lambda|^2$$

or equivalently

$$F/E = \operatorname{Re}\lambda \text{ and } G/E = |\lambda|^2.$$

Writing $\lambda = a + ib$, we find $a = F/E$ and $a^2 + b^2 = G/E$ so that

$$\lambda = \frac{F}{E} \pm i\sqrt{G/E - (F/E)^2} = \frac{1}{E}\left(F \pm i\sqrt{GE - F^2}\right).$$

We make a choice of the sign above, then we are looking for $u(x,y) \in \mathbb{C}$ such that $u_y = \lambda u_x$. Letting $u = \alpha + i\beta$, the equation $u_y = \lambda u_x$ may be written as the system of real equations

$$\alpha_y = \operatorname{Re}\left[(a+ib)(\alpha_x + i\beta_x)\right] = a\alpha_x - b\beta_x \text{ and}$$
$$\beta_y = \operatorname{Im}\left[(a+ib)(\alpha_x + i\beta_x)\right] = a\beta_x + b\alpha_x$$

which is equivalent to

$$\begin{pmatrix} \alpha \\ \beta \end{pmatrix}_y = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}_x .$$

So we may apply the Cauchy Kovalevskaya theorem 39.12 with $t = y$ to find a real analytic solution to this equation with (say) $u(x,0) = x$, i.e. $\alpha(x,0) = x$ and $\beta(x,0) = 0$. (We could take $u(x,0) = f(x)$ for any real analytic function f such that $f'(0) \neq 0$.) The only thing that remains to check is that $Du(0,0)$ is invertible. But

$$\begin{aligned} Du(0,0) &= \begin{pmatrix} \operatorname{Re} u_x & \operatorname{Re} u_y \\ \operatorname{Im} u_x & \operatorname{Im} u_y \end{pmatrix} = \begin{pmatrix} \alpha_x & \alpha_y \\ \beta_x & \beta_y \end{pmatrix} \\ &= \begin{pmatrix} \alpha_x & a\alpha_x - b\beta_x \\ \beta_x & a\beta_x + b\alpha_x \end{pmatrix} \end{aligned}$$

so that

$$\det [Du] = b\left(\alpha_x^2 + \beta_x^2\right) = \operatorname{Im} \lambda \left|u_x\right|^2 .$$

Thus

$$\det [Du(0,0)] = \operatorname{Im} \lambda(0,0) = \pm\sqrt{G/E - (F/E)^2}|_{(0,0)} \neq 0.$$

■

Example 39.14. Consider the linear PDE,

$$u_y = u_x \text{ with } u(x,0) = f(x) \tag{39.23}$$

where $f(x) = \sum_{m=0}^{\infty} a_m x^m$ as real analytic function near $x = 0$ with radius of convergence ρ. (So for any $r < \rho$, $|a_m| \leq Cr^{-n}$.) Formally the solution to Eq. (39.23) should be given by

$$u(x,y) = \sum_{n=0}^{\infty} \frac{1}{n!} \partial_y^n u(x,y)|_{y=0} y^n .$$

Now using the PDE (39.23),

$$\partial_y^n u(x,y)|_{y=0} = \partial_x^n u(x,0) = f^{(n)}(x).$$

Thus we get

$$u(x,y) = \sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(x) y^n . \tag{39.24}$$

By the Cauchy estimates,

$$\left|f^{(n)}(x)\right| \leq \frac{n!\rho}{(\rho - |x|)^{n+1}}$$

and so

$$\sum_{n=0}^{\infty} \frac{1}{n!}\left|f^{(n)}(x) y^n\right| \leq \rho \sum_{n=0}^{\infty} \frac{|y|^n}{(\rho - |x|)^{n+1}}$$

which is finite provided $|y| < \rho - |x|$, i.e. $|x| + |y| < \rho$. This of course makes sense because we know the solution to Eq. (39.23) is given by

$$u(x,y) = f(x+y).$$

Now we can expand Eq. (39.24) out to find

$$\begin{aligned} u(x,y) &= \sum_{n=0}^{\infty} \frac{1}{n!}\left(\sum_{m \geq n} m(m-1)\dots(m-n+1) a_m x^{m-n}\right) y^n \\ &= \sum_{m \geq n \geq 0} \binom{m}{n} a_m x^{m-n} y^n . \end{aligned} \tag{39.25}$$

Since

$$\begin{aligned} \sum_{m \geq n \geq 0} \binom{m}{n} \left|a_m x^{m-n} y^n\right| &\leq C \sum_{m \geq n \geq 0} \binom{m}{n} \left|r^{-m} x^{m-n} y^n\right| \\ &= C \sum_{m \geq 0} r^{-m} \left(|x| + |y|\right)^m < \infty \end{aligned}$$

provided $|x| + |y| < r$. Since $r < \rho$ was arbitrary, it follows that Eq. (39.25) is convergent for $|x| + |y| < \rho$.

Let us redo this example. By the PDE in Eq. (39.23), $\partial_y^m \partial_x^n u(x,y) = \partial_x^{n+m} u(x,y)$ and hence

$$\partial_y^m \partial_x^n u(0,0) = f^{(m+n)}(0).$$

Written another way

$$D^\alpha u(0,0) = f^{(|\alpha|)}(0)$$

and so the power series expansion for u must be given by

$$u(x,y) = \sum_\alpha \frac{f^{(|\alpha|)}(0)}{\alpha!} (x,y)^\alpha . \tag{39.26}$$

Using $f^{(m)}(0)/m! \leq Cr^{-m}$ we learn

$$\begin{aligned} \sum_\alpha \left|\frac{f^{(|\alpha|)}(0)}{\alpha!}(x,y)^\alpha\right| &\leq C \sum_\alpha \frac{\left|f^{(|\alpha|)}(0)\right|}{\alpha!} |x|^{\alpha_1} |y|^{\alpha_2} = C \sum_{m=0}^{\infty} \frac{\left|f^{(|\alpha|)}(0)\right|}{m!} \sum_{|\alpha| = m} \frac{m!}{\alpha!} |x \\ &\leq C \sum_{m=0}^{\infty} r^{-m} \left(|x| + |y|\right)^m = C \frac{r}{r - (|x| + |y|)} < \infty \end{aligned}$$

if $|x| + |y| < r$. Since $r < \rho$ was arbitrary, it follows that the series in Eq. (39.26) converges for $|x| + |y| < \rho$.

Now it is easy to check directly that Eq. (39.26) solves the PDE. However this is necessary since by construction $D^\alpha u_y(0,0) = D^\alpha u_x(0,0)$ for all α. This implies, because u_y and u_x are both real analytic, that $u_x = u_y$.

Part XII

Elliptic ODE

40

A very short introduction to generalized functions

Let U be an open subset of $\mathbb{R}^n$ and

$$C_c^\infty(U) = \cup_{K \sqsubset\sqsubset U} C^\infty(K) \tag{40.1}$$

denote the set of smooth functions on U with compact support in U.

Definition 40.1. *A sequence $\{\phi_k\}_{k=1}^\infty \subset \mathcal{D}(U)$ converges to $\phi \in \mathcal{D}(U)$, iff there is a compact set $K \sqsubset\sqsubset U$ such that $\operatorname{supp}(\phi_k) \subset K$ for all k and $\phi_k \to \phi$ in $C^\infty(K)$.*

Definition 40.2 (Distributions on $U \subset_o \mathbb{R}^n$). *A generalized function T on $U \subset_o \mathbb{R}^n$ is a continuous linear functional on $\mathcal{D}(U)$, i.e. $T : \mathcal{D}(U) \to \mathbb{C}$ is linear and $\lim_{n\to\infty} \langle T, \phi_k \rangle = 0$ for all $\{\phi_k\} \subset \mathcal{D}(U)$ such that $\phi_k \to 0$ in $\mathcal{D}(U)$. Here we have written $\langle T, \phi \rangle$ for $T(\phi)$. We denote the space of generalized functions by $\mathcal{D}'(U)$.*

Example 40.3. Here are a couple of examples of distributions.

1. For $f \in L^1_{loc}(U)$ define $T_f \in \mathcal{D}'(U)$ by $\langle T_f, \phi \rangle = \int_U \phi f dm$ for all $\phi \in \mathcal{D}(U)$. This is called the distribution associated to f.
2. More generally let μ be a complex measure on U, then $\langle \mu, \phi \rangle := \int_U \phi d\mu$ is a distribution. For example if $x \in U$, and $\mu = \delta_x$ then $\langle \delta_x, \phi \rangle = \phi(x)$ for all $\phi \in \mathcal{D}$.

Lemma 40.4. *Let $a_\alpha \in C^\infty(U)$ and $L = \sum_{|\alpha| \le m} a_\alpha \partial^\alpha$ – a m^{th} order linear differential operator on $\mathcal{D}(U)$. Then for $f \in C^m(U)$ and $\phi \in \mathcal{D}(U)$,*

$$\langle Lf, \phi \rangle := \langle T_{Lf}, \phi \rangle = \langle T, L^\dagger \phi \rangle$$

*where $L^\dagger$ is the **formal adjoint** of L defined by*

$$L^\dagger \phi = \sum_{|\alpha| \le m} (-1)^{|\alpha|} \partial^\alpha [a_\alpha \phi].$$

Proof. This is simply repeated integration by parts. No boundary terms arise since ϕ has compact support. ■

Definition 40.5 (Multiplication by smooth functions). *Suppose that $g \in C^\infty(U)$ and $T \in \mathcal{D}'(U)$ then we define $gT \in \mathcal{D}'(U)$ by*

$$\langle gT, \phi \rangle = \langle T, g\phi \rangle \text{ for all } \phi \in \mathcal{D}(U).$$

It is easily checked that gT is continuous.

Definition 40.6 (Differentiation). *For $T \in \mathcal{D}'(U)$ and $i \in \{1, 2, \ldots, n\}$ let $\partial_i T \in \mathcal{D}'(U)$ be the distribution defined by*

$$\langle \partial_i T, \phi \rangle = -\langle T, \partial_i \phi \rangle \text{ for all } \phi \in \mathcal{D}(U).$$

Again it is easy to check that $\partial_i T$ is a distribution.

Definition 40.7. *More generally if L is as in Lemma 40.4 and $T \in \mathcal{D}'$ we define $LT \in \mathcal{D}'$ by*

$$\langle LT, \phi \rangle = \langle T, L^\dagger \phi \rangle.$$

Example 40.8. Suppose that $f \in L^1_{loc}$ and $g \in C^\infty(U)$, then $gT_f = T_{gf}$. If further $f \in C^1(U)$, then $\partial_i T_f = T_{\partial_i f}$. More generally if $f \in C^m(U)$ then, by Lemma 40.4, $LT_f = T_{Lf}$.

Because of Definition 40.7 we may now talk about distributional or generalized solutions T to PDEs of the form $LT = S$ where $S \in \mathcal{D}'$.

Example 40.9. For the moment let us also assume that $U = \mathbb{R}$. $\langle T_f, \phi \rangle = \int_U \phi f dm$. Then we have

1. $\lim_{M\to\infty} T_{\sin Mx} = 0$
2. $\lim_{M\to\infty} T_{M^{-1}\sin Mx} = \pi\delta_0$ where δ_0 is the point measure at 0.
3. If $f \in L^1(\mathbb{R}^n, dm)$ with $\int_{\mathbb{R}^n} f dm = 1$ and $f_\epsilon(x) = \epsilon^{-n} f(x/\epsilon)$, then $\lim_{\epsilon\downarrow 0} T_{f_\epsilon} = \delta_0$. Indeed,

$$\begin{aligned}\lim_{\epsilon\downarrow 0} \langle T_{f_\epsilon}, \phi \rangle &= \lim_{\epsilon\downarrow 0} \int_{\mathbb{R}^n} \epsilon^{-n} f(x/\epsilon) \phi(x) dx \\ &= \lim_{\epsilon\downarrow 0} \int_{\mathbb{R}^n} f(x) \phi(\epsilon x) dx \overset{\text{D.C.T.}}{=} \int_{\mathbb{R}^n} f(x) \lim_{\epsilon\downarrow 0} \phi(\epsilon x) dx \\ &= \phi(0) \int_{\mathbb{R}^n} f(x) dx = \phi(0) = \langle \delta_0, \phi \rangle.\end{aligned}$$

As a concrete example we have

$$\lim_{\epsilon\downarrow 0} \frac{\epsilon}{\pi(x^2 + \epsilon^2)} = \delta_0 \text{ on } \mathbb{R},$$

i.e.

$$\lim_{\epsilon\downarrow 0} T_{\frac{\epsilon}{\pi(x^2+\epsilon^2)}} = \delta_0.$$

Example 40.10. Suppose that $a \in U$, then

$$\langle \partial_i \delta_a, \phi \rangle = -\partial_i \phi(a)$$

and more generally we have

$$\langle L\delta_a, \phi \rangle = \left(L^{\dagger} \phi\right)(a).$$

Lemma 40.11. *Suppose $f \in C^1([a,b])$ and $g \in PC^1([a,b])$, i.e. $g \in C^1([a,b] \setminus \Lambda)$ where Λ is a finite subset of (a,b) and $g(\alpha+)$, $g(\alpha-)$ exists for $\alpha \in \Lambda$. Then*

$$\int_a^b f'(x)g(x)dx = [f'(x)g(x)]\,|_a^b - \int_a^b f(x)g'(x)dx - \sum_{\alpha \in \Lambda} f(\alpha)\left(g(\alpha+) - g(\alpha-)\right). \tag{40.2}$$

In particular

$$\frac{d}{dx}T_g = T_{g'} + \sum_{\alpha \in \Lambda} \left(g(\alpha+) - g(\alpha-)\right)\delta_\alpha$$

Proof. Write $\Lambda \cup \{a,b\}$ as $\{a = \alpha_0 < \alpha_1 < \cdots < \alpha_n = b\}$, then

$$\begin{aligned}
\int_a^b f'(x)g(x)dx &= \sum_{k=0}^{n-1} \int_{\alpha_k}^{\alpha_{k+1}} f'(x)g(x)dx \\
&= \sum_{k=0}^{n-1} \left[[f(x)g(x)]\,|_{\alpha_k+}^{\alpha_{k+1}-} - \int_{\alpha_k}^{\alpha_{k+1}} f(x)g'(x)dx \right] \\
&= [f'(x)g(x)]\,|_a^b - \int_a^b f(x)g'(x)dx - \sum_{k=1}^{n-1} [f(x)g(x)]\,|_{\alpha_k-}^{\alpha_k+}
\end{aligned}$$

which is the same as Eq. (40.2). ■

41

Elliptic Ordinary Differential Operators

Let $\Omega \subset_o \mathbb{R}^n$ be a bounded connected open region. A function $u \in C^2(\Omega)$ is said to satisfy Laplace's equation if

$$\triangle u = 0 \text{ in } \Omega.$$

More generally if $f \in C(\Omega)$ is given we say u solves the **Poisson equation** if

$$-\triangle u = f \text{ in } \Omega.$$

In order to get a unique solution to either of these equations it is necessary to impose "boundary" conditions on u.

Example 41.1. For **Dirichlet boundary conditions** we impose $u = g$ on $\partial\Omega$ and for **Neumann boundary conditions** we impose $\frac{\partial u}{\partial \nu} = g$ on $\partial\Omega$, where $g : \partial\Omega \to \mathbb{R}$ is a given function.

Lemma 41.2. *Suppose $f : \Omega \xrightarrow{C^0} \mathbb{R}$, $\partial\Omega$ is C^2 and $g : \partial\Omega \to \mathbb{R}$ is continuous. Then if there exists a solution to $-\triangle u = f$ with $u = g$ on $\partial\Omega$ such that $u \in C^2(\Omega^o) \cap C^1(\Omega)$ then the solution is unique.*

Definition 41.3. *Given an open set $\Omega \subset \mathbb{R}^n$ we say $u \in C^1(\overline{\Omega})$ if $u \in C^1(\Omega) \cap C(\overline{\Omega})$ and ∇u extends to a continuous function on $\overline{\Omega}$.*

Proof. If $\tilde{u}$ is another solution then $v = \tilde{u} - u$ solves $\triangle v = 0, v = 0$ on $\partial\Omega$. By the divergence theorem,

$$0 = \int_\Omega \triangle v \cdot v dm = -\int_\Omega |\nabla v|^2 dm + \int_{\partial\Omega} v \nabla v \cdot n d\sigma = -\int_\Omega |\nabla v|^2 dm,$$

where the boundary terms are zero since $v = 0$ on $\partial\Omega$. This identity implies $\int_\Omega |\nabla u|^2 dx = 0$ which then shows $\nabla v \equiv 0$ and since Ω is connected we learn v is constant on Ω. Because v is zero on $\partial\Omega$ we conclude $v \equiv 0$, that is $u = \tilde{u}$. ■

For the rest of this section we will now restrict to $n = 1$. However we will allow for more general operators than Δ in this case.

41.1 Symmetric Elliptic ODE

Let $a \in C^1([0,1], (0,\infty))$ and

$$Lf = -(af')' = -af'' - a'f' \text{ for } f \in C^2([0,1]). \tag{41.1}$$

In the following theorem we will impose Dirichlet boundary conditions on L by restricting the domain of L to

$$D(L) := \{f \in C^2([0,1], \mathbb{R}) : f(0) = f(1) = 0\}.$$

Theorem 41.4. *The linear operator $L : D(L) \to C([0,1], \mathbb{R})$ is invertible and $L^{-1} : C([0,1], \mathbb{R}) \to D(L) \subset C^2([0,1], \mathbb{R})$ is a bounded operator.*

Proof.

1. (Uniqueness) If $f, g \in D(L)$ then by integration by parts

$$(Lf, g) := \int_0^1 (Lf)(x) g(x) dx = \int_0^1 a(x) f'(x) g'(x) \, dx. \tag{41.2}$$

Therefore if $Lf = 0$ then

$$0 = (Lf, f) = \int_0^1 a(x) f'(x)^2 \, dx$$

and hence $f' \equiv 0$ and since $f(0) = 0$, $f \equiv 0$. This shows L is injective.

2. (Existence) Given $g \in C([0,1], \mathbb{R})$ we are looking for $f \in D(L)$ such that $Lf = g$, i.e. $(af')' = g$. Integrating this equation implies

$$-a(x) f'(x) = -C + \int_0^x g(y) dy.$$

Therefore

$$f'(z) = \frac{C}{a(z)} - \int 1_{y \le z} \frac{1}{a(z)} g(y) dy$$

which upon integration and using $f(0) = 0$ gives

$$f(x) = \int_0^x \frac{C}{a(z)} \, dz - \int 1_{y \le z \le x} \frac{1}{a(z)} \, g(y) \, dz \, dy.$$

If we let

$$\alpha(x) := \int_0^x \frac{1}{a(z)} \, dz \tag{41.3}$$

the last equation may be written as

$$f(x) = C\alpha(x) - \int_0^x (\alpha(x) - \alpha(y)) g(y) \, dy. \tag{41.4}$$

It is a simple matter to work backwards to show the function f defined in Eq. (41.4) satisfies $Lf = g$ and $f(0) = 0$ for any constant C. So it only remains to choose C so that

$$0 = f(1) = C\,\alpha(1) - \int_0^1 (\alpha(1) - \alpha(y))g(y)dy.$$

Solving for C gives $C = \int_0^1 \left(1 - \frac{\alpha(y)}{\alpha(1)}\right) g(y)\ dy$ and the resulting function f may be written as

$$\begin{aligned} f(x) &= \int_0^1 \left[\left(1 - \frac{\alpha(y)}{\alpha(1)}\right)\alpha(x) - 1_{y\le x}(\alpha(x) - \alpha(y))\right] g(y)\ dy \\ &= \int_0^1 G(x,y)g(y)dy \end{aligned}$$

where

$$G(x,y) = \begin{cases} \alpha(x)\left(1 - \frac{\alpha(y)}{\alpha(1)}\right) & \text{if } x \le y \\ \alpha(y)\left(1 - \frac{\alpha(x)}{\alpha(1)}\right) & \text{if } y \le x. \end{cases} \tag{41.5}$$

For example when $a \equiv 1$,

$$G(x,y) = \begin{cases} x\,(1-y) \text{ if } x \le y \\ y\,(1-x) \text{ if } y \le x \end{cases}.$$

■

Definition 41.5. *The function G defined in Eq. (41.5) is called the **Green's function** for the operator $L : D(L) \to C([0,1],\mathbb{R})$.*

Remarks 41.6 *The proof of Theorem 41.4 shows*

$$\left(L^{-1}g\right)(x) := \int_0^1 G(x,y)g(y)dy \tag{41.6}$$

where G is defined in Eq. (41.5). The Green's function G has the following properties:

1. *Since L is invertible and G is a right inverse, G is also a left inverse, i.e. $GLf = f$ for all $f \in D(L)$.*
2. *G is continuous.*
3. *G is symmetric, $G(y,x) = G(x,y)$. (This reflects the symmetry in L, $(Lf,g) = (f,Lg)$ for all $f,g \in D(L)$, which follows from Eq. (41.2).)*
4. *G may be written as*

$$G(x,y) = \begin{cases} u(x)v(y) \text{ if } x \le y \\ u(y)v(x) \text{ if } y \le x. \end{cases}$$

where u and v are L – harmonic functions (i.e. and $Lu = Lv = 0$) with $u(0) = 0$ and $v(1) = 0$. In particular $L_xG(x,y) = 0 = L_yG(x,y)$ for all $y \ne x$.

5. *The first order derivatives of the Green's function have a jump discontinuity on the diagonal. Explicitly,*

$$G_y(x,x+) - G_y(x,x-) = -\frac{1}{a(x)}$$

which follows directly from

$$G_y(x,y) = \frac{1}{a(y)} \begin{cases} -\frac{\alpha(x)}{\alpha(1)} & \text{if } x < y \\ \left(1 - \frac{\alpha(x)}{\alpha(1)}\right) & \text{if } y < x. \end{cases} \tag{41.7}$$

By symmetry we also have

$$G_x(y+,y) - G_x(y-,y) = -\frac{1}{a(y)}.$$

6. *By Items 4. and 5. and Lemma 40.11 it follows that*

$$L_yG(x,y) := L_yT_{G(x,y)} = \frac{d}{dy}\left(a(y)G_y(x,y)\right) = \delta(y-x)$$

and similarly that

$$L_xT_{G(x,y)} = L_xG(x,y) = \delta(x-y).$$

As a consequence of the above remarks we have the following representation theorem for function $f \in C^2([0,1])$.

Theorem 41.7 (Representation Theorem). *For any $f \in C^2([0,1])$,*

$$f(x) = (GLf)(x) - G_y(x,y)a(y)f(y)\Big|_{y=0}^{y=1}. \tag{41.8}$$

Moreover if we are given $h : \partial[0,1] \to \mathbb{R}$ and $g \in C([0,1])$, then the unique solution to

$$Lf = g \text{ with } f = h \text{ on } \partial[0,1]$$

is

$$f(x) = (Gg)(x) - G_y(x,y)a(y)h(y)\Big|_{y=0}^{y=1}. \tag{41.9}$$

Proof. By repeated use of Lemma 40.11,

$$\begin{aligned} (GLf)(x) &= -\int_0^1 G(x,y)\frac{d}{dy}(a(y)f'(y))dy \\ &= \int_0^1 G_y(x,y)a(y)f'(y)dy \\ &= G_y(x,y)a(y)f(y)\Big|_{y=0}^{y=1} + \int_0^1 L_yG(x,y)f(y)dy \\ &= G_y(x,y)a(y)f(y)\Big|_{y=0}^{y=1} + \int_0^1 \delta(x-y)f(y)dy \\ &= G_y(x,y)a(y)f(y)\Big|_{y=0}^{y=1} + f(x) \end{aligned}$$

which proves Eq. (41.8). There are no boundary terms in the second equality above since $G(x,0) = G(x,1) = 0$.

Now suppose that f is defined as in Eq. (41.9). Observe from Eq. (41.7) that

$$\lim_{x\uparrow 1} a(1)G_y(x,1) = -1 \text{ and } \lim_{x\downarrow 0} a(0)G_y(x,0) = 1$$

and also notice that $G_y(x,1)$ and $G_y(x,0)$ are L_x – harmonic functions. Therefore by these remarks and Eq. (41.6), $f = h$ on $\partial[0,1]$ and

$$Lf(x) = g(x) - L_x G_y(x,y)a(y)h(y)\Big|_{y=0}^{y=1} = g(x)$$

as desired. ■

41.2 General Regular 2nd order elliptic ODE

Let $J = [r,s]$ be a closed bounded interval in $\mathbb{R}$.

Definition 41.8. *A second order linear operator of the form*

$$Lf = -af'' + bf' + cf \tag{41.10}$$

with $a \in C^2(J)$, $b \in C^1(J)$ and $c \in C^2(J)$ is said to be ***elliptic*** *if $a > 0$, (more generally if a is invertible if we are allowing for vector valued functions).*

For this section L will denote an elliptic ordinary differential operator. We will now consider the Dirichlet boundary valued problem for $f \in C^2([r,s])$,

$$Lf = -af'' + bf' + cf = 0 \text{ with } f = 0 \text{ on } \partial J. \tag{41.11}$$

Lemma 41.9. *Let $u, v \in C^2(J)$ be two L – harmonic functions, i.e. $Lu = 0 = Lv$ and let*

$$W := \det \begin{bmatrix} u & v \\ u' & v' \end{bmatrix} = uv' - vu'$$

be the Wronskian of u and v. Then W satisfies

$$W' = \frac{b}{a}W, \quad \frac{d}{dx}\frac{1}{W} = -\frac{b}{a}\frac{1}{W} \text{ and}$$
$$W(x) = W(r)e^{\int_r^x \frac{b}{a}(t)dt}.$$

Proof. By direct computation

$$aW' = a(uv'' - vu'') = u(bv' + cv) - v(bu' + cu) = bW.$$

■

Definition 41.10. *Let $H^k(J)$ denote those $f \in C^{k-1}(J)$ such that $f^{(k-1)}$ is absolutely continuous and $f^{(k)} \in L^2(J)$. We also let $H_0^2(J) = \{f \in H^2(J) : f|_{\partial J} =$ We make $H^k(J)$ into a Hilbert space using the following inner product*

$$(u,v)_{H^k} := \sum_{j=0}^k (D^j u, D^j v)_{L^2}.$$

Theorem 41.11. *As above, let $D(L) = \{f \in C^2(J) : f = 0 \text{ on } \partial J\}$. If the* $\mathrm{Nul}(L) \cap D(L) = \{0\}$, *i.e. if the only solution $f \in D(L)$ to $Lf = 0$ is $f = 0$, then $L : D(L) \to C(J)$ is an invertible. Moreover there exists a continuous function G on $J \times J$ (called the Dirichlet Green's function for L) such that*

$$(L^{-1}g)(x) = \int_J G(x,y)g(y)dy \text{ for all } g \in C(J). \tag{41.12}$$

Moreover if $g \in L^2(J)$ then $Gg \in H_0^2(J)$ and $L(G\ g) = g$ a.e. and more generally if $g \in H^k(J)$ then $Gg \in H_0^{k+2}(J)$

Proof. To prove the surjectivity of $L : D(L) \to C(J)$, (i.e. existence of solutions $f \in D(L)$ to $Lf = g$ with $g \in C(J)$) we are going to construct the Green's function G.

1. **Formal requirements on the Greens function.** Assuming Eq. (41.12) holds and working formally we should have

$$g(x) = L_x \int_J G(x,y)g(y)dy = \int_J L_x G(x,y)g(y)dy \tag{41.13}$$

for all $g \in C(J)$. Hence, again formally, this implies

$$L_x G(x,y) = \delta(y-x) \text{ with } G(r,y) = G(s,y) = 0. \tag{41.14}$$

This can be made more convincing by as follows. Let $\phi \in \mathcal{D} := \mathcal{D}(r,s)$, then multiplying

$$g(x) = L_x \int_J G(x,y)g(y)dy$$

by ϕ, integrating the result and then using integration by parts and Fubini's theorem gives

$$\begin{aligned}\int_J g(x)\phi(x)dx &= \int_J dx\phi(x)L_x \int_J dyG(x,y)g(y) \\ &= \int_J dxL_x\phi(x)\int_J dyG(x,y)g(y) \\ &= \int_J dyg(y)\int_J dx\ L_x\phi(x)G(x,y) \text{ for all } g \in C(J).\end{aligned}$$

From this we conclude

$$\int_J L_x\phi(x)G(x,y)dx = \phi(y),$$

i.e. $L_x T_{G(x,y)} = \delta(x-y)$.

2. **Constructing** G. In order to construct a solution to Eq. (41.14), let u,v be two non-zero L – harmonic functions chosen so that $u(r) = 0 = v(s)$ and $u'(r) = 1 = v'(s)$ and let W be the Wronskian of u and v. By Lemma 41.9, either W is never zero or is identically zero. If $W = 0$, then $(u(r),u'(r)) = \lambda(v(r),v'(r))$ for some $\lambda \in \mathbb{R}$ and by uniqueness of solutions to ODE it would follow that $u \equiv \lambda v$. In this case $u(r) = 0$ and $u(s) = \lambda v(s) = 0$, and hence $u \in D(L)$ with $Lu = 0$. However by assumption, this implies $u = 0$ which is impossible since $u'(0) = 1$. Thus W is never 0.
By Eq. (41.14) we should require $L_x G(x,y) = 0$ for $x \neq y$ and $G(r,y) = G(s,y) = 0$ which implies that

$$G(x,y) = \begin{cases} u(x)\phi(y) \text{ if } x < y \\ v(x)\psi(y) \text{ if } x > y \end{cases}$$

for some functions ϕ and ψ. We now want to choose ϕ and ψ so that G is continuous and $L_x G(x,y) = \delta(x-y)$. Using

$$G_x(x,y) = \begin{cases} u'(x)\phi(y) \text{ if } x < y \\ v'(x)\psi(y) \text{ if } x > y \end{cases}$$

Lemma 41.9, we are led to require

$$0 = G(y+,y) - G(y-,y) = u(y)\phi(y) - v(y)\psi(y)$$
$$1 = -\left[a(x)G_x(x,y)\right]\Big|_{x=y-}^{x=y+} = -a(y)\left[v'(y)\psi(y) - u'(y)\phi(y)\right].$$

Solving these equations for ϕ and ψ gives

$$\begin{pmatrix}\phi \\ \psi\end{pmatrix} = -\frac{1}{aW}\begin{pmatrix} v \\ u\end{pmatrix}$$

and hence

$$G(x,y) = -\frac{1}{a(y)W(y)}\begin{cases} u(x)v(y) \text{ if } x \leq y \\ v(x)u(y) \text{ if } x \geq y. \end{cases} \tag{41.15}$$

3. **With this** G**, Eq. (41.12) holds.** Given $g \in C(J)$, then f in Eq. (41.12) may be written as

$$f(x) = \int_J G(x,y)g(y)dy$$
$$= -v(x)\int_r^x \frac{u(y)}{a(y)W(y)}g(y)dy - u(x)\int_x^s \frac{v(y)}{a(y)W(y)}g(y)dy. \tag{41.16}$$

Differentiating this equation twice gives

$$f'(x) = -v'(x)\int_r^x \frac{u(y)}{a(y)W(y)}g(y)dy - u'(x)\int_x^s \frac{v(y)}{a(y)W(y)}g(y)dy \tag{41.17}$$

and

$$f''(x) = -v''(x)\int_r^x \frac{u(y)}{a(y)W(y)}g(y)dy - u''(x)\int_x^s \frac{v(y)}{a(y)W(y)}g(y)dy$$
$$- v'(x)\frac{u(x)}{a(x)W(x)}g(x) + u'(x)\frac{v(x)}{a(x)W(x)}g(x). \tag{41.18}$$

Using $Lv = 0 = Lu$, the definition of W and the last two equations we find

$$-a(x)f''(x) = \left[b(x)v'(x) + c(x)v(x)\right]\int_r^x \frac{u(y)}{a(y)W(y)}g(y)dy$$
$$+ \left[b(x)u'(x) + c(x)u(x)\right]\int_x^s \frac{v(y)}{a(y)W(y)}g(y)dy + g(x)$$
$$= -b(x)f'(x) - c(x)f(x) + g(x),$$

i.e. $Lf = g$.

Hence we have proved $L : D(L) \to C(J)$ is surjective and $L^{-1} : C(J) \to D(L)$ is given by Eq. (41.12).

Now suppose $g \in L^2(J)$, we will show that $f \in C^1(J)$ and Eq. (41.17) is still valid. The difficulty here is that it is clear that f is differentiable almost everywhere and Eq. (41.17) holds for almost every x. However this is not good enough, we need Eq. (41.17) to hold for all x. To remedy this, choose $g_n \in C(J)$ such that $g_n \to g$ in $L^2(J)$ and let $f_n := Gg_n$. Then by what we have just proved,

$$f_n'(x) = \int_J G_x(x,y)g_n(y)dy$$

Now by the Cauchy-Schwarz inequality,

$$\left|\int_J G_x(x,y)\left[g(y) - g_n(y)\right]dy\right|^2 \leq \|g - g_n\|_{L^2(J)}^2 \int_J |G_x(x,y)|^2 dy$$
$$\leq C\,\|g - g_n\|_{L^2(J)}^2$$

where $C := \sup_{x\in J}\int_J |G_x(x,y)|^2 dy < \infty$. From this inequality it follows that $f_n'(x)$ converges uniformly to $\int_J G_x(x,y)g(y)dy$ as $n \to \infty$ and hence $f \in C^1(J)$ and

$$f'(x) = \int_J G_x(x,y)g(y)dy \text{ for all } x \in J,$$

i.e. Eq. (41.17) is valid for all $x \in J$. It now follows from Eq. (41.17) that $f \in H^2(J)$ and Eq. (41.18) holds for almost every x. Working as before we may conclude $Lf = g$ a.e. Finally if $g \in H^k(J)$ for $k \geq 1$, the reader may easily show $f \in H_0^{k+2}(J)$ by examining Eqs. (41.17) and (41.18). ■

Remark 41.12. When L is given as in Eq. (41.1), $b = -a'$ and by Lemma 41.9

$$W(x) = W(0)e^{-\int_0^x \frac{a'}{a}(t)dt} = W(0)e^{-\ln(a(x)/a(0))} = \frac{W(0)a(0)}{a(x)}.$$

So in this case

$$G(x,y) = -\frac{1}{W(0)a(0)} \begin{cases} u(x)v(y) \text{ if } x \le y \\ v(x)u(y) \text{ if } x \ge y \end{cases}$$

where we may take

$$u(x) = \alpha(x) := \int_0^x \frac{1}{a(z)}dz \text{ and } v(x) = \left(1 - \frac{\alpha(x)}{\alpha(1)}\right).$$

Finally for this choice of u and v we have

$$W(0) = u(0)v'(0) - u'(0)v(0) = -\frac{1}{a(0)}$$

giving

$$G(x,y) = \begin{cases} u(x)v(y) \text{ if } x \le y \\ v(x)u(y) \text{ if } x \ge y \end{cases}$$

which agrees with Eq. (41.5) above.

Lemma 41.13. *Let $L^* f := -(af)'' - (bf)' + cf$ be the* ***formal adjoint*** *of L. Then*

$$(Lf, g) = (f, L^* g) \text{ for all } f, g \in D(L) \tag{41.19}$$

where $(f,g) := \int_J f(x)g(x)dx$. Moreover if $\mathrm{nul}(L) = \{0\}$ then $\mathrm{nul}(L^*) = \{0\}$ and the Greens function for L^* is G^* defined by $G^*(x,y) = G(y,x)$, where G is the Green's function in Eq. (41.15). Consequently $L_y^* G(x,y) = \delta(x-y)$.

Proof. First observe that G^* has been defined so that $(G^* g, f) = (g, Gf)$ for all $f \in L^2(J)$.Eq. (41.19) follows by two integration by parts after observing the boundary terms are zero because $f = g = 0$ on ∂J. If $g \in \mathrm{nul}(L^*)$ and $f \in D(L)$, we find

$$0 = (L^* g, f) = (g, Lf) \text{ for all } f \in D(L).$$

By Theorem 41.11, if $\mathrm{nul}(L) = \{0\}$ then $L : D(L) \to C(J)$ is invertible so the above equation implies $\mathrm{nul}(L^*) = \{0\}$. Another application of Theorem 41.11 then shows $L^* : D(L) \to C(J)$ is invertible and has a Green's function which we call $\tilde{G}(x,y)$. We will now complete the proof by showing $\tilde{G} = G^*$. To do this observe that

$$(f,g) = (L^* \tilde{G} f, g) = (\tilde{G} f, Lg) = (f, \tilde{G}^* Lg) \text{ for all } f, g \in D(L)$$

and this then implies $\tilde{G}^* L = Id_{D(L)} = GL$. Cancelling the L from this equation, show $\tilde{G}^* = G$ or equivalently that $\tilde{G} = G^*$. The remaining assertions of the Lemma follows from this observation.

Here is an **alternate proof** that $L_y^* G(x,y) = \delta(x-y)$, also see Using $GL = I_{D(L)}$, we learn for $u \in D(L)$ and $v \in C(J)$ that

$$(v,u) = (v, GLu) = (L^* G^* v, u)$$

which then implies $L^* G^* v = v$ for all $v \in C(J)$. This implies

$$f(x) = \int_J G(x,y) Lf(y) dy = \langle T_{G(x,\cdot)}, Lf \rangle = \langle L^* T_{G(x,\cdot)}, f \rangle \text{ for all } f \in D(L)$$

from which it follows that $L_y^* T_{G(x,y)} = \delta(x-y)$. ■

Definition 41.14. *A* ***Green's function for*** *L is a function $G(x,y)$ as defined as in Eq. (41.15) where u and v are* ***any*** *two linearly independent L – harmonic functions.*[1]

The following theorem in is a generalization of Theorem 41.7.

Theorem 41.15 (Representation Theorem). *Suppose and G is a Green's function for L then*

1. *$L_x T_{G(x,y)} = \delta(x-y)$ and $LG = I$ on $L^2(J)$. (However Gg and $G^* g$ may no longer satisfy the given Dirichlet boundary conditions.)*
2. *$L_y^* T_{G(x,y)} = \delta(x-y)$. More precisely we have the following representation formula. For any $f \in H^2(J)$,*
$$f(x) = (GLf)(x) + \left\{ G(x,y)a(y)f'(y) - [a(y)G(x,y)]_y f(y) \right\} \Big|_{y=r}^{y=s}. \tag{41.20}$$
3. *Let us now assume $\mathrm{nul}(L) = \{0\}$ and G is the Dirichlet Green's function for L. The Eq. (41.20) specializes to*
$$f(x) = (GLf)(x) - [a(y)G(x,y)]_y f(y) \Big|_{y=r}^{y=s}.$$
Moreover if we are given $h : \partial J \to \mathbb{R}$ and $g \in L^2(J)$, then the unique solution $f \in H^2(J)$ to
$$Lf = g \text{ a.e. with } f = h \text{ on } \partial J$$
is
$$f(x) = (Gg)(x) + H(x) \tag{41.21}$$
where, for $x \in J^0$,
$$H(x) := -[a(y)G(x,y)]_y h(y) \Big|_{y=r}^{y=s} \tag{41.22}$$
and $H(r) := H(r+)$ and $H(s) := H(s-)$.

[1] For example choose u, v so that $Lu = 0 = Lv$ and $u(\alpha) = v'(\alpha) = 0$ and $u'(\alpha) = v(\alpha) = 1$.

Proof. 1. The first item follows from the proof of Theorem 41.11 with out any modification.

2. Using Lemma 41.9,

$$\begin{aligned} L^*\left(\frac{u}{aW}\right) &= -(\frac{u}{W})'' - (\frac{bu}{aW})' + \frac{cu}{aW} \\ &= -(\frac{u'}{W} - \frac{b}{a}\frac{1}{W}u)' - (\frac{bu}{aW})' + \frac{cu}{aW} \\ &= -(\frac{u'}{W})' + \frac{cu}{aW} = -(\frac{u''}{W} - \frac{b}{a}\frac{1}{W}u) + \frac{cu}{aW} \\ &= \frac{1}{a}Lu = 0. \end{aligned}$$

Similarly $L^*(\frac{v}{aW}) = 0$ and therefore $L_y^* G(x,y) = 0$ for $y \neq x$. Since

$$\begin{aligned} G_y(x,y) = &-\left(\frac{d}{dy}\frac{1}{a(y)W(y)}\right)\begin{cases} u(x)v(y) \text{ if } x \le y \\ v(x)u(y) \text{ if } x \ge y \end{cases} \\ &- \frac{1}{a(y)W(y)}\begin{cases} u(x)v'(y) \text{ if } x \le y \\ v(x)u'(y) \text{ if } x \ge y \end{cases} \end{aligned} \tag{41.23}$$

we find

$$\begin{aligned} G_y(x,x+) - G_y(x,x-) &= \frac{1}{a(x)W(x)}\{v(x)u'(x) - u(x)v'(x)\} \\ &= -\frac{1}{a(x)}. \end{aligned}$$

Finally since

$$L_y^* = -a\frac{d^2}{dy^2} + \text{lower order terms}$$

we may conclude form Lemma 40.11 that $L_y^* G(x,y) = \delta(x-y)$.Using integration by parts for absolutely continuous functions and Lemma 41.13, for $f \in H^2(J)$,

$$\begin{aligned} (GLf)(x) &= \int_J G(x,y)Lf(y)dy \\ &= \int_J G(x,y)\left(-a(y)\frac{d^2}{dy^2} + b(y)\frac{d}{dy} + c(y)\right)f(y)dy \\ &= \int_J \left[\begin{array}{c} \frac{d}{dy}[a(y)G(x,y)]\, f'(y) \\ + \left(-\frac{d}{dy}[b(y)G(x,y)]\ f + c(y)\right) f(y) \end{array}\right] dy \\ &- G(x,y)a(y)f'(y)|_{y=r}^{y=s} \\ &= -G(x,y)a(y)f'(y)|_{y=r}^{y=s} + [a(y)G(x,y)]_y\, f(y)|_{y=r}^{y=s} \\ &+ \langle L_y^* G(x,y), f(y)\rangle \\ &= [a(y)G(x,y)]_y\, f(y)|_{y=r}^{y=s} - G(x,y)a(y)f'(y)|_{y=r}^{y=s} + f(x). \end{aligned}$$

This proves Eq. (41.20).

3. Now suppose G is the Dirichlet Green's function for L. By Eq. (41.15),

$$\begin{aligned} [-a(y)G(x,y)]_y &= \left(\frac{d}{dy}\frac{1}{W(y)}\right)\begin{cases} u(x)v(y) \text{ if } x \le y \\ v(x)u(y) \text{ if } x \ge y \end{cases} \\ &+ \frac{1}{W(y)}\begin{cases} u(x)v'(y) \text{ if } x \le y \\ v(x)u'(y) \text{ if } x \ge y \end{cases} \end{aligned}$$

and hence the function H defined in Eq. (41.22) is more explicitly given by

$$H(x) = \frac{1}{W(s)}\left(u(x)v'(s)\right)h(s) - \frac{1}{W(r)}\left(v(x)u'(r)\right)h(r). \tag{41.24}$$

From this equation or the fact that $L_x G(x,r) = 0 = L_x G(x,s)$, H is is $L-$ harmonic on J^0. Moreover, from Eq. (41.24),

$$\begin{aligned} H(r) &= -\frac{1}{W(r)}\left(v(r)u'(r)\right)h(r) \\ &= \frac{1}{W(r)}\left(u(r)v'(r) - v(r)u'(r)\right)h(r) = h(r) \end{aligned}$$

and

$$\begin{aligned} H(s) &= \frac{1}{W(s)}\left(u(s)v'(s)\right)h(s) \\ &= \frac{1}{W(s)}\left(u(s)v'(s) - v(s)u'(s)\right)h(s) = h(s). \end{aligned}$$

Therefore if f is defined by Eq. (41.21),

$$Lf = LGg - LH = g \text{ a.e. on } J^0$$

because $LG = I$ on $L^2(J)$ and

$$f|_{\partial J} = (Gg)|_{\partial J} + H|_{\partial J} = H|_{\partial J} = h$$

since $Gg \in H_0^2(J)$. ■

Corollary 41.16 (Elliptic Regularity I). *Suppose $-\infty \le r_0 < s_0 \le \infty$, $J_0 := (r_0, s_0)$ and L is as in Eq. (41.11) with the further assumption that $a, b, c \in C^\infty(\mathbb{R})$. If $f \in C^2(J_0)$ is a function such that $g := Lf \in C^k(J_0)$ for some $k \ge 0$, then $f \in C^{k+2}(J_0)$.*

Proof. Let $r < s$ be chosen so that $J := [r,s]$ is a bounded subinterval of J_0 and let G be a Green's function as in Definition 41.14. Since a, b, c are smooth, it follows from our general theory of ODE that $G(x,y) \in C^\infty(J \times J \setminus \Delta)$ where $\Delta = \{(x,x) : x \in J\}$ is the diagonal in $J \times J$. Now by Theorem 41.15, for $x \in J^0$,

$$f(x) = (Gg)(x) + \left\{ G(x,y)a(y)f'(y) - [a(y)G(x,y)]_y f(y) \right\} \Big|_{y=r}^{y=s}.$$

Since

$$x \to \left\{ G(x,y)a(y)f'(y) - [a(y)G(x,y)]_y f(y) \right\} \Big|_{y=r}^{y=s} \in C^\infty(J^0)$$

it suffices to show $Gg \in C^{k+2}(J^0)$. But this follows by examining the formula for $(Gg)''$ given on the right side of Eq. (41.18). ■

In fact we have the following rather striking version of this result.

Theorem 41.17 (Hypoellipticity). *Suppose* $-\infty \le r_0 < s_0 \le \infty$, $J_0 := (r_0, s_0)$ *and L is as in Eq. (41.11) with the further assumption that* $a, b, c \in C^\infty(\mathbb{R})$. *If* $u \in \mathcal{D}'(J_0)$ *is a generalized function such that* $v := Lu \in C^\infty(J_0)$, *then* $u \in C^\infty(J_0)$.

Proof. As in the proof of Corollary 41.16 let $r < s$ be chosen so that $J := [r,s]$ is a bounded subinterval of J_0 and let G be the Green's function constructed above.[2] Further suppose $\xi \in J^0$, $\theta \in C_c^\infty(J^0, [0,1])$ such that $\theta = 1$ in a neighborhood U of ξ and $\alpha \in C_c^\infty(V, [0,1])$ such that $\alpha = 1$ in a neighborhood V of ξ, see Figure 41.1. Finally suppose that $\phi \in C_c^\infty(V)$, then

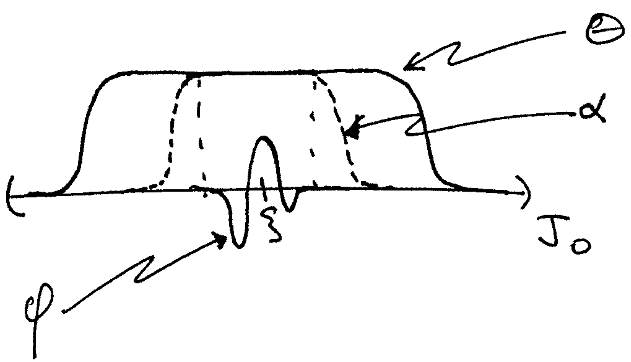


Fig. 41.1. Constructing the cutoff functions, θ and α.

$$\begin{aligned}\phi = \theta\phi = \theta L^* G^* \phi &= \theta L^* \left(M_\alpha + M_{1-\alpha}\right) G^* \phi \\ &= L^* M_\alpha G^* \phi + \theta L^* M_{1-\alpha} G^* \phi\end{aligned}$$

and hence

[2] Actually we can simply define G^* to be a Green's function for L^*. It is not necessary to know $G^*(x,y) = G(y,x)$ where G is a Green's function for L.

$$\begin{aligned}\langle u, \phi\rangle &= \langle u, L^* M_\alpha G^* \phi + \theta L^* M_{1-\alpha} G^* \phi\rangle \\ &= \langle Lu, M_\alpha G^* \phi\rangle + \langle u, \theta L^* M_{1-\alpha} G^* \phi\rangle.\end{aligned}$$

Now

$$\langle Lu, M_\alpha G^* \phi\rangle = \langle v, M_\alpha G^* \phi\rangle = \langle G M_\alpha v, \phi\rangle$$

and writing $u = D^n T_h$ for some continuous function h (which is always possible locally) we find

$$\begin{aligned}&\langle u, \theta L^* M_{1-\alpha} G^* \phi\rangle \\ &= (-1)^n \langle u, D^n M_\theta L^* M_{1-\alpha} G^* \phi\rangle \\ &= (-1)^n \int_{J\times J} h(x) D_x^n \left[\theta(x) L_x^* (1-\alpha(x)) G(y,x)\right] \phi(y) dy dx \\ &= \int_J \psi(y)\phi(y) dy\end{aligned}$$

where

$$\psi(y) := \int_J h(x) D_x^n \left[\theta(x) L_x^* (1-\alpha(x)) G(y,x)\right] dx$$

which is smooth for $y \in V$ because $1-\alpha(x) = 0$ on V and so $(1-\alpha(x))\, G(y,x)$ is smooth for $(x,y) \in J \times V$. Putting this altogether shows

$$\langle u, \phi\rangle = \langle G M_\alpha v + \psi, \phi\rangle \text{ for all } \phi \in C_c^\infty(V).$$

That is to say $u = G M_\alpha v + \psi$ on V which proves the theorem since $G M_\alpha v + \psi \in C^\infty(V)$. ■

Example 41.18. Let $L = \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial x^2}$ be the wave operator on $\mathbb{R}^2$ which is not elliptic. Given $f \in C^2(\mathbb{R})$ we have already seen that $Lf(y-x) = 0 \in C^\infty(\mathbb{R}^2)$. Clearly since f was arbitrary, it does not follow that $F(x,y) := f(y-x) \in C^\infty(\mathbb{R}^2)$. Moreover, if f is merely continuous and $F(x,y) := f(y-x)$, then $LT_F = 0$ with $F \notin C^2(\mathbb{R}^2)$. To check $LT_F = 0$ we first observe

$$\begin{aligned}-\langle (\partial_x + \partial_y) T_F, \phi\rangle &= \langle T_F, (\partial_x + \partial_y)\phi\rangle \\ &= \int_{\mathbb{R}^2} f(y-x)\, (\partial_x + \partial_y)\, \phi(x,y) dx dy \\ &= \int_{\mathbb{R}^2} f(y) \left[\phi_x(x, y+x) + \phi_y(x, y+x)\right] dx dy \\ &= \int_{\mathbb{R}^2} f(y) \frac{\partial}{\partial x} \left[\phi(x, y+x)\right] dx dy = 0.\end{aligned}$$

Therefore $LT_F = (\partial_x - \partial_y)(\partial_x + \partial_y) T_F = 0$ as well.

Corollary 41.19. *Suppose a, b, c are smooth and* $u \in \mathcal{D}'(J^0)$ *is an eigenvector for L, i.e.* $Lu = \lambda u$ *for some* $\lambda \in \mathbb{C}$. *Then* $u \in C^\infty(J)$.

Proof. Since $L - \lambda$ is an elliptic ordinary differential operator and $(L-\lambda)\, u = 0 \in C^\infty(J^0)$, it follows by Theorem 41.17 that $u \in C^\infty(J^0)$. ■

41.3 Elementary Sobolev Inequalities

Notation 41.20 *Let* $\overline{\int_J} f dm := \frac{1}{|J|}\int_J f dm$ *denote the average of* f *over* $J = [r,s]$.

Proposition 41.21. *For* $f \in H^1(J)$,

$$\begin{aligned}|f(x)| &\le \left|\overline{\int_J} f dm\right| + \|f'\|_{L^1(J)} \\ &\le \left|\overline{\int_J} f dm\right| + \sqrt{|J|}\left(\int_J |f'(y)|^2 dy\right)^{1/2} \le C(|J|)\|f\|_{H^1(J)}.\end{aligned}$$

where $C(|J|) = \max\left(\frac{1}{\sqrt{|J|}}, \sqrt{|J|}\right)$.

Proof. By the fundamental theorem of calculus for absolutely continuous functions

$$f(x) = f(a) + \int_a^x f'(y)dy$$

for any $a, x \in J$. Integrating this equation on a and then dividing by $|J| := s-r$ implies

$$f(x) = \overline{\int_J} f dm + \overline{\int_J} da \int_a^x f'(y)dy$$

and hence

$$\begin{aligned}|f(x)| &\le \left|\overline{\int_J} f dm\right| + \overline{\int_J} da \left|\int_a^x |f'(y)| dy\right| \\ &\le \left|\overline{\int_J} f dm\right| + \int_J |f'(y)| dy \\ &\le \left|\overline{\int_J} f dm\right| + \sqrt{|J|}\left(\int_J |f'(y)|^2 dy\right)^{1/2} \\ &\le \frac{1}{\sqrt{|J|}}\left(\int_J |f|^2 dm\right)^{1/2} + \sqrt{|J|}\left(\int_J |f'(y)|^2 dy\right)^{1/2}.\end{aligned}$$

■

Notation 41.22 *For the remainder of this section, suppose* $Lf = -\frac{1}{\rho}D(\rho a f') + cf$. *is an elliptic ordinary differential operator on* $J = [r,s]$, $\rho \in C^2(J,(0,\infty))$ *is a* ***positive weight*** *and*

$$(f,g)_\rho := \int_J f(x)g(x)\rho(x)dx.$$

We will also take $D(L) = H_0^2(J)$, *so that we are imposing Dirichlet boundary conditions on* L. *Finally let*

$$\mathcal{E}(f,g) := \int_J [af'g' + cfg]\,\rho dm \text{ for } f,g \in H^1(J).$$

Lemma 41.23. *For* $f, g \in D(L)$,

$$(Lf,g)_\rho = \mathcal{E}(f,g) = (f,Lg)_\rho. \tag{41.25}$$

Moreover

$$\mathcal{E}(f,f) \ge a_0\|f'\|_2^2 + c_0\|f\|_2^2 \text{ for all } f \in H^1(J)$$

where $c_0 := \min_J c$ *and* $a_0 = \min_J a$. *If* $\lambda_0 \in \mathbb{R}$ *with* $\lambda_0 + c_0 > 0$ *then*

$$\|f\|_{H^1(J)}^2 \le K\left[\mathcal{E}(f,f) + \lambda_0\|f\|_2^2\right] \tag{41.26}$$

where $K = [\min(a_0, c_0+\lambda_0)]^{-1}$.

Proof. Eq. (41.25) is a simple consequence of integration by parts. By elementary estimates

$$\mathcal{E}(f,f) \ge a_0\|f'\|_2^2 + c_0\|f\|_2^2$$

and

$$\mathcal{E}(f,f) + \lambda_0\|f\|_2^2 \ge a_0\|f'\|_2^2 + (c_0+\lambda_0)\|f\|_2^2 \ge \min(a_0, c_0+\lambda_0)\|f\|_{H^1(J)}^2$$

which proves Eq. (41.26). ■

Corollary 41.24. *Suppose* $\lambda_0 + c_0 > 0$ *then* $\mathrm{Nul}(L+\lambda_0) \cap D(L) = 0$ *and hence*

$$(L+\lambda_0) : H_0^2(J) \to L^2(J)$$

is invertible and the ***resolvent*** $(L+\lambda_0)^{-1}$ *has a continuous integral kernel* $G(x,y)$, *i.e.*

$$(L+\lambda_0)^{-1}u(x) = \int_J G(x,y)u(y)dy.$$

Moreover if we define $D(L^k)$ *inductively by*

$$D(L^k) := \left\{u \in D(L^{k-1}) : L^{k-1}u \in D(L)\right\}$$

we have $D(L^k) = H_0^{2k}(J)$.

Proof. By Lemma 41.23, for all $u \in D(L)$,

$$\|u\|_{H^1(J)}^2 \le K\left((Lu,u) + \lambda_0\|u\|_2^2\right) = K\left(((L+\lambda_0)u,u)\right)$$

so that if $(L+\lambda_0)u = 0$, then $\|u\|_{H^1(J)}^2 = 0$ and hence $u = 0$. The remaining assertions except for $D(L^k) = H_0^k(J)$ now follow directly from Theorem 41.11

applied with L replaced by $L+\lambda_0$. Finally if $u \in D(L)$ then $(L+\lambda_0)\,u = Lu + \lambda_0 u \in L^2(J)$ and therefore

$$u = (L+\lambda_0)^{-1}(Lu + \lambda_0 u) \in H_0^2(J).$$

Now suppose we have shown, $D(L^k) = H_0^{2k}(J)$ and $u \in D\left(L^{k+1}\right)$, then

$$(L+\lambda_0)\,u = Lu + \lambda_0 u \in D(L^k) + D(L^{k+1}) \subset D(L^k) = H_0^{2k}(J)$$

and so by Theorem 41.11, $u \in (L+\lambda_0)^{-1} H_0^{2k}(J) \subset H_0^{2k+2}(J)$. ■

Corollary 41.25. *There exists an orthonormal basis* $\{\phi_n\}_{n=0}^\infty$ *for* $L^2(J, \rho dm)$ *of eigenfunctions of* L *with eigenvalues* $\lambda_n \in \mathbb{R}$ *such that* $-c_0 \le \lambda_0 < \lambda_1 < \lambda_2 < \dots.$

Proof. Let $\lambda_0 > -c_0$ and let $G := (L+\lambda_0)^{-1} : L^2(J) \to H_0^2(J) = D(L) \subset L^2(J)$. From the theory of compact operators to be developed later, G is a compact symmetric positive definite operator on $L^2(J)$ and hence there exists an orthonormal basis $\{\phi_n\}_{n=0}^\infty$ for $L^2(J, \rho dm)$ of eigenfunctions of G with eigenvalues $\mu_n > 0$ such that $\mu_0 \ge \mu_1 \ge \mu_2 \ge \dots \to 0$.[3] Since

$$\mu_n \phi_n = G\phi_n = (L+\lambda_0)^{-1}\phi_n,$$

it follows that $\mu_n (L+\lambda_0)\,\phi_n = \phi_n$ for all n and therefore $L\phi_n = \lambda_n \phi_n$ with $\lambda_n = (\mu_n^- 1 - \lambda_0) \uparrow \infty$. Finally since L is a second order ordinary differential equation there can be at most one linearly independent eigenvector for a given eigenvalue λ_n and hence $\lambda_n < \lambda_n + 1$ for all n. ■

Example 41.26. Let $J = [0,\pi]$, $\rho = 1$ and $L = -D^2$ on $H_0^2(J)$. Then $L\phi = \lambda\phi$ implies $\phi'' + \lambda\phi = 0$. Since L is positive, we need only consider the case where $\lambda \ge 0$ in which case $\phi(x) = a\cos\left(\sqrt{\lambda}x\right) + b\sin\left(\sqrt{\lambda}x\right)$. The boundary conditions for f imply $a = 0$ and $0 = \sin\left(\sqrt{\lambda}\pi\right)$, i.e. $\sqrt{\lambda} \in \mathbb{N}_+$. Therefore in this example

$$\phi_k(x) = \sqrt{\frac{2}{\pi}}\sin(kx) \text{ with } \lambda_k = k^2.$$

The collection of functions $\{\phi_k\}_{k=1}^\infty$ is an orthonormal basis for $L^2(J)$.

Theorem 41.27. *Let* $J = [r,s]$ *and* $\rho, a \in C^2(J,(0,\infty))$, $c \in C^2(J)$ *and* L *be defined by*

$$Lf = -\frac{1}{\rho}D(\rho a f') + cf.$$

[3] In fact G is "Hilbert Schmidt" which then implies

$$\sum_{n=0}^\infty \mu_n^2 < \infty.$$

and for $\lambda \in \mathbb{R}$ *let*

$$E^\lambda := \left\{\phi \in H_0^2(J) : L\phi = \alpha\phi \text{ for some } \alpha < \lambda\right\}.$$

Then there are constants $d_1, d_2 > 0$ *such that*

$$\dim(E^\lambda) \le d_1\lambda + d_2. \tag{41.27}$$

Proof. For $\lambda \in \mathbb{R}$ let $E_\lambda := \left\{\phi \in H_0^2(J) : L\phi = \lambda\phi\right\}$. By Corollary 41.24, $E_\lambda = \{0\}$ if $\lambda < c_0$ and since $(Lf,g)_\rho = (f,Lg)_\rho$ for all $f,g \in H_0^2(J)$ it follows that $E_\lambda \perp E_\beta$ for all $\lambda \ne \beta$. Indeed, if $f \in E_\lambda$ and $g \in E_\beta$, then

$$(\beta - \lambda)(f,g)_\rho = (f,Lg)_\rho - (Lf,g)_\rho = 0.$$

Thus it follow that any finite dimensional subspace $W \subset E^\lambda$ has an orthonormal basis (relative to $(\cdot,\cdot)_\rho$ – inner product) of eigenvectors $\{\phi_k\}_{k=1}^n \subset E^\lambda$ of L, say $L\phi_k = \lambda_k\phi_k$. Let $u = \sum_{k=1}^n u_k\phi_k$ where $u_k \in \mathbb{R}$. By Proposition 41.21 and Lemma 41.23,

$$\|u\|_u^2 \le C\|u\|_{H^1(J)}^2 \le C((L+\lambda_0)\,u,u)_\rho = C\left(\sum_{k=1}^n u_k(\lambda_k+\lambda_0)\,\phi_k, u\right)_\rho$$

(where C is a constant varying from place to place but independent of u) and hence for any $x \in J$,

$$\left|\sum_{k=1}^n u_k\phi_k(x)\right|^2 \le \|u\|_u^2 \le C(\lambda+\lambda_0)\sum_{k=1}^n |u_k|^2.$$

Now choose $u_k = \phi_k(x)$ in this equation to find

$$\left|\sum_{k=1}^n |\phi_k(x)|^2\right|^2 \le C(\lambda+\lambda_0)\sum_{k=1}^n |\phi_k(x)|^2$$

or equivalently that

$$\sum_{k=1}^n |\phi_k(x)|^2 \le C(\lambda+\lambda_0).$$

Multiplying this equation by ρ and then integrating shows

$$\dim(W) = n = \sum_{k=1}^n (\phi_k,\phi_k)_\rho \le C(\lambda+\lambda_0)\int_J \rho dm = C'(\lambda+\lambda_0).$$

Since $W \subset E^\lambda$ is arbitrary, it follows that

$$\dim(E^\lambda) \le C'(\lambda+\lambda_0).$$

■

Remarks 41.28 *Notice that for all* $\lambda \in \mathbb{R}$, $\dim(E_\lambda) \le 1$ *because if* $u, v \in E_\lambda$ *then by uniqueness of solutions to ODE,* $u = [u'(r)/v'(r)]\,v$. *Let* $\{\phi_k\}_{k=1}^\infty \subset H_0^2(J) \cap C^\infty(J)$ *be the eigenvectors of* L *ordered so that the corresponding eigenvalues are increasing. With this ordering we have* $k = \dim(E^{\lambda_k}) \le d_1\lambda_k + d_2$ *and therefore,*

$$\lambda_k \ge d_1^{-1}(k - d_2). \tag{41.28}$$

The estimates in Eqs. (41.27) and (41.28) are not particularly good as Example 41.26 illustrates.

41.4 Associated Heat and Wave Equations

Lemma 41.29. L *is a closed operator, i.e. if* $s_n \in D(L)$ *and* $s_n \to s$ *and* $Ls_n \to g$ *in* L^2, *then* $s \in D(L)$ *and* $Ls = g$. *In particular if* $f_k \in D(L)$ *and* $\sum_{k=1}^\infty f_k$ *and* $\sum_{k=1}^\infty Lf_k$ *exists in* L^2, *then* $\sum_{k=1}^\infty f_k \in D(L)$ *and*

$$L\sum_{k=1}^\infty f_k = \sum_{k=1}^\infty Lf_k.$$

Proof. Let $\lambda_0 + c_0 > 0$ and $G = (L+\lambda_0)^{-1}$. Then by assumption $(L+\lambda_0)\,s_n \to g + \lambda_0 s$ and so

$$s \leftarrow s_n = G\,(L+\lambda_0)\,s_n \to G\,(g+\lambda_0 s) \text{ as } n \to \infty$$

showing $s = Gg \in D(L+\lambda_0) = D(L)$ and

$$(L+\lambda_0)\,s = (L+\lambda_0)\,G\,(g+\lambda_0 s) = g + \lambda_0 s$$

and hence $Ls = g$ as desired. The assertions about the sums follow by applying the sequence results to $s_n = \sum_{k=1}^n f_k$. ∎

Theorem 41.30. *Given* $f \in L^2$, *let*

$$u(t) = e^{-tL}f = \sum_{n=0}^\infty (f,\phi_n)e^{-t\lambda_n}\phi_n. \tag{41.29}$$

Then for $t > 0$, $u(t,x)$ *is smooth in* (t,x) *and solves the heat equation*

$$u_t(t,x) = -Lu(t,x),\ \ u(t,x) = 0 \text{ for } x \in \partial J \tag{41.30}$$

$$\text{and } f = L^2 - \lim_{t\downarrow 0} u(t) \tag{41.31}$$

Moreover, $u(t,x) = \int_J p_t(x,y)f(y)\rho(y)dy$ *where*

$$p_t(x,y) := \sum_{n=0}^\infty e^{-t\lambda_n}\phi_n(x)\phi_n(y) \tag{41.32}$$

is a smooth function in $t > 0$ *and* $x, y \in J$. *The function* p_t *is called the* ***Diurichlet Heat Kernel*** *for* L.

Proof. (Sketch.) For any $t > 0$ and $k \in \mathbb{N}$, $\sup_n\left(e^{-t\lambda_n}\lambda_n^k\right) < \infty$ and so by Lemma 41.29, for $t > 0$, $u(t) \in D\left(L^k\right) = H_0^{2k}(J)$[4] (Corollary 41.24) and

$$L^k u(t) = \sum_{n=0}^\infty (f,\phi_n)e^{-t\lambda_n}\lambda_n^k\phi_n.$$

Also we have $L^k u^{(m)}(t)$ exists in L^2 for all $k, m \in \mathbb{N}$ and

$$L^k u^{(m)}(t) = (-1)^m\sum_{n=0}^\infty (f,\phi_n)e^{-t\lambda_n}\lambda_n^{k+m}\phi_n.$$

By Sobolev inequalities and elliptic estimates such as Proposition 41.21 and Lemma 41.23, one concludes that $u \in C^\infty((0,\infty), H_0^k(J))$ for all k and then that $u \in C^\infty((0,\infty)\times J, \mathbb{R})$. Eq. (41.30) is now relatively easy to prove and Eq. (41.31) follows from the following computation

$$\|f - u(t)\|_2^2 = \sum_{n=1}^\infty |(f,\phi_n)|^2\left|1 - e^{-t\lambda_n}\right|^2$$

which goes to 0 as $t \downarrow 0$ by the D.C.T. for sums.

Finally from Eq. (41.29)

$$\begin{aligned} u(t,x) &= \sum_{n=0}^\infty \int_J f(y)\phi(y)\rho(y)dy e^{-t\lambda_n}\phi_n(x) \\ &= \int_J \sum_{n=0}^\infty e^{-t\lambda_n}\phi_n(x)\phi(y)f(y)\rho(y)dy \end{aligned}$$

where the interchange of the sum and the integral is permissible since

$$\begin{aligned} &\int_J \sum_{n=0}^\infty e^{-t\lambda_n}\left|\phi_n(x)\phi(y)f(y)\right|\rho(y)dy \\ &\qquad \le C\int_J \sum_{n=0}^\infty e^{-t\lambda_n}(\lambda_0+\lambda_n)^2\left|f(y)\right|\rho(y)dy < \infty \end{aligned}$$

since $\sum_{n=0}^\infty e^{-t\lambda_n}(\lambda_0+\lambda_n)^2 < \infty$ because λ_n grows linearly in n. Moreover one similarly shows

$$\left(\frac{\partial}{\partial t}\right)^j \partial_x^{2k-1}\partial_y^{2l-1}p_t(x,y) = \sum_{n=0}^\infty (-\lambda_n)^j e^{-t\lambda_n}\partial_x^{2k-1}\phi_n(x)\partial_y^{2l-1}\phi(y)$$

where the above operations are permissible since

[4] Basically, if $L^k u = g \in L^2(J)$ then $u = G^k g \in H_0^{2k}(J)$.

$$\left\|\phi_n^{(2k-1)}\right\|_u \le C\|\phi_n\|_{H_0^{2k}(J)} \le C\left\|(L+\lambda_0)^k \phi_n\right\|_2 = C(\lambda_n+\lambda_0)^k$$

and therefore

$$\sum_{n=0}^{\infty}\left|(-\lambda_n)^j e^{-t\lambda_n}\partial_x^{2k-1}\phi_n(x)\partial_y^{2l-1}\phi(y)\right| \le C\sum_{n=0}^{\infty}|\lambda_n|^j(\lambda_n+\lambda_0)^{k+l}e^{-t\lambda_n}<\infty.$$

Again we use λ_n grows linearly with n. From this one may conclude that $p_t(x,y)$ is smooth for $t>0$ and $x,y\in J$. (We will do this in more detail when we work out the higher dimensional analogue.) ■

Remark 41.31 (Wave Equation). Suppose $f \in D(L^k)$, then

$$|(f,\phi_n)| = \left|\frac{1}{\lambda_n^k}(f,L^k\phi_n)\right| = \left|\frac{1}{\lambda_n^k}(L^k f,\phi_n)\right| \le \frac{1}{|\lambda_n^k|}\left\|L^k f\right\|_2$$

and therefore

$$\cos\left(t\sqrt{L}\right)f := \sum_{n=0}^{\infty}\cos\left(t\sqrt{\lambda_n}\right)(f,\phi_n)\phi_n$$

will be convergent in L^2 but moreover

$$L^k\cos\left(t\sqrt{L}\right)f = \sum_{n=0}^{\infty}\cos\left(t\sqrt{\lambda_n}\right)(f,\phi_n)\lambda_n^k\phi_n = \sum_{n=0}^{\infty}\cos\left(t\sqrt{\lambda_n}\right)(L^kf,\phi_n)\phi_n$$

will also be convergent. Therefore if we let

$$u(t) = \cos\left(t\sqrt{L}\right)f + \frac{\sin\left(t\sqrt{L}\right)}{\sqrt{L}}g$$

where $f,g\in D(L^k)$ for all k. Then we will get a solution to the wave equation

$$u_{tt}(t,x)+Lu(t,x)=0 \text{ with } u(0)=f \text{ and } \dot{u}(0)=g.$$

More on all of this later.

41.5 Extensions to Other Boundary Conditions

In this section, we will assume $\rho\in C^2(J,(0,\infty))$,

$$Lu = -\rho^{-1}(\rho a u')' + bu' + cu \tag{41.33}$$

is an elliptic ODE on $L^2(J)$ with smooth coefficients and

$$(u,v) = (u,v)_\rho = \int_J u(x)v(x)\rho(x)dx. \tag{41.34}$$

Theorem 41.32. *For $v\in H^2(J)$ let*

$$L^*v = -\rho^{-1}(\rho a v')' - bv' + \left[c-\rho^{-1}(\rho b)'\right]v. \tag{41.35}$$

Then for $u,v\in H^2(J)$,

$$(Lu,v) = (u,L^*v) + \mathcal{B}(u,v)|_{\partial J} \tag{41.36}$$

where

$$\mathcal{B}(u,v) = \rho a\left\{(u',u)\cdot(-v,v'+\frac{b}{a}v)\right\}. \tag{41.37}$$

Proof. This is an exercise in integration by parts,

$$\begin{aligned}(Lu,v) &= \int_J\left(-(\rho a u')' + \rho b u' + \rho c u\right)vdm\\ &= \int_J\left(\rho a u' v' - (\rho b v)' u + \rho c u\right)dm + [\rho b u v - \rho a u' v]\,|_{\partial J}\\ &= \int_J\left(-u(\rho a v')' - (\rho b v)' u + \rho c v u\right)dm\\ &\quad + [\rho b u v + \rho a u v' - \rho a u' v]\,|_{\partial J}\\ &= \int_J\left(-u\rho^{-1}(\rho a v')' - \rho^{-1}(\rho b v)' u + c v u\right)\rho dm\\ &\quad + \left[\rho a\left(\frac{b}{a}uv + uv' - vu'\right)\right]|_{\partial J}\\ &= (u,L^*v) + \left[\rho a(u',u)\cdot(-v,v'+\frac{b}{a}v)\right]|_{\partial J}.\end{aligned}$$

■

Notation 41.33 *Given $(\alpha,\beta):\partial J\to\mathbb{R}^2\setminus\{0\}$ and $u,v\in H^2(J)$ let*

$$Bu = \alpha u' + \beta u = (\alpha,\beta)\cdot(u',u) \text{ on } \partial J$$

and

$$B^*v = \alpha v' + \left(\beta + \frac{b}{a}\alpha\right)v = \alpha v' + \tilde{\beta}v \text{ on } \partial J$$

where $\tilde{\beta} := \left(\beta + \frac{b}{a}\alpha\right)$.

Remarks 41.34 *The function $(\alpha,\tilde{\beta}):\partial J\to\mathbb{R}^2$ also takes values in $\mathbb{R}^2\setminus\{0\}$ because $(\alpha,\tilde{\beta})=0$ iff $(\alpha,\beta)=0$. Furthermore if $\alpha=0$ then $\tilde{\beta}=\beta$.*

Proposition 41.35. *Let B and B^* be as defined in Notation 41.33 and define*

$$D(L) = \left\{u\in H^2(J): Bu=0 \text{ on } \partial J\right\}.$$
$$D(L^*) = \left\{u\in H^2(J): B^*u=0 \text{ on } \partial J\right\},$$

Then $v \in H^2(J)$ *satisfies*

$$(Lu, v) = (u, L^* v) \text{ for all } u \in D(L) \tag{41.38}$$

iff $v \in D(L^*)$. *(This result will be substantially improved on in Theorem 41.41 below.)*

Proof. We have to check that $\mathcal{B}(u,v)$ appearing in Eq. (41.36) is 0. (Actually we must check that $\mathcal{B}(u,v)|_{\partial J} = 0$ which we might arrange by using something like "periodic boundary conditions." I am not considering this type of condition at the moment. Since u may be chosen to be zero near r or s we must require $\mathcal{B}(u,v) = 0$ on ∂J.) Now $\mathcal{B}(u,v) = 0$ iff

$$(u', u) \cdot \left(-v, v' + \frac{b}{a} v\right) = 0 \tag{41.39}$$

which happens iff (u', u) is parallel to $\left(v' + \frac{b}{a}v, v\right)$. The boundary condition $Bu = 0$ may be rewritten as saying $(u', u) \cdot (\alpha, \beta) = 0$ or equivalently that (u', u) is parallel to $(-\beta, \alpha)$ on ∂J. Therefore the condition in Eq. (41.39) is equivalent to $(-\beta, \alpha)$ is parallel to $\left(v' + \frac{b}{a}v, v\right)$ or equivalently that

$$0 = (\alpha, \beta) \cdot \left(v' + \frac{b}{a} v, v\right) = B^* v.$$

■

Corollary 41.36. *The formulas for* L *and* L^* *agree iff* $b = 0$ *in which case*

$$Lu = -\rho^{-1} D (a \rho u') + cu,$$

$B = B^*$, $D(L) = D(L^*)$ *and*

$$(Lu, v) = (u, Lv) \text{ for all } u, v \in D(L). \tag{41.40}$$

(In fact L *is a "self-adjoint operator," as we will see later by showing* $(L + \lambda_0)^{-1}$ *exists for* λ_0 *sufficiently large. Eq. (41.40) then may be used to deduce* $(L + \lambda_0)^{-1}$ *is a bounded self-adjoint operator with a symmetric Green's functions* G.*)*

41.5.1 Dirichlet Forms Associated to $(L, D(L))$

For the rest of this section let $a, b_1, b_2, c_0, \rho \in C^2(J)$, with $a > 0$ and $\rho > 0$ on J and for $u, v \in H^1(J)$, let

$$\mathcal{E}(u,v) := \int_J (a u' v' + b_1 u v' + b_2 u' v + c_0 u v) \rho dm \text{ and} \tag{41.41}$$

$$\|u\|_{H^1(J)} := \left(\|u'\|^2 + \|u\|^2\right)^{1/2}$$

where $\|u\|^2 = (u,u)_\rho$ as defined in Eq. (41.34).

Lemma 41.37 (A Coercive inequality for $\mathcal{E}$). *There is a constant* $K < \infty$ *such that*

$$|\mathcal{E}(u,v)| \le K \|u\|_{H^1(J)} \|v\|_{H^1(J)} \text{ for } u, v \in H^1(J). \tag{41.42}$$

Let $a_0 = \min_J a$, $\bar{c} = \min_J c_0$ *and* $B := \max_J |b_1 + b_2|$, *then for* $u \in H^1(J)$,

$$\mathcal{E}(u,u) \ge \frac{a_0}{2} \|u'\|^2 + \left(\bar{c} - \frac{B^2}{2a_0}\right) \|u\|^2. \tag{41.43}$$

Proof. Let $A = \max_J a$, $B_i = \max_J |b_i|$ and $C_0 := \max_J |c_0|$, then

$$\begin{aligned} |\mathcal{E}(u,v)| &\le \int_J (a |u'| |v'| + |b_1| |u| |v'| + |b_2| |u'| |v| + |c_0| |u| |v|) \rho dm \\ &\le A \|u'\| \|v'\| + B_1 \|u\| \|v'\| + B_2 \|u'\| \|v\| + C_0 \|u\| \|v\| \\ &\le K \left(\|u'\|^2 + \|u\|^2\right)^{1/2} \left(\|v'\|^2 + \|v\|^2\right)^{1/2}. \end{aligned}$$

Let $a_0 = \min_J a$, $\bar{c} = \min_J c$ and $B := \max_J |b_1 + b_2|$, then for any $\delta > 0$,

$$\begin{aligned} \mathcal{E}(u,u) &= \int_J \left(a |u'|^2 + (b_1 + b_2) u u' + c_0 |u|^2\right) \rho dm \\ &\ge a_0 \|u'\|^2 + \bar{c} \|u\|^2 - B \int_J |u| |u'| \rho dm \\ &\ge a_0 \|u'\|^2 + \bar{c} \|u\|^2 - \frac{B}{2} \left(\delta \|u'\|^2 + \delta^{-1} \|u\|^2\right) \\ &= \left(a_0 - \frac{B\delta}{2}\right) \|u'\|^2 + \left(\bar{c} - \frac{B}{2}\delta^{-1}\right) \|u\|^2. \end{aligned}$$

Taking $\delta = a_0/B$ in this equation proves Eq. (41.43). ■

Theorem 41.38. *Let*

$$\begin{aligned} b &= (b_2 - b_1), \quad c := c_0 - \rho^{-1} (\rho b_1)', \\ Lu &= -\rho^{-1} (a \rho u')' + b u' + c u \text{ and} \\ Bu &= (\rho a u' + \rho b_1 u)|_{\partial J}. \end{aligned} \tag{41.44}$$

Then for $u \in H^2(J)$ *and* $v \in H^1(J)$

$$\mathcal{E}(u,v) = (Lu, v) + [(Bu) v]_{\partial J}$$

and for $u \in H^1(J)$ *and* $v \in H^2(J)$,

$$\mathcal{E}(u,v) = (u, L^* v) + [(B^* v) u]_{\partial J}.$$

Here (as in Eq. (41.35)

$$L^*v = -\rho^{-1}\left(a\rho u'\right)' - \rho^{-1}\left[\rho bu\right]' + cu$$

and (as in Notation 41.33)

$$B^*v = \rho a v' + \left(\rho b_1 + \frac{b}{a}\rho a\right) v = \rho a v' + \rho b_2 v.$$

Proof. Let $u \in H^2(J)$ and $v \in H^1(J)$ and integrating Eq. (41.41) by parts to find

$$\begin{aligned}\mathcal{E}(u,v) &= \int_J \left(-\rho^{-1}\left(a\rho u'\right)' v - \rho^{-1}\left(\rho b_1 u\right)' v + b_2 u' v + c_0 uv\right)\rho dm \\ &\quad + \left[\rho a u' v + \rho b_1 uv\right]_{\partial J} \\ &= (Lu, v) + \left[Bu \cdot v\right]_{\partial J} \qquad (41.45)\end{aligned}$$

where

$$\begin{aligned}Lu &= -\rho^{-1}\left(a\rho u'\right)' - \rho^{-1}\left(\rho b_1 u\right)' + b_2 u' + c_0 u \\ &= -\rho^{-1}\left(a\rho u'\right)' + \left(b_2 - b_1\right) u' + \left[c_0 - \rho^{-1}\left(\rho b_1\right)'\right] u \\ &= -\rho^{-1}\left(a\rho u'\right)' + bu' + cu\end{aligned}$$

and

$$Bu = \rho a u' + \rho b_1 u.$$

Similarly

$$\begin{aligned}\mathcal{E}(u,v) &= \int_J \left(-u\rho^{-1}\left(a\rho v'\right)' + b_1 u v' - u\rho^{-1}\left(\rho b_2 v\right)' + c_0 uv\right)\rho dm \\ &\quad + \left[\left(\rho a u v' + \rho b_2 uv\right)\right]_{\partial J} \\ &= (u, L^{\dagger} v) + \left[B^{\dagger} v \cdot u\right]_{\partial J}\end{aligned}$$

where

$$\begin{aligned}L^{\dagger} v &= -\rho^{-1}\left(a\rho v'\right)' + b_1 v' - \rho^{-1}\left(\rho b_2 v\right)' + c_0 v \\ &= -\rho^{-1}\left(a\rho v'\right)' + \left(b_1 - b_2\right) v' + \left[c_0 - \rho^{-1}\left(\rho b_2\right)'\right] v \\ &= -\rho^{-1}\left(a\rho v'\right)' - bv' + \left[c + \rho^{-1}\left(\rho\left(b_1 - b_2\right)\right)^{-1}\right] v \\ &= -\rho^{-1}\left(a\rho v'\right)' - bv' + \left[c - \rho^{-1}\left(\rho b\right)'\right] v = L^* v.\end{aligned}$$

and

$$B^{\dagger} v = \left(\rho a v' + \rho b_2 v\right) = B^* v.$$

■

Remark 41.39. As a consequence of Theorem 41.38, the mapping

$$\left(a, b_1, b_2, c_0\right) \to \left[\begin{array}{c}(u,v) \to \\ \mathcal{E}(u,v) = \int_J \left(a u' v' + b_1 u v' + b_2 u' v + c_0 uv\right)\rho dm\end{array}\right]$$

is highly **non-**injective. In fact $\mathcal{E}$ depends only on a, $b = b_2 - b_1$ and $c := c_0 - \rho^{-1}\left(\rho b_1\right)'$ on J and b_1 on ∂J.

Corollary 41.40. *As above let $(\alpha, \beta) : \partial J \to \mathbb{R}^2 \setminus \{0\}$ and let*

$$\begin{aligned}D(L) &= \left\{u \in H^2(J) : Bu = \alpha u' + \beta u = 0 \text{ on } \partial J\right\} \text{ and} \\ Lu &= -\rho^{-1}\left(a\rho u'\right)' + bu' + cu.\end{aligned}$$

Given $\lambda_0 > 0$ sufficiently large, $(L + \lambda_0) : D(L) \to L^2(J)$ and $(L^* + \lambda_0) : D(L^*\ ) \to L^2(J)$ are invertible and there is a continuous Green's function $G(x,y)$ such that

$$(L + \lambda_0)^{-1} f(x) = \int_J G(x,y) f(y) dy.$$

Proof. Let us normalize α so that $\alpha = a$ whenever $\alpha \neq 0$. The boundary term in Eq. (41.45) will be zero whenever

$$au' + b_1 u = 0 \text{ when } v \neq 0 \text{ on } \partial J.$$

This suggests that we define a subspace χ of $H^1(J)$ by

$$\chi := \left\{u \in H^1(J) : u = 0 \text{ on } \partial J \text{ where } \alpha = 0 \text{ on } \partial J\right\}.$$

Hence χ is either $H_0^1(J)$, $H^1(J)$, $\left\{u \in H^1(J) : u(r) = 0\right\}$ or $\left\{u \in H^1(J) : u(s) = 0\right.$ Now choose a function $b_1 \in C^2(J)$ such that $b_1 = \beta$ on ∂J, then set $b_2 := b + b_1$ and $c_0 = c + \rho^{-1}\left(\rho b_1\right)'$, then

$$D(L) = \chi \cap \left\{u \in H^2(J) : Bu = au' + b_1 u = 0 \text{ on } \partial J\right\}$$

and

$$(Lu, v) = \mathcal{E}(u,v) \text{ for all } u \in D(L) \text{ and } v \in \chi.$$

Using this observation, it follows from Eq. (41.43) of Lemma 41.37, for λ_0 sufficiently large and any $u \in D(L)$, that

$$\begin{aligned}\left(\left(L + \lambda_0\right) u, u\right) &= \mathcal{E}(u,u) + \lambda_0 (u,u) \\ &\geq \frac{a_0}{2}\left\|u'\right\|^2 + \left(\bar{c} - \frac{B^2}{2a_0} + \lambda_0\right)\|u\|^2 \geq \frac{a_0}{2}\|u\|_{H^1(J)}^2.\end{aligned}$$

As usual this equation shows $\mathrm{Nul}(L + \lambda_0) = \{0\}$. Similarly on shows

$$(u, L^* v) = \mathcal{E}(u,v) \text{ for all } v \in D(L^*) \text{ and } u \in \chi$$

and working as above we conclude that $\mathrm{Nul}(L^* + \lambda_0) = \{0\}$. The remaining assertions are now proved as in the proof of Corollary 41.24. ■

With this result in hand we may now improve on Proposition 41.35.

Theorem 41.41. *Let L, B, $D(L)$, L^*, B^* and $D(L^*)$ be as in Proposition 41.35 and $v \in L^2(J)$. Then there exists $g \in L^2(J)$ such that*

$$(Lu, v) = (u, g) \text{ for all } u \in D(L) \qquad (41.46)$$

iff $v \in D(L^*)$ and in which case $g = L^* v$.

Proof. Choose $\lambda_0 > 0$ so that $L_0 := (L + \lambda_0 I) : D(L) \to L^2(J)$ is invertible. Then Eq. (41.46) is equivalent to

$$(L_0 u, v) = (u, g + \lambda_0 v) \text{ for all } u \in D(L). \tag{41.47}$$

Taking $u = L_0^{-1} w$ with $w \in L^2(J)$ in this equation implies

$$(w, v) = (L_0^{-1} w, g + \lambda_0 v) = \left(w, \left(L_0^{-1}\right)^* (g + \lambda_0 v)\right) \text{ for all } w \in L^2(J)$$

which shows

$$v = \left(L_0^{-1}\right)^* (g + \lambda_0 v). \tag{41.48}$$

Since $(L_0 u, v) = (u, L_0^* v)$ for all $u \in D(L)$ and $v \in D(L^*)$, by replacing u by $L_0^{-1} u$ and v by $(L_0^*)^{-1} v$ in this equation we learn

$$\left(u, (L_0^*)^{-1} v\right) = (L_0^{-1} u, v) \text{ for all } u, v \in L^2(J).$$

From this equation it follows that $(L_0^*)^{-1} = \left(L_0^{-1}\right)^*$ and hence from Eq. (41.48) it follows that $v \in D(L_0^*) = D(L^*)$.

■

Part XIII

Constant Coefficient Equations

42

Convolutions, Test Functions and Partitions of Unity

42.1 Convolution and Young's Inequalities

Letting δ_x denote the "delta–function" at x, we wish to define a product $(*)$ on functions on $\mathbb{R}^n$ such that $\delta_x * \delta_y = \delta_{x+y}$. Now formally any function f on $\mathbb{R}^n$ is of the form

$$f = \int_{\mathbb{R}^n} f(x)\delta_x dx$$

so we should have

$$\begin{aligned} f * g &= \int_{\mathbb{R}^n\times\mathbb{R}^n} f(x)g(y)\delta_x * \delta_y dxdy = \int_{\mathbb{R}^n\times\mathbb{R}^n} f(x)g(y)\delta_{x+y}dxdy \\ &= \int_{\mathbb{R}^n\times\mathbb{R}^n} f(x-y)g(y)\delta_x dxdy \\ &= \int_{\mathbb{R}^n}\left[\int_{\mathbb{R}^n} f(x-y)g(y)dy\right]\delta_x dx \end{aligned}$$

which suggests we make the following definition.

Definition 42.1. *Let $f, g : \mathbb{R}^n \to \mathbb{C}$ be measurable functions. We define*

$$f * g(x) = \int_{\mathbb{R}^n} f(x-y)g(y)dy$$

whenever the integral is defined, i.e. either $f(x-\cdot)g(\cdot) \in L^1(\mathbb{R}^n, m)$ or $f(x-\cdot)g(\cdot) \geq 0$. Notice that the condition that $f(x-\cdot)g(\cdot) \in L^1(\mathbb{R}^n, m)$ is equivalent to writing $|f| * |g|(x) < \infty$.

Notation 42.2 *Given a multi-index $\alpha \in \mathbb{Z}_+^n$, let $|\alpha| = \alpha_1 + \cdots + \alpha_n$,*

$$x^\alpha := \prod_{j=1}^n x_j^{\alpha_j}, \text{ and } \partial_x^\alpha = \left(\frac{\partial}{\partial x}\right)^\alpha := \prod_{j=1}^n \left(\frac{\partial}{\partial x_j}\right)^{\alpha_j}.$$

Remark 42.3 (The Significance of Convolution). Suppose that $L = \sum_{|\alpha|\leq k} a_\alpha \partial^\alpha$ is a constant coefficient differential operator and suppose that we can solve (uniquely) the equation $Lu = g$ in the form

$$u(x) = Kg(x) := \int_{\mathbb{R}^n} k(x,y)g(y)dy$$

where $k(x,y)$ is an "integral kernel." (This is a natural sort of assumption since, in view of the fundamental theorem of calculus, integration is the inverse operation to differentiation.) Since $\tau_z L = L\tau_z$ for all $z \in \mathbb{R}^n$, (this is another way to characterize constant coefficient differential operators) and $L^{-1} = K$ we should have $\tau_z K = K\tau_z$. Writing out this equation then says

$$\begin{aligned} \int_{\mathbb{R}^n} k(x-z,y)g(y)dy &= (Kg)(x-z) = \tau_z Kg(x) = (K\tau_z g)(x) \\ &= \int_{\mathbb{R}^n} k(x,y)g(y-z)dy = \int_{\mathbb{R}^n} k(x,y+z)g(y)dy. \end{aligned}$$

Since g is arbitrary we conclude that $k(x-z,y) = k(x,y+z)$. Taking $y=0$ then gives

$$k(x,z) = k(x-z,0) =: \rho(x-z).$$

We thus find that $Kg = \rho * g$. Hence we expect the convolution operation to appear naturally when solving constant coefficient partial differential equations. More about this point later.

The following proposition is an easy consequence of Minkowski's inequality for integrals.

Proposition 42.4. *Suppose $q \in [1,\infty]$, $f \in L^1$ and $g \in L^q$, then $f * g(x)$ exists for almost every x, $f * g \in L^q$ and*

$$\|f * g\|_p \leq \|f\|_1 \|g\|_p .$$

For $z \in \mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{C}$, let $\tau_z f : \mathbb{R}^n \to \mathbb{C}$ be defined by $\tau_z f(x) = f(x-z)$.

Proposition 42.5. *Suppose that $p \in [1,\infty)$, then $\tau_z : L^p \to L^p$ is an isometric isomorphism and for $f \in L^p$, $z \in \mathbb{R}^n \to \tau_z f \in L^p$ is continuous.*

Proof. The assertion that $\tau_z : L^p \to L^p$ is an isometric isomorphism follows from translation invariance of Lebesgue measure and the fact that $\tau_{-z} \circ \tau_z = id$. For the continuity assertion, observe that

$$\|\tau_z f - \tau_y f\|_p = \|\tau_{-y}(\tau_z f - \tau_y f)\|_p = \|\tau_{z-y} f - f\|_p$$

from which it follows that it is enough to show $\tau_z f \to f$ in L^p as $z \to 0 \in \mathbb{R}^n$.

When $f \in C_c(\mathbb{R}^n)$, $\tau_z f \to f$ uniformly and since the $K := \cup_{|z|\le 1}\mathrm{supp}(\tau_z f)$ is compact, it follows by the dominated convergence theorem that $\tau_z f \to f$ in L^p as $z \to 0 \in \mathbb{R}^n$. For general $g \in L^p$ and $f \in C_c(\mathbb{R}^n)$,

$$\begin{aligned}\|\tau_z g - g\|_p &\le \|\tau_z g - \tau_z f\|_p + \|\tau_z f - f\|_p + \|f - g\|_p \\ &= \|\tau_z f - f\|_p + 2\|f - g\|_p\end{aligned}$$

and thus

$$\limsup_{z\to 0} \|\tau_z g - g\|_p \le \limsup_{z\to 0} \|\tau_z f - f\|_p + 2\|f - g\|_p = 2\|f - g\|_p .$$

Because $C_c(\mathbb{R}^n)$ is dense in L^p, the term $\|f - g\|_p$ may be made as small as we please. ■

Definition 42.6. *Suppose that (X, τ) is a topological space and μ is a measure on $\mathcal{B}_X = \sigma(\tau)$. For a measurable function $f : X \to \mathbb{C}$ we define the essential support of f by*

$$\mathrm{supp}_\mu(f) = \{x \in U : \mu(\{y \in V : f(y) \neq 0\}) > 0 \ \forall \tau \ni V \ni x\} . \tag{42.1}$$

Lemma 42.7. *Suppose (X, τ) is second countable and $f : X \to \mathbb{C}$ is a measurable function and μ is a measure on $\mathcal{B}_X$. Then $X := U \setminus \mathrm{supp}_\mu(f)$ may be described as the largest open set W such that $f1_W(x) = 0$ for μ – a.e. x. Equivalently put, $C := \mathrm{supp}_\mu(f)$ is the smallest closed subset of X such that $f = f1_C$ a.e.*

Proof. To verify that the two descriptions of $\mathrm{supp}_\mu(f)$ are equivalent, suppose $\mathrm{supp}_\mu(f)$ is defined as in Eq. (42.1) and $W := X \setminus \mathrm{supp}_\mu(f)$. Then

$$\begin{aligned} W &= \left\{ \begin{array}{c} x \in X : \mu(\{y \in V : f(y) \neq 0\}) = 0 \\ \text{for some neighborhood } V \text{ of } x \end{array} \right\} \\ &= \cup \{V \subset_o X : \mu(f1_V \neq 0) = 0\} \\ &= \cup \{V \subset_o X : f1_V = 0 \text{ for } \mu - \text{a.e.}\} . \end{aligned}$$

So to finish the argument it suffices to show $\mu(f1_W \neq 0) = 0$. To to this let $\mathcal{U}$ be a countable base for τ and set

$$\mathcal{U}_f := \{V \in \mathcal{U} : f1_V = 0 \text{ a.e.}\}.$$

Then it is easily seen that $W = \cup \mathcal{U}_f$ and since $\mathcal{U}_f$ is countable $\mu(f1_W \neq 0) \le \sum_{V\in\mathcal{U}_f} \mu(f1_V \neq 0) = 0$. ■

Lemma 42.8. *Suppose $f, g, h : \mathbb{R}^n \to \mathbb{C}$ are measurable functions and assume that x is a point in $\mathbb{R}^n$ such that $|f| * |g|(x) < \infty$ and $|f| * (|g| * |h|)(x) < \infty$, then*

1. *$f * g(x) = g * f(x)$*

2. *$f * (g * h)(x) = (f * g) * h(x)$*
3. *If $z \in \mathbb{R}^n$ and $\tau_z(|f| * |g|)(x) = |f| * |g|(x - z) < \infty$, then*
$$\tau_z(f * g)(x) = \tau_z f * g(x) = f * \tau_z g(x)$$
4. *If $x \notin \overline{\mathrm{supp}_m(f) + \mathrm{supp}_m(g)}$ then $f * g(x) = 0$ and in particular, $\mathrm{supp}_m(f * g) \subset \overline{\mathrm{supp}_m(f) + \mathrm{supp}_m(g)}$ where in defining $\mathrm{supp}_m(f * g)$ we will use the convention that "$f * g(x) \neq 0$" when $|f| * |g|(x) = \infty$.*

Proof. For item 1.,

$$|f| * |g|(x) = \int_{\mathbb{R}^n} |f|(x - y)|g|(y)dy = \int_{\mathbb{R}^n} |f|(y)|g|(y - x)dy = |g| * |f|(x)$$

where in the second equality we made use of the fact that Lebesgue measure invariant under the transformation $y \to x - y$. Similar computations prove all of the remaining assertions of the first three items of the lemma.

Item 4. Since $f * g(x) = \tilde{f} * \tilde{g}(x)$ if $f = \tilde{f}$ and $g = \tilde{g}$ a.e. we may, by replacing f by $f1_{\mathrm{supp}_m(f)}$ and g by $g1_{\mathrm{supp}_m(g)}$ if necessary, assume that $\{f \neq 0\} \subset \mathrm{supp}_m(f)$ and $\{g \neq 0\} \subset \mathrm{supp}_m(g)$. So if $x \notin (\mathrm{supp}_m(f) + \mathrm{supp}_m(g))$ then $x \notin (\{f \neq 0\} + \{g \neq 0\})$ and for all $y \in \mathbb{R}^n$, either $x - y \notin \{f \neq 0\}$ or $y \notin \{g \neq 0\}$. That is to say either $x - y \in \{f = 0\}$ or $y \in \{g = 0\}$ and hence $f(x - y)g(y) = 0$ for all y and therefore $f * g(x) = 0$. This shows that $f * g = 0$ on $\mathbb{R}^n \setminus \left(\overline{\mathrm{supp}_m(f) + \mathrm{supp}_m(g)}\right)$ and therefore

$$\mathbb{R}^n \setminus \left(\overline{\mathrm{supp}_m(f) + \mathrm{supp}_m(g)}\right) \subset \mathbb{R}^n \setminus \mathrm{supp}_m(f * g),$$

i.e. $\mathrm{supp}_m(f * g) \subset \overline{\mathrm{supp}_m(f) + \mathrm{supp}_m(g)}$. ■

Remark 42.9. Let A, B be closed sets of $\mathbb{R}^n$, it is not necessarily true that $A + B$ is still closed. For example, take

$$A = \{(x, y) : x > 0 \text{ and } y \ge 1/x\} \text{ and } B = \{(x, y) : x < 0 \text{ and } y \ge 1/|x|\},$$

then every point of $A + B$ has a positive y - component and hence is not zero. On the other hand, for $x > 0$ we have $(x, 1/x) + (-x, 1/x) = (0, 2/x) \in A + B$ for all x and hence $0 \in \overline{A + B}$ showing $A + B$ is not closed. Nevertheless if one of the sets A or B is compact, then $A + B$ is closed again. Indeed, if A is compact and $x_n = a_n + b_n \in A + B$ and $x_n \to x \in \mathbb{R}^n$, then by passing to a subsequence if necessary we may assume $\lim_{n\to\infty} a_n = a \in A$ exists. In this case

$$\lim_{n\to\infty} b_n = \lim_{n\to\infty} (x_n - a_n) = x - a \in B$$

exists as well, showing $x = a + b \in A + B$.

Proposition 42.10. *Suppose that $p, q \in [1, \infty]$ and p and q are conjugate exponents, $f \in L^p$ and $g \in L^q$, then $f * g \in BC(\mathbb{R}^n)$, $\|f * g\|_u \le \|f\|_p \|g\|_q$ and if $p, q \in (1, \infty)$ then $f * g \in C_0(\mathbb{R}^n)$.*

Proof. The existence of $f*g(x)$ and the estimate $|f*g|(x) \le \|f\|_p \|g\|_q$ for all $x \in \mathbb{R}^n$ is a simple consequence of Holders inequality and the translation invariance of Lebesgue measure. In particular this shows $\|f*g\|_u \le \|f\|_p \|g\|_q$. By relabeling p and q if necessary we may assume that $p \in [1,\infty)$. Since

$$\begin{aligned}\|\tau_z(f*g) - f*g\|_u &= \|\tau_z f * g - f*g\|_u \\ &\le \|\tau_z f - f\|_p \|g\|_q \to 0 \text{ as } z \to 0\end{aligned}$$

it follows that $f*g$ is uniformly continuous. Finally if $p,q \in (1,\infty)$, we learn from Lemma 42.8 and what we have just proved that $f_m * g_m \in C_c(\mathbb{R}^n)$ where $f_m = f1_{|f|\le m}$ and $g_m = g1_{|g|\le m}$. Moreover,

$$\begin{aligned}\|f*g - f_m*g_m\|_u &\le \|f*g - f_m*g\|_u + \|f_m*g - f_m*g_m\|_u \\ &\le \|f-f_m\|_p \|g\|_q + \|f_m\|_p \|g-g_m\|_q \\ &\le \|f-f_m\|_p \|g\|_q + \|f\|_p \|g-g_m\|_q \to 0 \text{ as } m\to\infty\end{aligned}$$

showing $f*g \in C_0(\mathbb{R}^n)$. ■

Theorem 42.11 (Young's Inequality). *Let $p,q,r \in [1,\infty]$ satisfy*

$$\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}. \tag{42.2}$$

If $f \in L^p$ and $g \in L^q$ then $|f|*|g|(x) < \infty$ for m – a.e. x and

$$\|f*g\|_r \le \|f\|_p \|g\|_q. \tag{42.3}$$

In particular L^1 is closed under convolution. (The space $(L^1,*)$ is an example of a "Banach algebra" without unit.)

Remark 42.12. Before going to the formal proof, let us first understand Eq. (42.2) by the following scaling argument. For $\lambda > 0$, let $f_\lambda(x) := f(\lambda x)$, then after a few simple change of variables we find

$$\|f_\lambda\|_p = \lambda^{-1/p}\|f\| \text{ and } (f*g)_\lambda = \lambda f_\lambda * g_\lambda.$$

Therefore if Eq. (42.3) holds for some $p,q,r \in [1,\infty]$, we would also have

$$\begin{aligned}\|f*g\|_r &= \lambda^{1/r}\|(f*g)_\lambda\|_r \le \lambda^{1/r}\lambda \|f_\lambda\|_p \|g_\lambda\|_q \\ &= \lambda^{(1+1/r-1/p-1/q)} \|f\|_p \|g\|_q\end{aligned}$$

for all $\lambda > 0$. This is only possible if Eq. (42.2) holds.

Proof. Let $\alpha,\beta \in [0,1]$ and $p_1,p_2 \in [0,\infty]$ satisfy $p_1^{-1} + p_2^{-1} + r^{-1} = 1$. Then by Hölder's inequality,

$$\begin{aligned}|f*g(x)| &= \left|\int f(x-y)g(y)dy\right| \\ &\le \int |f(x-y)|^{(1-\alpha)} |g(y)|^{(1-\beta)} |f(x-y)|^\alpha |g(y)|^\beta dy \\ &\le \left(\int |f(x-y)|^{(1-\alpha)r} |g(y)|^{(1-\beta)r} dy\right)^{1/r} \left(\int |f(x-y)|^{\alpha p_1} dy\right)^{1/p_1} \times \\ &\quad \left(\int |g(y)|^{\beta p_2} dy\right)^{1/p_2} \\ &= \left(\int |f(x-y)|^{(1-\alpha)r} |g(y)|^{(1-\beta)r} dy\right)^{1/r} \|f\|_{\alpha p_1}^\alpha \|g\|_{\beta p_2}^\beta.\end{aligned}$$

Taking the r^{th} power of this equation and integrating on x gives

$$\begin{aligned}\|f*g\|_r^r &\le \int \left(\int |f(x-y)|^{(1-\alpha)r} |g(y)|^{(1-\beta)r} dy\right) dx \cdot \|f\|_{\alpha p_1}^\alpha \|g\|_{\beta p_2}^\beta \\ &= \|f\|_{(1-\alpha)r}^{(1-\alpha)r} \|g\|_{(1-\beta)r}^{(1-\beta)r} \|f\|_{\alpha p_1}^{\alpha r} \|g\|_{\beta p_2}^{\beta r}.\end{aligned} \tag{42.4}$$

Let us now suppose, $(1-\alpha)r = \alpha p_1$ and $(1-\beta)r = \beta p_2$, in which case Eq. (42.4) becomes,

$$\|f*g\|_r^r \le \|f\|_{\alpha p_1}^r \|g\|_{\beta p_2}^r$$

which is Eq. (42.3) with

$$p := (1-\alpha)r = \alpha p_1 \text{ and } q := (1-\beta)r = \beta p_2. \tag{42.5}$$

So to finish the proof, it suffices to show p and q are arbitrary indices in $[1,\infty]$ satisfying $p^{-1} + q^{-1} = 1 + r^{-1}$.

If α,β,p_1,p_2 satisfy the relations above, then

$$\alpha = \frac{r}{r+p_1} \text{ and } \beta = \frac{r}{r+p_2}$$

and

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{p_1}\frac{r+p_1}{r} + \frac{1}{p_2}\frac{r+p_2}{r} = \frac{1}{p_1} + \frac{1}{p_2} + \frac{2}{r} = 1 + \frac{1}{r}.$$

Conversely, if p,q,r satisfy Eq. (42.2), then let α and β satisfy $p = (1-\alpha)r$ and $q = (1-\beta)r$, i.e.

$$\alpha := \frac{r-p}{r} = 1 - \frac{p}{r} \le 1 \text{ and } \beta = \frac{r-q}{r} = 1 - \frac{q}{r} \le 1.$$

From Eq. (42.2), $\alpha = p(1-\frac{1}{q}) \ge 0$ and $\beta = q(1-\frac{1}{p}) \ge 0$, so that $\alpha,\beta \in [0,1]$. We then define $p_1 := p/\alpha$ and $p_2 := q/\beta$, then

$$\frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{r} = \beta\frac{1}{q} + \alpha\frac{1}{p} + \frac{1}{r} = \frac{1}{q} - \frac{1}{r} + \frac{1}{p} - \frac{1}{r} + \frac{1}{r} = 1$$

as desired. ■

Theorem 42.13 (Approximate δ – functions). *Let $p \in [1,\infty]$, $\phi \in L^1(\mathbb{R}^n)$, $a := \int_{\mathbb{R}^n} f(x)dx$, and for $t>0$ let $\phi_t(x) = t^{-n}\phi(x/t)$. Then*

1. *If $f \in L^p$ with $p<\infty$ then $\phi_t * f \to af$ in L^p as $t \downarrow 0$.*
2. *If $f \in BC(\mathbb{R}^n)$ and f is uniformly continuous then $\|\phi_t * f - f\|_\infty \to 0$ as $t \downarrow 0$.*
3. *If $f \in L^\infty$ and f is continuous on $U \subset_o \mathbb{R}^n$ then $\phi_t * f \to af$ uniformly on compact subsets of U as $t \downarrow 0$.*

See Theorem 8.15 if Folland for a statement about almost everywhere convergence.

Proof. Making the change of variables $y = tz$ implies

$$\phi_t * f(x) = \int_{\mathbb{R}^n} f(x-y)\phi_t(y)dy = \int_{\mathbb{R}^n} f(x-tz)\phi(z)dz$$

so that

$$\begin{aligned}\phi_t * f(x) - af(x) &= \int_{\mathbb{R}^n} [f(x-tz)-f(x)]\,\phi(z)dz \\ &= \int_{\mathbb{R}^n} [\tau_{tz}f(x)-f(x)]\,\phi(z)dz. \end{aligned} \tag{42.6}$$

Hence by Minkowski's inequality for integrals, Proposition 42.5 and the dominated convergence theorem,

$$\|\phi_t * f - af\|_p \le \int_{\mathbb{R}^n} \|\tau_{tz}f - f\|_p \,|\phi(z)|\, dz \to 0 \text{ as } t \downarrow 0.$$

Item 2. is proved similarly. Indeed, form Eq. (42.6)

$$\|\phi_t * f - af\|_\infty \le \int_{\mathbb{R}^n} \|\tau_{tz}f - f\|_\infty \,|\phi(z)|\, dz$$

which again tends to zero by the dominated convergence theorem because $\lim_{t\downarrow 0}\|\tau_{tz}f - f\|_\infty = 0$ uniformly in z by the uniform continuity of f.

Item 3. Let $B_R = B(0,R)$ be a large ball in $\mathbb{R}^n$ and $K \sqsubset\sqsubset U$, then

$$\begin{aligned}\sup_{x\in K}|\phi_t * f(x) - af(x)| &\le \left|\int_{B_R}[f(x-tz)-f(x)]\,\phi(z)dz\right| \\ &\quad + \left|\int_{B_R^c}[f(x-tz)-f(x)]\,\phi(z)dz\right| \\ &\le \int_{B_R}|\phi(z)|\,dz \cdot \sup_{x\in K, z\in B_R}|f(x-tz)-f(x)| \\ &\quad + 2\|f\|_\infty \int_{B_R^c}|\phi(z)|\,dz \\ &\le \|\phi\|_1 \cdot \sup_{x\in K, z\in B_R}|f(x-tz)-f(x)| \\ &\quad + 2\|f\|_\infty \int_{|z|>R}|\phi(z)|\,dz \end{aligned}$$

so that using the uniform continuity of f on compact subsets of U,

$$\limsup_{t\downarrow 0}\sup_{x\in K}|\phi_t * f(x) - af(x)| \le 2\|f\|_\infty \int_{|z|>R}|\phi(z)|\,dz \to 0 \text{ as } R\to\infty.$$

■

Remark 42.14 (Another Proof of part of Theorem 42.13). By definition of the convolution and Hölder's or Jensen's inequality we have

$$\begin{aligned}\int_{\mathbb{R}^n}|v*\phi_t(x)|^p dx &\le \int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n}(v(x-y))|\phi_t(y)|dy\right)^p dx \\ &\le \int_{\mathbb{R}^n\times\mathbb{R}^n}|v(x-y)|^p\phi_t(y)dy\,dx = \|v\|_{L^p}^p.\end{aligned}$$

Therefore $\|v*\phi_t\|_{L^p} \le \|v\|_{L^p}$ which implies $v*\phi_t \in L^p$. If $\phi_t \in C_c^\infty(\mathbb{R}^n)$, by differentiating under the integral (see Theorem 42.18 below) it is easily seen that $v*\phi_t \in C^\infty$. Finally for $u \in C_c(\mathbb{R}^n)$,

$$\begin{aligned}\|v - v*\phi_t\|_{L^p} &\le \|v-u\|_{L^p} + \|u - u*\phi_t\|_{L^p} + \|u*\phi_t - v*\phi_t\|_{L^p} \\ &\le \|u-u*\phi_t\|_{L^p} + 2\|v-u\|_{L^p}\end{aligned}$$

and hence

$$\limsup_{t\downarrow 0}\|v - v*\phi_t\|_{L^p} \le 2\|v-u\|_{L^p}$$

which may be made arbitrarily small since $C_c(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n, m)$.

Exercise 42.15. Let

$$f(t) = \begin{cases} e^{-1/t} & \text{if } t>0 \\ 0 & \text{if } t\le 0.\end{cases}$$

Show $f \in C^\infty(\mathbb{R},[0,1])$.

Lemma 42.16. *There exists $\phi \in C_c^\infty(\mathbb{R}^n, [0,\infty))$ such that $\phi(0) > 0$, $\operatorname{supp}(\phi) \subset \bar{B}(0,1)$ and $\int_{\mathbb{R}^n} \phi(x)dx = 1$.*

Proof. Define $h(t) = f(1-t)f(t+1)$ where f is as in Exercise 42.15. Then $h \in C_c^\infty(\mathbb{R}, [0,1])$, $\operatorname{supp}(h) \subset [-1,1]$ and $h(0) = e^{-2} > 0$. Define $c = \int_{\mathbb{R}^n} h(|x|^2)dx$. Then $\phi(x) = c^{-1}h(|x|^2)$ is the desired function. ■

Definition 42.17. *Let $X \subset \mathbb{R}^n$ be an open set. A **Radon** measure on $\mathcal{B}_X$ is a measure μ which is finite on compact subsets of X. For a Radon measure μ, we let $L^1_{loc}(\mu)$ consists of those measurable functions $f : X \to \mathbb{C}$ such that $\int_K |f|\, d\mu < \infty$ for all compact subsets $K \subset X$.*

Theorem 42.18 (Differentiation under integral sign). *Let $\Omega \subset \mathbb{R}^n$ and $f : \mathbb{R}^m \times \Omega \to \mathbb{R}$ be given. **Assume:***

1. *$x \to f(x,y)$ is differentiable for all $y \in \Omega$.*
2. *$\left|\frac{\partial f}{\partial x^i}(x,y)\right| \le g(y)$ for some g such that $\int_\Omega |g(y)|dy < \infty$.*
3. *$\int |f(x,y)|dy < \infty$.*

Then $\frac{\partial}{\partial x^i}\int_\Omega f(x,y)dy = \int_\Omega \frac{\partial f}{\partial x^i}(x,y)dy$ and moreover if $x \to \frac{\partial f}{\partial x^i}(x,y)$ is continuous then so is $x \to \int \frac{\partial f}{\partial x^i}(x,y)dy$.

The reader asked to use Theorem 42.18 to verify the following proposition.

Proposition 42.19. *Suppose that $f \in L^1_{loc}(\mathbb{R}^n, m)$ and $\phi \in C_c^1(\mathbb{R}^n)$, then $f * \phi \in C^1(\mathbb{R}^n)$ and $\partial_i(f*\phi) = f * \partial_i\phi$. Moreover if $\phi \in C_c^\infty(\mathbb{R}^n)$ then $f*\phi \in C^\infty(\mathbb{R}^n)$.*

Corollary 42.20 (C^∞ – Uryhson's Lemma). *Given $K \sqsubset\sqsubset U \subset_o \mathbb{R}^n$, there exists $f \in C_c^\infty(\mathbb{R}^n, [0,1])$ such that $\operatorname{supp}(f) \subset U$ and $f = 1$ on K.*

Proof. Let ϕ be as in Lemma 42.16, $\phi_t(x) = t^{-n}\phi(x/t)$ be as in Theorem 42.13, d be the standard metric on $\mathbb{R}^n$ and $\epsilon = d(K, U^c)$. Since K is compact and U^c is closed, $\epsilon > 0$. Let $V_\delta = \{x \in \mathbb{R}^n : d(x,K) < \delta\}$ and $f = \phi_{\epsilon/3} * 1_{V_{\epsilon/3}}$, then

$$\operatorname{supp}(f) \subset \overline{\operatorname{supp}(\phi_{\epsilon/3}) + V_{\epsilon/3}} \subset \bar{V}_{2\epsilon/3} \subset U.$$

Since $\bar{V}_{2\epsilon/3}$ is closed and bounded, $f \in C_c^\infty(U)$ and for $x \in K$,

$$f(x) = \int_{\mathbb{R}^n} 1_{d(y,K)<\epsilon/3} \cdot \phi_{\epsilon/3}(x-y)dy = \int_{\mathbb{R}^n} \phi_{\epsilon/3}(x-y)dy = 1.$$

The proof will be finished after the reader (easily) verifies $0 \le f \le 1$. ■

Here is an application of this corollary whose proof is left to the reader.

Lemma 42.21 (Integration by Parts). *Suppose f and g are measurable functions on $\mathbb{R}^n$ such that $t \to f(x_1,\dots,x_{i-1},t,x_{i+1},\dots,x_n)$ and $t \to g(x_1,\dots,x_{i-1},t,x_{i+1},\dots,x_n)$ are continuously differentiable functions on $\mathbb{R}$ for each fixed $x = (x_1,\dots,x_n) \in \mathbb{R}^n$. Moreover assume $f \cdot g$, $\frac{\partial f}{\partial x_i} \cdot g$ and $f \cdot \frac{\partial g}{\partial x_i}$ are in $L^1(\mathbb{R}^n, m)$. Then*

$$\int_{\mathbb{R}^n} \frac{\partial f}{\partial x_i} \cdot g dm = -\int_{\mathbb{R}^n} f \cdot \frac{\partial g}{\partial x_i} dm.$$

Exercise 42.22 (Integration by Parts). Suppose that $(x,y) \in \mathbb{R}\times\mathbb{R}^{n-1} \to f(x,y) \in \mathbb{C}$ and $(x,y) \in \mathbb{R}\times\mathbb{R}^{n-1} \to g(x,y) \in \mathbb{C}$ are measurable functions such that for each fixed $y \in \mathbb{R}^{n-1}$, $x \to f(x,y)$ and $x \to g(x,y)$ are continuously differentiable. Also assume $f\cdot g$, $\partial_x f \cdot g$ and $f \cdot \partial_x g$ are integrable relative to Lebesgue measure on $\mathbb{R}\times\mathbb{R}^{n-1}$, where $\partial_x f(x,y) := \frac{d}{dt} f(x+t,y)|_{t=0}$. Show

$$\int_{\mathbb{R}\times\mathbb{R}^{n-1}} \partial_x f(x,y)\cdot g(x,y)dxdy = -\int_{\mathbb{R}\times\mathbb{R}^{n-1}} f(x,y)\cdot \partial_x g(x,y)dxdy. \quad (42.7)$$

(Note: this result and Fubini's theorem proves Lemma 42.21.)

Hints: Let $\psi \in C_c^\infty(\mathbb{R})$ be a function which is 1 in a neighborhood of $0 \in \mathbb{R}$ and set $\psi_\epsilon(x) = \psi(\epsilon x)$. First verify Eq. (42.7) with $f(x,y)$ replaced by $\psi_\epsilon(x) f(x,y)$ by doing the x – integral first. Then use the dominated convergence theorem to prove Eq. (42.7) by passing to the limit, $\epsilon \downarrow 0$.

Solution 42.23 (42.22). By assumption, $\partial_x[\psi_\epsilon(x)f(x,y)] \cdot g(x,y)$ and $\psi_\epsilon(x)f(x,y)\partial_x g(x,y)$ are in $L^1(\mathbb{R}^n)$, so we may use Fubini's theorem and follow the hint to learn

$$\begin{aligned}\int_{\mathbb{R}^{n-1}} dy \int_{\mathbb{R}} \partial_x[\psi_\epsilon(x)f(x,y)] \cdot g(x,y)dx \\ = -\int_{\mathbb{R}^{n-1}} dy \int_{\mathbb{R}} [\psi_\epsilon(x)f(x,y)] \cdot \partial_x g(x,y)dx, \quad (42.8)\end{aligned}$$

wherein we have done and integration by parts. (There are no boundary terms because ψ_ϵ is compactly supported.) Now

$$\begin{aligned}\partial_x[\psi_\epsilon(x)f(x,y)] &= \partial_x\psi_\epsilon(x)\cdot f(x,y) + \psi_\epsilon(x)\partial_x f(x,y) \\ &= \epsilon\psi'(\epsilon x)f(x,y) + \psi_\epsilon(x)\partial_x f(x,y)\end{aligned}$$

and by the dominated convergence theorem and the given assumptions we have, as $\epsilon \downarrow 0$, that

$$\begin{aligned}&\left|\int_{\mathbb{R}^n} \epsilon\psi'(\epsilon x)f(x,y)\cdot g(x,y)dxdy\right| \le C\epsilon \int_{\mathbb{R}^n} |f(x,y)\cdot g(x,y)|\, dxdy \to 0 \\ &\int_{\mathbb{R}^n} \psi_\epsilon(x)\partial_x f(x,y)\cdot g(x,y)dxdy \to \int_{\mathbb{R}^n} \partial_x f(x,y)\cdot g(x,y)dxdy \text{ and} \\ &\int_{\mathbb{R}^n} \psi_\epsilon(x)f(x,y)\cdot \partial_x g(x,y)dxdy \to \int_{\mathbb{R}^n} f(x,y)\cdot\partial_x g(x,y)dxdy\end{aligned}$$

where $C = \sup_{x\in\mathbb{R}} |\psi'(x)|$. Combining the last three equations with Eq. (42.8) shows

$$\int_{\mathbb{R}\times\mathbb{R}^{n-1}} \partial_x f(x,y) \cdot g(x,y) dx dy = -\int_{\mathbb{R}\times\mathbb{R}^{n-1}} f(x,y) \cdot \partial_x g(x,y) dx dy$$

as desired.

With this result we may give another proof of the Riemann Lebesgue Lemma.

Lemma 42.24. *For $f \in L^1(\mathbb{R}^n, m)$ let*

$$\hat{f}(\xi) := (2\pi)^{-n/2} \int_{\mathbb{R}^n} f(x) e^{-i\xi\cdot x} dm(x)$$

be the Fourier transform of f. Then $\hat{f} \in C_0(\mathbb{R}^n)$ and $\left\|\hat{f}\right\|_u \le (2\pi)^{-n/2} \|f\|_1$. (The choice of the normalization factor, $(2\pi)^{-n/2}$, in $\hat{f}$ is for later convenience.)

Proof. The fact that $\hat{f}$ is continuous is a simple application of the dominated convergence theorem. Moreover,

$$\left|\hat{f}(\xi)\right| \le \int |f(x)| \, dm(x) \le (2\pi)^{-n/2} \|f\|_1$$

so it only remains to see that $\hat{f}(\xi) \to 0$ as $|\xi| \to \infty$.

First suppose that $f \in C_c^\infty(\mathbb{R}^n)$ and let $\Delta = \sum_{j=1}^n \frac{\partial^2}{\partial x_j^2}$ be the Laplacian on $\mathbb{R}^n$. Notice that $\frac{\partial}{\partial x_j} e^{-i\xi\cdot x} = -i\xi_j e^{-i\xi\cdot x}$ and $\Delta e^{-i\xi\cdot x} = -|\xi|^2 e^{-i\xi\cdot x}$. Using Lemma 42.21 repeatedly,

$$\int \Delta^k f(x) e^{-i\xi\cdot x} dm(x) = \int f(x) \Delta_x^k e^{-i\xi\cdot x} dm(x) = -|\xi|^{2k} \int f(x) e^{-i\xi\cdot x} dm(x)$$
$$= -(2\pi)^{n/2} |\xi|^{2k} \hat{f}(\xi)$$

for any $k \in \mathbb{N}$. Hence $(2\pi)^{n/2} \left|\hat{f}(\xi)\right| \le |\xi|^{-2k} \left\|\Delta^k f\right\|_1 \to 0$ as $|\xi| \to \infty$ and $\hat{f} \in C_0(\mathbb{R}^n)$. Suppose that $f \in L^1(m)$ and $f_k \in C_c^\infty(\mathbb{R}^n)$ is a sequence such that $\lim_{k\to\infty} \|f - f_k\|_1 = 0$, then $\lim_{k\to\infty} \left\|\hat{f} - \hat{f}_k\right\|_u = 0$ and hence Hence $\hat{f} \in C_0(\mathbb{R}^n)$ because $C_0(\mathbb{R}^n)$ is complete. ∎

Corollary 42.25. *Let $X \subset \mathbb{R}^n$ be an open set and μ be a Radon measure on $\mathcal{B}_X$.*

1. Then $C_c^\infty(X)$ is dense in $L^p(\mu)$ for all $1 \le p < \infty$.

2. If $h \in L^1_{loc}(\mu)$ satisfies

$$\int_X f h d\mu = 0 \text{ for all } f \in C_c^\infty(X) \tag{42.9}$$

then $h(x) = 0$ for μ – a.e. x.

Proof. Let $f \in C_c(X)$, ϕ be as in Lemma 42.16, ϕ_t be as in Theorem 42.13 and set $\psi_t := \phi_t * (f 1_X)$. Then by Proposition 42.19 $\psi_t \in C^\infty(X)$ and by Lemma 42.8 there exists a compact set $K \subset X$ such that $\text{supp}(\psi_t) \subset K$ for all t sufficiently small. By Theorem 42.13, $\psi_t \to f$ uniformly on X as $t \downarrow 0$

1. The dominated convergence theorem (with dominating function being $\|f\|_\infty 1_K$), shows $\psi_t \to f$ in $L^p(\mu)$ as $t \downarrow 0$. This proves Item 1. because of the measure theoretic fact that $C_c(X)$ is dense in $L^p(\mu)$.
2. Keeping the same notation as above, the dominated convergence theorem (with dominating function being $\|f\|_\infty |h| 1_K$) implies

$$0 = \lim_{t\downarrow 0} \int_X \psi_t h d\mu = \int_X \lim_{t\downarrow 0} \psi_t h d\mu = \int_X f h d\mu.$$

Since this is true for all $f \in C_c(X)$, it follows by measure theoretic arguments that $h = 0$ a.e.

∎

42.2 Smooth Partitions of Unity

Theorem 42.26. *Let $V_1, \dots, V_k \subset_o \mathbb{R}^n$ and $\phi \in C_c^\infty\left(\cup_{i=1}^k V_i\right)$. Then there exists $\phi_j \in C_c^\infty(V_j)$ such that $\phi = \sum_i^k \phi_j$. If $\phi \ge 0$ one can choose $\phi_j \ge 0$.*

Proof. The proof will be split into two steps.

1. There exists $K_j \sqsubset\sqsubset V_j$ such that $\text{supp}\,\phi \subset \cup K_j$. Indeed, for all $x \in \text{supp}\,\phi$ there exists an open neighborhood N_x of x such that $\overline{N}_x \subset V_j$ for some j and $\overline{N}_x$ is compact. Now $\{N_x\}_{x\in\text{supp}\,\phi}$ covers $K := \text{supp}\,\phi$ and hence there exists a finite set $\Lambda \subset\subset K$ such that $K \subset \cup_{x\in\Lambda} N_x$.. Let $K_j := \cup\left\{\overline{N}_x : x \in \Lambda \text{ and } \overline{N}_x \subset V_j\right\}$. Then each K_j is compact, $K_j \subset V_j$ and $\text{supp}\,\phi = K \subset \bigcup_{j=1}^k K_j$.
2. By Corollary 42.20 there exists $\psi_j \in C_c^\infty(V_j, [0,1])$ such that $\psi_j := 1$ in the neighborhood of K_j. Now define

$$
\begin{aligned}
\phi_1 &= \phi\psi_1 \\
\phi_2 &= (\phi - \phi_1)\psi_2 = \phi(1 - \psi_1)\psi_2 \\
\phi_3 &= (\phi - \phi_1 - \phi_2)\psi_3 = \phi\{(1 - \psi_1) - (1 - \psi_1)\psi_2\}\psi_3 \\
&= \phi(1 - \psi_1)(1 - \psi_2)\psi_3 \\
&\vdots \\
\phi_k &= (\phi - \phi_1 - \phi_2 - \cdots - \phi_{k-1})\psi_k \\
&= \phi(1 - \psi_1)(1 - \psi_2)\dots(1 - \psi_{k-1})\psi_k
\end{aligned}
$$

By the above computations one finds that (a) $\phi_i \geq 0$ if $\phi \geq 0$ and (b)

$$\phi - \phi_1 - \phi_2 - \cdots - \phi_k = \phi(1 - \psi_1)(1 - \psi_2)\dots(1 - \psi_k) = 0.$$

since either $\phi(x) = 0$ or $x \notin \operatorname{supp}\phi = K$ and $1 - \psi_i(x) = 0$ for some i.

■

Corollary 42.27. *Let $V_1, \dots, V_k \subset_0 \mathbb{R}^n$ and K be a compact subset of $\cup_{i=1}^k V_i$. Then there exists $\phi_i \in C_c^\infty(V_i, [0,1])$ such $\sum_{i=1}^k \phi_i \leq 1$ with $\sum_{i=1}^k \phi_i = 1$ on a neighborhood of K.*

Proof. By Corollary 42.20 there exists $\phi \in C_c^\infty(\cup_{i=1}^k V_i, [0,1])$ such that $\phi = 1$ on a neighborhood of K. Now let $\{\phi_i\}_{i=1}^k$ be the functions constructed in Theorem 42.26. ■

43

Poisson and Laplace's Equation

For the majority of this section we will assume $\Omega \subset \mathbb{R}^n$ is a compact manifold with C^2 – boundary. Let us record a few consequences of the divergence theorem in Proposition 22.30 in this context. If $u, v \in C^2(\Omega^o) \cap C^1(\Omega)$ and $\int_\Omega |\triangle u| dx < \infty$ then

$$\int_\Omega \triangle u \cdot v dm = -\int_\Omega \nabla u \cdot \nabla v dm + \int_{\partial\Omega} v \frac{\partial u}{\partial n} d\sigma \tag{43.1}$$

and if further $\int_\Omega \{|\triangle u| + |\triangle v|\} dx < \infty$ then

$$\int_\Omega (\triangle u v - \triangle v\, u) dm = \int_{\partial\Omega} \left(v\frac{\partial u}{\partial n} - \frac{\partial v}{\partial n} u \right) d\sigma. \tag{43.2}$$

Lemma 43.1. *Suppose $u \in C^2(\Omega^o) \cap C^1(\Omega)$, $\Delta u = 0$ on Ω^o and $u = 0$ on $\partial\Omega$. Then $u \equiv 0$. Similarly if $\Delta u = 0$ on Ω^o and $\partial_n u = 0$ on $\partial\Omega$, then u is constant on each connected component of Ω.*

Proof. Letting $v = u$ in Eq. (43.1) shows in either case that

$$0 = -\int_\Omega \nabla u \cdot \nabla u dm + \int_{\partial\Omega} u \frac{\partial u}{\partial n} d\sigma = -\int_\Omega |\nabla u|^2 dm.$$

This then implies $\nabla u = 0$ on Ω^o and hence u is constant on the connected component of Ω^o. If $u = 0$ on $\partial\Omega$, these constants must all be zero. ■

Proposition 43.2 (Laplacian on radial functions). *Suppose $f(x) = F(|x|)$, then*

$$\triangle f(x) = \frac{1}{r^{n-1}} \frac{d}{dr}(r^{n-1}F'(r))\Big|_{r=|x|} = F''(|x|) + \frac{(n-1)}{|x|}F'(|x|). \tag{43.3}$$

In particular $\Delta F(|x|) = 0$ implies $\frac{d}{dr}(r^{n-1}F'(r)) = 0$ and hence $F'(r) = \tilde{A}r^{1-n}$. That is to say

$$F(r) = \begin{cases} Ar^{2-n} + B & \text{if } n \neq 2 \\ A\ln r + B & \text{if } n = 2. \end{cases}$$

Proof. Since $(\partial_v f)(x) = F'(|x|)\ \partial_v|x| = F'(|x|)\hat{x} \cdot v$ where $\hat{x} = \frac{x}{|x|}$, $\nabla f(x) = F'(|x|)\hat{x}$. Hence for $g \in C_c^1(\mathbb{R}^n)$,

$$\begin{aligned}
\int_{\mathbb{R}^n} \triangle f(x) g(x) dx &= -\int_{\mathbb{R}^n} \nabla f(x) \cdot \nabla g(x)\ dx \\
&= -\int_{\mathbb{R}^n} F'(r)\hat{x} \cdot \nabla g(r\hat{x})\ dx \\
&= -\int_{S^{n-1}\times[0,\infty)} F'(r) \frac{d}{dr} g(r\omega)\ r^{n-1} dr\ d\sigma(\omega) \\
&= \int_{S^{n-1}\times[0,\infty)} \frac{d}{dr}(r^{n-1}F'(r)) g(r\omega) dr\ d\sigma(\omega) \\
&= \int_{S^{n-1}\times[0,\infty)} \frac{1}{r^{n-1}} \frac{d}{dr}\ (r^{n-1}F'(r))\ g(r\omega) r^{n-1}\ dr\ d\sigma(\omega) \\
&= \int_{\mathbb{R}^n} \frac{1}{r^{n-1}} \frac{d}{dr}(r^{n-1}F'(r))\Big|_{r=|x|} g(x)\ dx.
\end{aligned}$$

Since this is valid for all $g \in C_c^1(\mathbb{R}^n)$, Eq. (43.3) is valid. Alternatively, we may simply compute directly as follows:

$$\begin{aligned}
\triangle f(x) &= \nabla \cdot [F'(|x|)\hat{x}] = \nabla F'(|x|) \cdot \hat{x} + F'(|x|)\nabla \cdot \hat{x} \\
&= F''(|x|)\hat{x} \cdot \hat{x} + F'(|x|)\nabla \cdot \frac{x}{|x|} \\
&= F''(|x|) + F'(|x|) \left\{ \frac{n}{|x|} - \frac{x}{|x|^2} \cdot \hat{x} \right\} \\
&= F''(|x|) + \frac{(n-1)}{|x|} F'(|x|).
\end{aligned}$$

■

Notation 43.3 *For $t > 0$, let*

$$\alpha(t) := \alpha_n(t) := c_n \begin{cases} \frac{1}{t^{n-2}} & \text{if } n \neq 2 \\ \ln t & \text{if } n = 2, \end{cases} \tag{43.4}$$

where $c_n = \begin{cases} \frac{1}{(n-2)\sigma(S^{n-1})} & \text{if } n \neq 2 \\ -\frac{1}{2\pi} & \text{if } n = 2. \end{cases}$ *Also let*

$$\phi(y) = \phi_n(y) := \alpha(|y|) = c_n \begin{cases} \frac{1}{|y|^{n-2}} & \text{if } n \neq 2 \\ \ln|y| & \text{if } n = 2. \end{cases} \tag{43.5}$$

An important feature of α is that

$$\alpha'(t) = c_n \begin{cases} -(n-2)\frac{1}{t^{n-1}} & \text{if } n \neq 2 \\ \frac{1}{t} & \text{if } n = 2 \end{cases} = -\frac{1}{\sigma(S^{n-1})}\frac{1}{t^{n-1}} \tag{43.6}$$

for all n. This then implies, for all n, that

$$\nabla\phi(x) = \nabla\left[\alpha\left(|x|\right)\right] = \alpha'(|x|)\hat{x} = -\frac{1}{\sigma(S^{n-1})}\frac{1}{|x|^{n-1}}\hat{x} = -\frac{1}{\sigma(S^{n-1})}\frac{1}{|x|^{n}}x. \tag{43.7}$$

One more piece of notation will be useful in the sequel.

Notation 43.4 (Averaging operator) *Suppose μ is a finite measure on some space Ω, we will define*

$$\fint_{\Omega} f d\mu := \frac{1}{\mu(\Omega)}\int_{\Omega} f d\mu.$$

For example if Ω is a compact manifold with C^2 – boundary in $\mathbb{R}^n$ then

$$\fint_{\Omega} f(x)dx = \frac{1}{m(\Omega)}\int_{\Omega} f(x)dx = \frac{1}{Vol(\Omega)}\int_{\Omega} f(x)dx$$

and

$$\fint_{\partial\Omega} f d\sigma = \frac{1}{\sigma(\partial\Omega)}\int_{\partial\Omega} f(x)dx = \frac{1}{Area(\partial\Omega)}\int_{\partial\Omega} f(x)dx.$$

Theorem 43.5. *Let Ω be a compact manifold with C^2- boundary, $u \in C^2(\Omega^o) \cap C^1(\Omega)$ with $\int_{\Omega}|\Delta u(y)|\, dy < \infty$. Then for $x \in \Omega$*

$$u(x) = \int_{\partial\Omega}\left(\phi(x-y)\frac{\partial u}{\partial n}(y) - u(y)\frac{\partial\phi(x-y)}{\partial n_y}\right)d\sigma - \int_{\Omega}\phi(x-y)\triangle u(y)dy. \tag{43.8}$$

Proof. Let $\psi(y) := \phi(x-y)$ and $\epsilon > 0$ be small so that $B_x(\epsilon) \subset \Omega$ and let $\Omega_\epsilon := \Omega \setminus B_x(\epsilon)$, see Figure 43.1 below.

Let us begin by observing

$$\int_{|x-y|\le\epsilon}\psi(y)\, dy = \int_{|y|\le\epsilon}\frac{1}{|y|^{n-2}}\, dy = \sigma(S^{n-1})\int_0^\epsilon \frac{1}{r^{n-2}}\, r^{n-1}\, dr$$
$$= \sigma(S^{n-1})\int_0^\epsilon r\, dr = \sigma(S^{n-1})\frac{\epsilon^2}{2}$$

when $n \neq 2$ and for $n = 2$ that

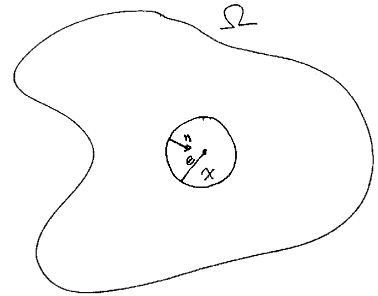


Fig. 43.1. Removing the region where ψ is singular from Ω.

$$\int_{|x-y|\le\epsilon}\psi(y)\, dy = \int_{|y|\le\epsilon}\ln|y|\, dy = \sigma(S^1)\int_0^\epsilon r\ln r\, dr$$
$$= 2\pi\left[\frac{1}{2}r^2\ln r - \frac{1}{4}r^2\right]_0^\epsilon = \pi\epsilon^2\left[\ln\epsilon - 1/2\right].$$

This shows $\psi \in L^1_{loc}(\Omega)$ and hence that $\psi\Delta u \in L^1(\Omega)$ and by dominated convergence theorem,

$$\int_{\Omega}\psi(y)\, \triangle u(y)\, dy = \lim_{\epsilon\downarrow 0}\int_{\Omega_\epsilon}\psi(y)\triangle u(y)\, dy.$$

Using Green's identity (Eq. (43.2) and Proposition 43.2) and $\Delta\psi = 0$ on Ω_ϵ, we find

$$\int_{\Omega_\epsilon}\triangle u(y)\psi(y)\, dy = \int_{\Omega_\epsilon}\triangle\psi(y)u(y)\, dy + \int_{\partial\Omega_\epsilon}\left(\psi\frac{\partial u}{\partial n} - \frac{\partial\psi}{\partial n}u\right)d\sigma$$
$$= \int_{\partial\Omega}\left(\psi\frac{\partial u}{\partial n} - \frac{\partial\psi}{\partial n}u\right)d\sigma$$
$$+ \int_{\partial\Omega_\epsilon\setminus\partial\Omega}\left(\psi\frac{\partial u}{\partial n} - \frac{\partial\psi}{\partial n}u\right)d\sigma. \tag{43.9}$$

Working on the last term in Eq. (43.9) we have, for $n \neq 2$,

$$
\begin{aligned}
\int_{\partial B(x,\epsilon)} \psi(y)\,\frac{\partial u}{\partial \mathbf{n}}(y) &= \int_{|y|=\epsilon} \psi(x+\omega)\,\frac{\partial u}{\partial \mathbf{n}}(x+\omega)d\sigma(\omega) \\
&= \int_{|\omega|=1} \psi(x+\epsilon\omega)\,\frac{\partial u}{\partial \mathbf{n}}(x+\epsilon\omega)\epsilon^{n-1}d\sigma(\omega) \\
&= \int_{|\omega|=1} \frac{1}{\epsilon^{n-2}}\,\frac{\partial u}{\partial \mathbf{n}}(x+\epsilon\omega)\epsilon^{n-1}d\sigma(\omega) \\
&= \epsilon\int_{|\omega|} \frac{\partial u}{\partial \mathbf{n}}\,(x+\epsilon\omega)d\sigma(\omega) \to 0 \text{ as } \epsilon \downarrow 0.
\end{aligned}
$$

Similarly when $n=2$,

$$
\int_{\partial B(x,\epsilon)} \psi(y)\,\frac{\partial u}{\partial \mathbf{n}}(y) = \epsilon\ln\epsilon \int_{|\omega|=1} \frac{\partial u}{\partial \mathbf{n}}(x+\epsilon\omega)d\sigma(\omega) \to 0 \text{ as } \epsilon \downarrow 0.
$$

Using Eq. (43.7) and $n(y) = -\widehat{(y-x)}$ as in Figure 43.2 we find

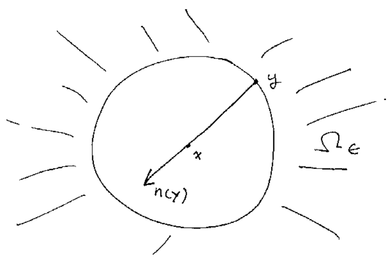


Fig. 43.2. The outward normal to Ω_ϵ is the inward normal to $B(x,\epsilon)$.

$$
\begin{aligned}
\frac{\partial \psi}{\partial \mathbf{n}}(y) &= \nabla_y\phi(y-x)\cdot n(y) = -\frac{1}{\sigma(S^{n-1})}\frac{1}{|y-x|^n}\,(y-x)\cdot\left(-\widehat{(y-x)}\right) \\
&= \frac{1}{\sigma(S^{n-1})}\frac{1}{\epsilon^{n-1}}
\end{aligned}
\tag{43.10}
$$

and therefore

$$
\begin{aligned}
-\int_{\partial\Omega_\epsilon\setminus\partial\Omega} u\,\frac{\partial\psi}{\partial \mathbf{n}}\,d\sigma(y) &= -\frac{1}{\sigma(S^{n-1})}\frac{1}{\epsilon^{n-1}}\int_{\partial B(x,\epsilon)} u(y)d\sigma(y) \\
&= -\fint_{|\omega|=1} u(x+\epsilon\omega)d\sigma(\omega) \to -u(x) \text{ as } \epsilon\downarrow 0
\end{aligned}
$$

by the dominated convergence theorem. So we may pass to the limit in Eq. (43.9) to find

$$
\int_\Omega \psi(y)\,\triangle u(y)dy = \int_{\partial\Omega}\left(\psi(y)\frac{\partial u}{\partial \mathbf{n}} - u\,\frac{\partial\psi}{\partial \mathbf{n}}\right)d\sigma(y) - u(x)
$$

which is equivalent to Eq. (43.8). ∎

The following Corollary gives an easy but useful extension of Theorem 43.5It will be us

Corollary 43.6. *Keeping the same notation as in Theorem 43.5. Further assume that $h \in C^2(\Omega^o)\cap C^1(\Omega)$ and $\Delta h = 0$ and set $G(y) := \phi(x-y)+h(y)$. Then we still have the representation formula*

$$
u(x) = \int_{\partial\Omega}\left(G(y)\frac{\partial u}{\partial \mathbf{n}}(y) - u(y)\frac{\partial G(y)}{\partial \mathbf{n}}\right)d\sigma - \int_\Omega G(y)\triangle u(y)dy. \tag{43.11}
$$

Proof. By Green's identity (Proposition 22.30) with $v=h$,

$$
\int_\Omega \triangle u\ hdm = \int_\Omega(\triangle uh - \triangle h\ u)dm = \int_{\partial\Omega}\left(h\frac{\partial u}{\partial n} - \frac{\partial h}{\partial n}u\right)d\sigma,
$$

i.e.

$$
0 = -\int_\Omega \triangle u\ h\ dm + \int_{\partial\Omega}\left(h\frac{\partial u}{\partial n} - \frac{\partial h}{\partial n}u\right)d\sigma. \tag{43.12}
$$

Eq. (43.11) now follows by adding Eqs. (43.8) and (43.12). ∎

Corollary 43.7. *For all $u \in C_c^2(\mathbb{R}^n)$,*

$$
-\int_{\mathbb{R}^n} \triangle u(y)\phi(y)dy = u(0). \tag{43.13}
$$

Proof. Let $\Omega = B(0,R)$ where R is chosen so large that $\mathrm{supp}(g)\subset\Omega$, then by Theorem 43.5,

$$
\begin{aligned}
u(0) &= \int_{\partial\Omega}\left(\phi(y)\frac{\partial u}{\partial v}(y) - u(y)\frac{\partial\phi(y)}{\partial v_y}\right)d\sigma - \int_\Omega \phi(y)\triangle u(y)dy \\
&= -\int_\Omega \phi(y)\triangle u(y)dy.
\end{aligned}
$$

∎

Remark 43.8. We summarize (43.13) by saying $-\triangle\phi = \delta$.

Formally we expect for reasonable functions ρ that

$$\Delta(\phi * \rho) = \Delta\phi * \rho = -\delta * \rho = -\rho.$$

Theorem 43.9. *Suppose $\Omega \subset_o \mathbb{R}^n$, $\rho \in C^2(\Omega) \cap L^1(\Omega)$ and*

$$u(x) := \int_\Omega \phi(x-y)\rho(y)dy = (\phi * 1_\Omega \rho)(x),$$

then

$$-\triangle u = \rho \text{ on } \Omega.$$

Proof. First assume that $\rho \in C_c^2(\Omega)$ in which case we may set $\rho := 1_\Omega \rho \in C_c^2(\mathbb{R}^n)$. Therefore

$$u(x) = \int_{\mathbb{R}^n} \rho(y) \frac{1}{|x-y|^{n-2}} dy = \int_{\mathbb{R}^n} \rho(x-y) \frac{1}{|y|^{n-2}} dy$$

and so we may differentiate under the integral to find

$$\triangle u(x) = \int_{\mathbb{R}^n} \triangle_x \rho(x-y) \, \frac{1}{|y|^{n-2}} \, dy = -\rho(x)$$

where the last equality follows from Corollary 43.7.

For $\rho \in C^2(\Omega) \cap L^1(\Omega)$ and $x_0 \in \Omega$, choose $\alpha \in C_c^\infty(\Omega, [0,1])$ such that $\alpha = 1$ in a neighborhood of x_0 and let $\beta := 1 - \alpha$. Then $u = (\phi * \alpha\rho) + (\phi * \beta 1_\Omega \rho)$ and so

$$\Delta u = \Delta(\phi * \alpha\rho) + \Delta(\phi * \beta 1_\Omega \rho). \tag{43.14}$$

By what we have just proved

$$\Delta(\phi * \alpha\rho)(x) = -(\alpha\rho)(x) = -\rho(x) \text{ for } x \text{ near } x_0. \tag{43.15}$$

Since $\beta = 0$ near x_0 and

$$(\phi * \beta 1_\Omega \rho)(x) = \int_\Omega \phi(x-y)\beta(y)\rho(y)dy,$$

we may differentiate past the integral to learn

$$\Delta(\phi * \beta 1_\Omega \rho)(x) = \int_\Omega \Delta_x \phi(x-y)\beta(y)\rho(y)dy = 0 \tag{43.16}$$

for x near x_0. and this completes the proof. The combination of Eqs. (43.14 – 43.16) completes the proof. ■

43.1 Harmonic and Subharmonic Functions

Definition 43.10 (HarmonicFunctions). *Let $\Omega \subset_o \mathbb{R}^n$. A function $u \in C^2(\Omega)$ is said to be* ***harmonic (subharmonic)*** *on Ω if $\Delta u = 0$ ($\Delta u \geq 0$) on Ω.*

Because of the Cauchy Riemann equations, the real and imaginary parts of holomorphic functions are harmonic. For example $z^2 = (x^2 - y^2) + 2ixy$ implies $(x^2 - y^2)$ and xy are harmonic functions on the plane. Similarly,

$$e^z = e^x \cos y + ie^x \sin y \text{ and}$$
$$\ln(z) = \ln r + i\theta$$

implies

$$e^x \cos y, \; e^x \sin y, \; \ln r, \text{ and } \theta(x,y)$$

are harmonic functions on their domains of definition.

Remark 43.11. If we can choose h in Corollary 43.6 so that $G = 0$ on $\partial\Omega$, then Eq. (43.11) gives

$$u(x) = -\int_\Omega G(y)\triangle u(y)dy - \int_{\partial\Omega} u \frac{\partial G(y)}{\partial v} d\sigma \tag{43.17}$$

which shows how to recover $u(x)$ from Δu on Ω and u on $\partial\Omega$. The next theorem is a consequence of this remark.

Theorem 43.12 (Mean Value Property). *If $\triangle u = 0$ on Ω and $\overline{B(x,r)} \subset \Omega$ then*

$$u(x) = \frac{1}{\sigma(\partial B(x,r))} \int_{\partial B(x,r)} u(y)\, d\sigma(y) =: \fint_{\partial B(x,r)} u \; d\sigma \tag{43.18}$$

More generally if $\Delta u \geq 0$ on Ω, then

$$u(x) \leq \fint_{\partial B(x,r)} u \; d\sigma \tag{43.19}$$

Proof. For $y \in B(x,r)$,

$$G(y) = \phi(x-y) - \alpha(r) = \alpha(|x-y|) - \alpha(r)$$

where α is defined as in Eq. (43.4). Then $G(y) = 0$ for $y \in \partial B(x,r)$ and $G(y) > 0$ for all $y \in B(x,r)$ because α is decreasing as is seen from Eq. (43.6). From Eq. (43.10) (using now that n is the outward normal to $B(x,r)$),

$$\frac{\partial G}{\partial \mathbf{n}}(x + r\omega) = -\frac{1}{\sigma(S^{n-1})r^{n-1}} \text{ for } |\omega| = 1$$

and so according to Eq. (43.17) we have

$$u(x) = \frac{1}{r^{n-1}\sigma(S^{n-1})} \int_{\partial B(x,r)} u d\sigma - \int_{B(x,r)} G(y)\triangle u(y)dy$$
$$= \fint_{\partial B(x,r)} u \, d\sigma - \int_{B(x,r)} G(y)\triangle u(y)dy. \tag{43.20}$$

This completes the proof since $G(y) > 0$ for all $y \in B(x,r)$. ■

Remark 43.13 (Mean value theorem). Assuming $\overline{B(x,R)} \subset \Omega$ and multiplying Eq. (43.18) (Eq. (43.19)) by

$$\sigma(\partial B(x,r)) = \sigma(S^{n-1})r^{n-1}$$

and then integrating on $0 \le r \le R$, implies

$$u(x)m(B(x,R)) = (\text{or } \le) \int_0^R dr \int_{\partial B(x,r)} u(y)\, d\sigma(y)$$
$$= \int_0^R dr\; r^{n-1} \int_{S^{n-1}} u(x+r\omega)\, d\sigma(\omega) = \int_{B(x,R)} udm.$$

Therefore if $\Delta u = 0$ or $\Delta u \ge 0$ then

$$u(x) = \fint_{B(x,R)} udm \text{ or } u(x) \le \fint_{B(x,R)} udm \text{ respectively} \tag{43.21}$$

for all $\overline{B(x,R)} \subset \Omega$.

Proposition 43.14 (Converse of the mean value property). *If $u \in C(\Omega)$ (or more generally measurable and locally bounded) and*

$$u(x) = \fint_{\partial B(x,r)} u(y)d\sigma(y) \tag{43.22}$$

for all $x \in \Omega$ and $r > 0$ such that $\overline{B(x,r)} \subset \Omega$, then $u \in C^\infty(\Omega)$ and $\triangle u = 0$. Similarly, if $u \in C^2(\Omega)$ and $x \in \Omega$ and

$$u(x) \le \fint_{\partial B(x,r)} u(y)d\sigma(y) \tag{43.23}$$

for all r sufficiently small, then $\Delta u(x) \ge 0$.

Proof. First assume $u \in C(\Omega)$ and Eq. (43.22) hold which implies

$$u(x) = \fint_S u(x+r\omega)d\sigma(\omega) \tag{43.24}$$

for all $x \in \Omega$ and r sufficiently small, where $S = S^{n-1}$ denotes the unit sphere in $\mathbb{R}^n$. Let $\eta \in C_c^\infty(\mathbb{R}^n, [0,\infty))$ such that $\eta(0) > 0$ and

$$1 = \int_{\mathbb{R}^n} \eta(|x|^2)dx = \sigma(S)\int_0^\infty \eta(r^2)r^{n-1}dr$$

and for $\epsilon > 0$ let $\eta_\epsilon(x) = \epsilon^{-n}\eta\left(\frac{|x|^2}{\epsilon^2}\right) \in C_c^\infty(\mathbb{R}^n)$ and $u_\epsilon(x) = \eta_\epsilon * u(x)$. Then for any $x_0 \in \Omega$ and $\epsilon > 0$ sufficiently small, u_ϵ is a well defined smooth function near x_0. Moreover for x near x_0 we have

$$u_\epsilon(x) = \int_{\mathbb{R}^n} \eta_\epsilon(x-y)u(y)dy = \int_0^\infty dr\; r^{n-1} \int_{|\omega|=1} \eta_\epsilon(r\omega)u(x+r\omega)d\sigma(\omega)$$
$$= \int_0^\infty dr r^{n-1} \int_{|\omega|=1} \epsilon^{-n}\eta\left(\frac{r^2}{\epsilon^2}\right)u(x+r\omega)d\sigma(\omega)$$
$$= u(x)\sigma(S)\int_0^\infty dr\; r^{n-1}\epsilon^{-n}\eta\left(\frac{r^2}{\epsilon^2}\right) = u(x)$$

which shows u is smooth near x_0.

Now suppose that $u \in C^2$, and u satisfies Eq. (43.23), $x \in \Omega$ and $|r| < \epsilon$ with ϵ sufficiently small so that

$$f(r) := \fint_{\partial B(x,r)} u\; d\sigma = \fint_{S^{n-1}} u(x+r\omega)d\sigma(\omega)$$

is well defined. Clearly $f \in C^2(-\epsilon,\epsilon)$, f is an even function of r so $f'(0) = 0$, $f(0) = u(x)$ and $f(r) \ge f(0)$. From these conditions it follows that $f''(0) \ge 0$ for otherwise we would find from Taylor's theorem that $f(r) < f(0)$ for $0 < |r| < \epsilon$. On the other hand

$$0 \le f''(0) = \fint_{S^{n-1}} \left(\partial_\omega^2 u\right)(x)d\sigma(\omega)$$
$$= \fint_{S^{n-1}} \left(\partial_i\partial_j u\right)(x)\omega_i\omega_j d\sigma(\omega)$$
$$= \left(\partial_i\partial_j u\right)(x)\delta_{ij} \fint_{S^{n-1}} \omega_i^2 d\sigma(\omega) = \frac{1}{n}\Delta u(x). \tag{43.25}$$

wherein we have used the symmetry of $d\sigma$ on S^{n-1} to conclude

$$\fint_{S^{n-1}} \omega_i \omega_j d\sigma(\omega) = 0 \text{ if } i \neq j$$

and

$$\begin{aligned} \fint_{S^{n-1}} \omega_i^2 d\sigma(\omega) &= \frac{1}{n} \sum_{j=1}^{n} \fint_{S^{n-1}} \omega_j^2 d\sigma(\omega) \\ &= \frac{1}{n} \fint_{S^{n-1}} |\omega|^2 \, d\sigma(\omega) = \frac{1}{n} \;\forall\; i. \end{aligned}$$

Alternatively, by the divergence theorem,

$$\begin{aligned} \fint_{S^{n-1}} \omega_i \omega_j d\sigma(\omega) &= \fint_{S^{n-1}} \omega_i e_j \cdot n(\omega) d\sigma(\omega) \\ &= \frac{1}{\sigma(S^{n-1})} \int_{B(0,1)} \nabla \cdot (x_i e_j) \, dm \\ &= \frac{1}{\sigma(S^{n-1})} m(B(0,1)) \delta_{ij} = \frac{1}{n} \delta_{ij}. \end{aligned}$$

This completes the proof since if u satisfies (43.22) then f is constant and it follows from Eq. (43.25) that $\Delta u(x) = 0$.

Second proof of the last statement. Now that we know u is C^2 we have by Eq. (43.20) that

$$\int_{B(x,r)} G(y) \triangle u(y) dy = \fint_{\partial B(x,r)} u \, d\sigma - u(x) \geq 0$$

and since with α as in Eq. (43.4),

$$\begin{aligned} \int_{B(x,r)} G(y) \triangle u(y) dy &= \int_{B(0,r)} G(x+y) \triangle u(x+y) dy \\ &= \int_0^r \rho^{n-1} d\rho \int_{S^n} d\omega G(x+\rho\omega) \triangle u(x+\rho\omega) \\ &= \int_0^r \rho^{n-1} d\rho \, (\alpha(\rho) - \alpha(r)) \int_{S^n} d\omega \triangle u(x+\rho\omega) \\ &\cong \Delta u(x) \sigma(S^{n-1}) \int_0^r \rho^{n-1} d\rho \, (\alpha(\rho) - \alpha(r)) \\ &= \Delta u(x) \sigma(S^{n-1}) c_n \left\{ \frac{r^2}{2} - \frac{r^n}{n r^{n-2}} \right\} \\ &= b_n r^2 \Delta u(x) \end{aligned}$$

where b_n is a positive constant. From this it follows that $\Delta u(x) \geq 0$.

Third proof of the last statement. If $u \in C^2(\Omega)$ satisfies expand $u(x+r\omega)$ in a Taylor series

$$u(x+r\omega) = u(x) + r\nabla u(x) \cdot \omega + \frac{r^2}{2} \partial_\omega^2 u(x) + o(r^3),$$

and integrate on ω to find

$$\begin{aligned} \fint_{\partial B(x,r)} u d\sigma &= \fint_{S^{n-1}} u(x+r\omega) d\sigma(\omega) \\ &= \fint_{S^{n-1}} \left[u(x) + r\nabla u(x) \cdot \omega + r^2 \frac{1}{2} \partial_\omega^2 u(x) + \dots \right] d\sigma(\omega) \\ &= u(x) + \frac{1}{2} r^2 \Delta u(x) + o(r^2). \end{aligned}$$

Thus if u satisfies Eq. (43.22) Eq. (43.23) we conclude

$$\begin{aligned} u(x) &= u(x) + \frac{1}{2} r^2 \Delta u(x) + o(r^2) \text{ or} \\ u(x) &\leq u(x) + \frac{1}{2} r^2 \Delta u(x) + o(r^2) \end{aligned}$$

from which we conclude $\Delta u(x) = 0$ or $\Delta u(x) \geq 0$ respectively.

Fourth proof of the statement: If u satisfies Eq. (43.22) then $\Delta u = 0$. Since we already know u is smooth, it is permissible to differentiate Eq. (43.24) in r to learn,

$$\begin{aligned} 0 &= \fint_{S^{n-1}} \nabla u(x+r\omega) \cdot \omega \, d\sigma(\omega) = \fint_{S^{n-1}} \frac{\partial u}{\partial \mathbf{n}}(x+r\omega) \, d\sigma(\omega) \\ &= \frac{1}{\sigma(S^{n-1}) r^{n-1}} \int_{\partial B(x,r)} \nabla u \cdot \mathbf{n} \, d\sigma = \frac{1}{\sigma(S^{n-1}) r^{n-1}} \int_{B(x,r)} \triangle u \, dm. \end{aligned}$$

Dividing this equation by r and letting $r \downarrow 0$ shows $\Delta u(x) = 0$. ∎

Corollary 43.15 (Smoothness of Harmonic Functions). *If $u \in C^2(\Omega)$ and $\triangle u = 0$ then $u \in C^\infty(\Omega)$. (Soon we will show u is real analytic, see Theorem 43.16 of Corollary 43.34 below.)*

Theorem 43.16 (Bounds on Harmonic functions). *Suppose u is a Harmonic function on $\Omega \subset \mathbb{R}^n$, $x_0 \in \Omega$, α is a multi-index with $k := |\alpha|$ and* $0 < r < \operatorname{dist}(x_o, \partial\Omega)$. *Then*

$$|D^\alpha u(x_0)| \leq \frac{C_k}{r^{n+k}} \|u\|_{L^1(B(x_0,r))} \leq \frac{C_k}{\operatorname{dist}(x_o, \partial\Omega)^{n+k}} \|u\|_{L^1(\Omega)} \tag{43.26}$$

where $C_k = \frac{(2^{n+1} n \; k)^k}{\alpha(n)}$. In particular one shows that u is real analytic in Ω.

Proof. Let $\eta_\epsilon(x)$ be constructed as in the proof of Proposition 43.14 so that $u(x) = u * \eta_\epsilon(x)$. Therefore, $D^\alpha u(x) = u_* D^\alpha \eta_\epsilon(x)$ and hence

$$|D^\alpha u(x_0)| \le \|u\|_{L^1(B(x_0,\epsilon))} \|D^\alpha \eta_\epsilon\|_{L^\infty}.$$

Now

$$D^\alpha \eta_\epsilon(x) = \epsilon^{-n} \frac{1}{\epsilon^{|\alpha|}} (D^\alpha \eta)(\frac{x}{\epsilon})$$

so that

$$|D^\alpha \eta_\epsilon(x)| = \epsilon^{-n} \frac{1}{\epsilon^{|\alpha|}} \left|(D^\alpha \eta)(\frac{x}{\epsilon})\right| \le C_\alpha \frac{1}{\epsilon^{|\alpha|+n}} \cong C_\alpha \frac{1}{r^{|\alpha|+n}}$$

where the last identity is gotten by taking ϵ comparable to r. Putting this all together then implies that

$$|D^\alpha u(x_0)| \le \frac{1}{r^{n+|\alpha|}} \|D^\alpha \eta\|_{L^\infty} \|u\|_{L^1(B(x_0,r))}$$

which is an inequality of the form in Eq. (43.26). To get the desired constant we will have to work harder. This is done in Theorem 7. on p. 29 of the book. The idea is to use $D^\alpha u$ is harmonic for all α and therefore,

$$\begin{aligned} D^\alpha u(x_0) &= \fint_{B(x_0,\rho)} D^\alpha u dm = \fint_{B(x_0,\rho)} \partial_i D^\beta u dm \\ &= \frac{n}{\sigma(S^{n-1})\rho^n} \int_{B(x_0,\rho)} \partial_i D^\beta u dm \\ &= \frac{n}{\sigma(S^{n-1})\rho^n} \int_{\partial B(x_0,\rho)} D^\beta u n_i d\sigma \end{aligned}$$

so that

$$|D^\alpha u(x_0)| \le \frac{n}{\rho} \left\|D^\beta u\right\|_{L^\infty(B(x_0,\rho))}$$

and for $\alpha = 0$ and $x \in B(x_0, r/2)$ we have

$$|u(x)| \le \fint_{B(x,r/2)} |u| \, dm \le \frac{1}{|B(0,1)|} \left(\frac{2}{r}\right)^n \|u\|_{L^1(B(x_0,r)}.$$

Using this and similar inequalities along with a tricky induction argument one gets the desired constants. The details are in Theorem 7. p. 29 and Theorem 10 p.31 of the book. (See also Corollary 43.34 below for another proof of analyticity of u.) ■

Corollary 43.17 (Liouville's Theorem). *Suppose $u \in C^2(\mathbb{R}^n)$, $\Delta u = 0$ on $\mathbb{R}^n$ and $|u(x)| \le C(1 + |x|^N)$ for all $x \in \mathbb{R}^n$. Then u is a polynomial of degree at most N.*

Proof. We have seen there are constants $C_{|\alpha|} < \infty$ such that

$$\begin{aligned} |D^\alpha u(x_0)| &\le C_{|\alpha|} \|u\|_{L^1(B(x_0,r))} \frac{1}{r^{n+|\alpha|}} \\ &\le \widetilde{C}_{|\alpha|} r^n \|u\|_{L^\infty(B(x_0,r))} \cdot \frac{1}{r^{n+|\alpha|}} \\ &\cong C \frac{(1+r^N)}{r^{|\alpha|}} \to 0 \text{ as } r \to \infty \end{aligned}$$

when if $|\alpha| > N$. Therefore $D^\alpha u := 0$ for all $|\alpha| > N$ and the the result follows by Taylor's Theorem with remainder,

$$u(x) = \sum_{|\alpha| \le N} \frac{D^\alpha u(x_0)(x - x_0)^\alpha}{\alpha!}.$$

■

Corollary 43.18 (Compactness of Harmonic Functions). *Suppose $\Omega \subset_o \mathbb{R}^n$ and $u_n \in C^2(\Omega)$ is a sequence of harmonic functions such that for each compact set $K \subset \Omega$,*

$$C_K := \sup \left\{ \int_K |u_n| \, dm : n \in \mathbb{N} \right\} < \infty.$$

Then there is a subsequence $\{v_n\} \subset \{u_n\}$ which converges, along with all of its derivatives, uniformly on compact subsets of Ω to a harmonic function u.

Proof. An application of Theorem 43.16 shows that for each compact set $K \subset \Omega$, $\sup_n |\nabla u_n|_{L^\infty(K)} < \infty$ and hence by the locally compact form of the Arzela-Ascolli theorem, there is a subsequence $\{v_n\} \subset \{u_n\}$ which converges uniformly on compact subsets of Ω to a continuous function $u \in C(\Omega)$. Passing to the limit in the mean value theorem for harmonic functions along with the converse to the mean value theorem, Proposition 43.14, shows u is harmonic on Ω. Since $v_m \to u$ uniformly on compacts it follows for any $K \sqsubset\sqsubset \Omega$ that $\int_K |u - v_n| \, dm \to 0$. Another application of Theorem 43.16 then shows $D^\alpha v_n \to D^\alpha u$ uniformly on compacts. ■

In light of Proposition 43.14, we will extend the notion of subharmonicity as follows.

Definition 43.19 (Subharmonic Functions). *A function $u \in C(\Omega)$ is said to be **subharmonic** if for all $x \in \Omega$ and all $r > 0$ sufficiently small,*

$$u(x) \le \fint_{\partial B(x,r)} u \, d\sigma.$$

The reason for the name subharmonic should become apparent from Corollary 43.26 below.

Remark 43.20. Suppose that $u, v \in C(\Omega)$ are subharmonic functions then so is $u+v$. Indeed,

$$u(x)+v(x) \le \fint_{\partial B(x,r)} u\, d\sigma + \fint_{\partial B(x,r)} v\, d\sigma = \fint_{\partial B(x,r)} (u+v)\, d\sigma.$$

Theorem 43.21 (Harnack's Inequality). *Let V be a precompact open and connected subset of Ω. Then there exists $C = C(V,\Omega)$ such that*

$$\sup_V u \le C \inf_V u \tag{43.27}$$

for all non-negative harmonic functions, u, on Ω.

Proof. Let $r = \frac{1}{4}\mathrm{dist}(V, \Omega^c)$ and $x \in V$ (as in Figure 43.3) and $|y-x| \le r$,

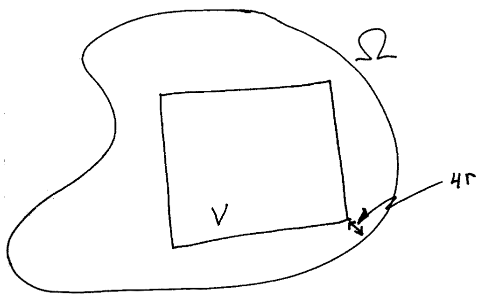


Fig. 43.3. A pre-compact region $V \subset \Omega$.

then by the mean value equality in Eq. (43.21) of Remark 43.13,

$$\begin{aligned} u(x) &= \fint_{B(x,2r)} u(z)dz = \frac{1}{m(B(0,1))(2r)^n} \int_{B(x,2r)} u(z)dz \\ &\ge \frac{1}{m(B(0,1))(2r)^n} \int_{B(y,r)} u(z)dz = \frac{1}{2^n} \fint_{B(y,r)} u(z)dz = \frac{1}{2^n} u(y), \end{aligned}$$

see Figure 43.4. Therefore

$$u(x) \ge \frac{1}{2^n} u(y) \text{ for all } x, y \in V \text{ with } |x-y| \le r. \tag{43.28}$$

Since $\bar{V}$ is compact there exists a finite cover $\mathcal{S} := \{W_i\}_{i=1}^M$ of $\bar{V}$ consisting of balls with of radius r with centers $x_i \in \bar{V}$.

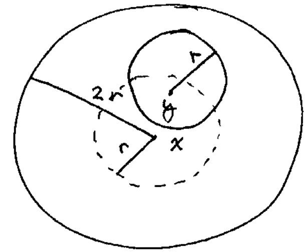


Fig. 43.4. Nested balls.

Claim: For all $x, y \in V$, there exists a chain $\{B_i\}_{i=1}^k \subset \mathcal{S}$ of distinct balls such that $x \in B_1$, $y \in B_k$ and $B_i \cap B_{i+1} \ne \phi$ for all $i = 1, \dots, k-1$.

Indeed, by connectedness of V there exists $\gamma \in C([0,1], V)$ such that $\gamma(0) = x$ while $\gamma(1) = y$. For sake of contradiction, suppose

$$T := \sup\{t \in [0,1] : \exists \text{ a chain as above } \ni \gamma(t) \in B_k\} < 1.$$

Since there are only finitely many possible chains (at most $\sum_{k=1}^M \frac{M!}{(M-k)!}$) there must be a chain $\{B_i\}_{i=1}^k \subset \mathcal{S}$ such that $\gamma(T) \in \bar{B}_k$. Let $B_{k+1} \in \mathcal{S}$ such that $B_{k+1} \ni \gamma(T)$. If $B_{k+1} = B_j$ with $j \le k$, then $\{B_i\}_{i=1}^j \subset \mathcal{S}$ is a chain such that $\gamma(T) \in B_j$. Otherwise, since $\gamma(T) \in B_{k+1} \cap \bar{B}_k$, it follows that $B_{k+1} \cap B_k \ne \emptyset$ and $\{B_i\}_{i=1}^{k+1} \subset \mathcal{S}$ is a chain such that $\gamma(T) \in B_{k+1}$. In either case we will have violated the definition of T and hence we must conclude $T = 1$. This proves the claim, since again using the fact that there are only a finite number of possible chains, there must be at least one chain for which $\gamma(1) \in B_k$.

To complete the proof, for any $x, y \in V$ use a chain as in the above claim to find a sequence of points $\{x_j\}_{j=1}^N \subset V$ with $N \le 2M$, $x_1 = x$, $x_N = y$ and $|x_{i+1} - x_i| < r$ for all i. Then by repeated use of Eq. (43.28) we may conclude

$$u(y) \le (2^n)^{2M}\, u(x).$$

Since $x, y \in V$ are arbitrary, this equation implies Eq. (43.27) with $C := 2^{4M}$. ■

Remark 43.22. It is not sufficient to assume u is sub-harmonic in Theorem 43.21. For example if $M > 0$, then $u(x) = Mx^2 + 1$ is sub-harmonic on $\mathbb{R}$, $\inf_{(-1,1)} u = 1$ while $\sup_{(-1,1)} u = M+1$. Since M is arbitrary, this would force $C = \infty$.

Also it is important that Ω is connected. Indeed, if $\Omega = \Omega_1 \cup \Omega_2$ with Ω_1 and Ω_2 being disjoint open sets, then let $u \equiv 1$ on Ω_1 and $u \equiv M$ on Ω_2 for any $M > 0$. This function is harmonic on Ω and hence $C \ge M$ for all $M > 0$, i.e. $C = \infty$.

Theorem 43.23 (Strong Maximum Principle). *Let $\Omega \subset \mathbb{R}^n$ be connected and open and $u \in C(\Omega)$ be a subharmonic function (see Definition 43.19). If $M = \sup_{x\in\Omega} u(x)$ is attained in Ω then $u := M$. (Notice that $u \in C^2(\Omega)$ and $\Delta u = 0$, then u is harmonic and hence in particular sub-harmonic.)*

Proof. Suppose there exists $x \in \Omega$ such that $M = u(x)$. If $\epsilon > 0$ is chosen so that $\overline{B(x,\epsilon)} \subset \Omega$ as in Figure 43.1 and $u(y) < M$ for some $y \in \partial B(x,\epsilon)$, then by the mean value inequality,

$$M = u(x) \le \fint_{\partial B(x,\epsilon)} u(y)d\sigma(y) < M$$

which is nonsense. Therefore $u := M$ on $\partial B(x,\epsilon)$ and since $\epsilon \in (0, \operatorname{dist}(x,\partial\Omega))$ we concluded that $u := M$ on $B(x,R)$ provided $\overline{B(x,R)} \subset \Omega$. Therefore $\{x \in \Omega : u(x) = M\}$ is both open and relatively closed in Ω and hence $\{x \in \Omega : u(x) = M\} = \Omega$ because Ω is connected. ■

Corollary 43.24. *If Ω is bounded open set $u \in C(\overline{\Omega})$ is subharmonic, then*

$$M := \max_{x\in\overline{\Omega}} u(x) = \max_{x\in\operatorname{bd}(\Omega)} u(x).$$

Again this corollary applies to $u \in C(\overline{\Omega}) \cap C^2(\Omega)$ such that $\Delta u = 0$.

Proof. By Theorem 43.23, if $x \in \Omega$ is an interior maximum of u, then $u = M$ on the connected component Ω_x of Ω which contains x. By continuity, u is constant on $\bar{\Omega}_x$ and in particular u takes on the value M on $\operatorname{bd}(\Omega)$. ■

Corollary 43.25. *Given $g \in C(\operatorname{bd}(\Omega))$, $f \in C(\Omega)$ there exists at most one function $u \in C^2(\Omega) \cap C(\overline{\Omega})$ such that $\triangle u = f$ on Ω and $u = g$ on $\operatorname{bd}(\Omega)$.*

Proof. If $v \in C^2(\Omega) \cap C(\overline{\Omega})$ is another such function then $w := u - v \in C^2(\Omega) \cap C(\overline{\Omega})$ satisfies $\Delta w = 0$ in Ω and $w = 0$ on $\operatorname{bd}(\Omega)$. Therefore applying Corollary 43.24 to w and $-w$ implies

$$\max_{x\in\overline{\Omega}} w(x) = \max_{x\in\operatorname{bd}(\Omega)} w(x) = 0 \text{ and } \min_{x\in\overline{\Omega}} w(x) = \min_{x\in\operatorname{bd}(\Omega)} w(x) = 0.$$

■

Corollary 43.26. *Suppose $g \in C(\operatorname{bd}(\Omega))$ and $u \in C^2(\Omega) \cap C(\overline{\Omega})$ such that $\triangle u = 0$ on Ω. Then $w \le u$ for any subharmonic function $w \in C(\bar{\Omega})$ such that $w \le g$ on $\operatorname{bd}(\Omega)$.*

Proof. The function $-u$ is subharmonic and so is $v = w - u$ by Remark 43.20. Since $v = w - g \le 0$ on $\operatorname{bd}(\Omega)$, it follows by Corollary 43.24 that $v \le 0$ on Ω, i.e. $w \le g$ on Ω. ■

43.2 Green's Functions

Notation 43.27 *Unless otherwise stated, for the rest of this section assume $\Omega \subset \mathbb{R}^n$ is a compact manifold with C^2 – boundary.*

For $x \in \Omega$, suppose there exists $h \in C^2(\Omega^o) \cap C^1(\Omega)$ which solves

$$\triangle h_x = 0 \text{ on } \Omega \text{ with } h_x(y) = \phi(x-y) \text{ for } y \in \partial\Omega. \tag{43.29}$$

Hence if we define

$$G(x,y) = \phi_x(y) - h_x(y) \tag{43.30}$$

then by the representation formula (Eq. (43.11) also see Remark 43.11) implies

$$u(x) = -\int_{\Omega} G(x,y)\triangle u(y)\,dy - \int_{\partial\Omega} \frac{\partial G}{\partial \mathbf{n}_y}(x,y)\,u(y)\,d\sigma(y) \tag{43.31}$$

for all $u \in C^2(\Omega^o) \cap C^1(\Omega)$.

Throughout the rest of this subsection we will make the following assumption.

Assumption 2 (Solvability of Dirichlet Problem) *We assume that for each $g \in C(\partial\Omega)$ there exists $h = h_g \in C^2(\Omega^o) \cap C^1(\Omega)$ such that*

$$\Delta h = 0 \text{ on } \Omega \text{ with } h = g \text{ on } \partial\Omega.$$

In this case we define $G(x,y)$ as in Eq. (43.30). We will (almost) verify that this assumption holds in Section 43.5 below. The full verification will come later when we study Hilbert space methods.

Theorem 43.28. *Let $G(x,y)$ be given as in Eq (43.30). Then*

1. *$G(x,y)$ is smooth on $(\Omega^o \times \Omega^o)\backslash\triangle$ where $\triangle = \{(x,x) : x \in \Omega^o\}$.*
2. *$G(x,y) = G(y,x)$ for all $x, y \in \Omega$. In particular the function $h(x,y) := h_x(y)$ is symmetric in x, y and $x \in \Omega^o \to h_x \in C(\Omega)$ is a smooth mapping.*
3. *If Ω is connected, then $G(x,y) > 0$ for all $(x,y) \in (\Omega^o \times \Omega^o)\backslash\triangle$.*

Proof. Let $\epsilon > 0$ be small and $\Omega_\epsilon := \Omega \setminus (B(x,\epsilon) \cup B(z,\epsilon))$ as in Figure 43.5, then by Green's theorem and the fact that $\Delta_y G(x,y) = 0$ if $y \neq x$,

$$\begin{aligned} 0 &= \int_{\Omega_\epsilon} \triangle_y G(x,y) G(z,y) dy \\ &= \int_{\partial\Omega_\epsilon} \left(\frac{\partial}{\partial \mathbf{n}} G(x,y) G(z,y) - G(x,y) \frac{\partial G}{\partial \mathbf{n}}(z,y)\right) d\sigma \\ &+ \int_{\Omega_\epsilon} G(x,y) \triangle_y G(z,y) dy \\ &= \int_{\partial\Omega_\epsilon} \left(\frac{\partial}{\partial \mathbf{n}} G(x,y) G(z,y) - G(x,y) \frac{\partial G}{\partial \mathbf{n}}(z,y)\right) d\sigma \end{aligned}$$

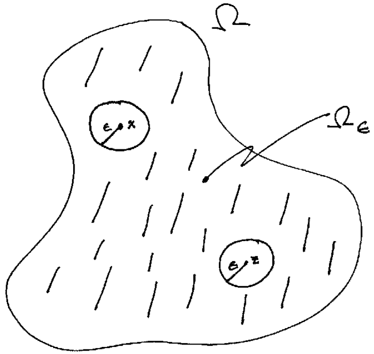


Fig. 43.5. Excising the singular region from Ω.

Since $G(x,y)$ and $G(z,y)=0$ for $y\in\partial\Omega$, the previous equation implies,

$$\int_{\partial(B(x,\epsilon)\cup B(z,\epsilon)).}\left\{\frac{\partial}{\partial \mathbf{n}_y}G(x,y)G(z,y)-G(z,y)\frac{\partial}{\partial \mathbf{n}_y}G(z,y)\right\}d\sigma=0.$$

We now let $\epsilon\downarrow 0$ in the above equations to find

$$\lim_{\epsilon\downarrow 0}\int_{\partial(B(x,\epsilon))}\frac{\partial\phi(x-y)}{\partial \mathbf{n}_y}G(z,y)d\sigma(y)=\lim_{\epsilon\downarrow 0}\int_{\partial B(z,\epsilon)}G(x,y)\frac{\partial}{\partial \mathbf{n}_y}\phi(z-y)\,d\sigma. \tag{43.32}$$

Moreover as we have seen above,

$$\lim_{\epsilon\downarrow 0}\int_{\partial B(z,\epsilon)}G(x,y)\frac{\partial}{\partial \mathbf{n}_y}\phi(z-y)\,d\sigma=G(x,z)\text{ and}$$
$$\lim_{\epsilon\downarrow 0}\int_{\partial(B(x,\epsilon))}\frac{\partial\phi(x-y)}{\partial \mathbf{n}_y}G(z,y)d\sigma(y)=G(z,x)$$

and hence $G(x,z)=G(z,x)$. Since $G(x,y)=\phi(x-y)-h_x(y)$ and $\phi(x-y)=\phi(y-x)$ it follows that $h_x(y)=h_y(x)=:h(x,y)$. Therefore $y\to h_x(y)$ and $x\to h_x(y)$ are smooth functions. Now by the maximum principle (Theorem 43.23):

$$\begin{aligned}|h_x(y)-h_z(y)|&\le\max_{y\in\partial\Omega}|h_x(y)-h_z(y)|\\&=\max_{y\in\partial\Omega}|\phi(x-y)-\phi(z-y)|\to 0\text{ as }x\to z.\end{aligned}$$

Therefore the map $x\in\Omega\to h_x\in C(\Omega)$ is continuous and in particular the map $(x,y)\to h(x,y)$ is jointly continuous. Letting η be as in the proof of Proposition 43.14, we find

$$\begin{aligned}h(x,y)&=\int_\Omega h(\tilde{x},y)\eta(x-\tilde{x})d\tilde{x}\\&=\int_{\Omega\times\Omega}h(\tilde{x},\tilde{y})\;\eta(y-\tilde{y})\;\eta(x-\tilde{x})\;d\tilde{x}\;d\tilde{y}\end{aligned}$$

from which it follows that in fact h is smooth on $\Omega\times\Omega$.

It only remains to show $x\to h_x\in C(\Omega)$ is smooth as well. Fix $x\in\Omega$ and for $v\in\mathbb{R}^n$, let $H_v\in C^2(\Omega^o)\cap C^1(\Omega)$ denote the solution to

$$\Delta H_v=0\text{ on }\Omega\text{ with }H_v(y)=v\cdot\nabla\phi(x-y)\text{ for }y\in\partial\Omega.$$

Notice that $v\to H_v$ is linear and by the maximum principle,

$$\begin{aligned}\|h_{x+v}-h_x-H_v\|_{L^\infty(\Omega)}&\le\|h_{x+v}-h_x-H_v\|_{L^\infty(\partial\Omega)}\\&=\|\phi(x+v-\cdot)-\phi(x-\cdot)-v\cdot\nabla\phi(x-\cdot)\|_{L^\infty(\partial\Omega)}.\end{aligned}$$

Now,

$$\begin{aligned}&\phi(x+v-y)-\phi(x-y)-v\cdot\nabla\phi(x-y)\\&\qquad=\int_0^1[\nabla\phi(x+tv-y)-\nabla\phi(x-y)]\cdot v dt\end{aligned}$$

so that, by the dominated convergence theorem,

$$\begin{aligned}&||\phi(x+v-\cdot)-\phi(x-\cdot)-v\cdot\nabla\phi(x-\cdot)||_{L^\infty(\partial\Omega)}\\&\qquad\le|v|\int_0^1\|\nabla\phi(x+tv-\cdot)-\nabla\phi(x-\cdot)\|_{L^\infty(\partial\Omega)}\,dt=o\left(|v|\right).\end{aligned}$$

This proves $x\to h_x$ is differentiable and that $\partial_v h_x=H_v$. Similarly one shows that $x\to h_x$ has higher derivatives as well.

For the last item, let $x\in\Omega^o$ and choose $\epsilon>0$ sufficiently small so that $\overline{B(x,\epsilon)}\subset\Omega^o\setminus\{y\}$ and $G(x,z)>0$ for all $z\in\overline{B(x,\epsilon)}$. Then the function $u(y):=G(x,y)$ is Harmonic on $\Omega^0\setminus\overline{B(x,\epsilon)}$, $u\in C(\Omega\setminus B(x,\epsilon))$, $u=0$ on $\partial\Omega$ and $u>0$ on $\partial\overline{B(x,\epsilon)}$. Hence by the maximum principle, $0\le u$ on $\Omega\setminus B(x,\epsilon)$ and since u is not constant we must also have $u>0$ on $\Omega^0\setminus\overline{B(x,\epsilon)}$. Since $\epsilon>0$ was any sufficiently small number, it follows $G(x,y)>0$ for all $y\in\Omega^o\setminus\{x\}$. ■

Corollary 43.29. *Keeping the above hypothesis and assuming $\rho\in C^2(\Omega^o)\cap L^1(\Omega)$ and $g\in C(\partial\Omega)$, then there is (a necessarily unique) solution $u\in C^2(\Omega^o)\cap C(\Omega)$ to*

$$\Delta u=-\rho\text{ with }u=g\text{ on }\partial\Omega \tag{43.33}$$

which is given by Eq. (43.31).

Proof. According to the remarks just before Eq. (43.31), if a solution to Eq. (43.33) exists it must be given by

$$u(x) = \int_\Omega G(x,y)\rho(y)\,dy - \int_{\partial\Omega} \frac{\partial G}{\partial \mathbf{n}_y}(x,y)\,g(y)\,d\sigma(y). \tag{43.34}$$

From Assumption 2, there exists a solution $v \in C^2(\Omega) \cap C^1(\bar{\Omega})$ such that $\Delta v = 0$ and $v = g$ on $\partial\Omega$. So replacing u by $u - v$ if necessary, it suffices to prove there is a solution $u \in C^2(\Omega^o) \cap C(\Omega)$ such that Eq. (43.33) holds with $g \equiv 0$. To produce this solution, let

$$u(x) := \int_\Omega G(x,y)\rho(y)\,dy = \int_\Omega \phi(x-y)\rho(y)\,dy - H(x)$$

where

$$H(x) := \int_\Omega h(x,y)\rho(y)dy.$$

Using the result in Theorem 43.28, one easily shows $H \in C^\infty(\Omega^o) \cap C^1(\Omega)$and $\Delta H = 0$. By Theorem 43.9,

$$\Delta_x \int_\Omega \phi(x-y)\rho(y)\,dy = -\rho(x) \text{ for } x \in \Omega$$

and therefore $u \in C^2(\Omega)$ and $\Delta u = -\rho$. ■

Remark 43.30. Because of the maximum principle, for any $x \in \Omega$ the map $g \in C(\partial\Omega) \to h_g(x) \in C(\bar{\Omega})$ is a positive linear functional. So by the Riesz representation theorem, there exists a unique positive probability measure σ_x on $\partial\Omega$ such that

$$h_g(x) = \int_{\partial\Omega} g(y)d\sigma_x(y) \text{ for all } g \in C(\partial\Omega).$$

Evidently this measure is given by

$$d\sigma_x(y) = -\frac{\partial G}{\partial \mathbf{n}_y}(x,y)d\sigma(y)$$

and in particular $-\frac{\partial G}{\partial \mathbf{n}_y}(x,y) \geq 0$ for all $x \in \Omega$ and $y \in \partial\Omega$. It is in fact easy to see that $-\frac{\partial G}{\partial \mathbf{n}_y}(x,y) > 0$ for all $x \in \Omega$ and $y \in \partial\Omega$.

43.3 Explicit Green's Functions and Poisson Kernels

In this section we will use the *method of images* to construct explicit formula for the Green's functions and Poisson Kernels for the half plane[1], $\bar{\mathbb{H}}^n = \{x \in$

[1] We will do this again later using the Fourier transform.

$\mathbb{R}^n : x_n \geq 0\}$and Balls $\overline{B(0,a)}$. For $x = (x', z) \in \mathbb{R}^{n-1} \times (0,\infty) = \mathbb{H}^n$ let $Rx := (x', -z)$. It is simple to verify $|x-y| = |Rx - y|$ for all $x \in \mathbb{H}^n$ and $y \in \partial\bar{\mathbb{H}}^n$. Form this and the properties of ϕ, one concluded, for $x \in \mathbb{H}^n$, that $h_x(y) := \phi(y - Rx)$ is Harmonic in $y \in \mathbb{H}^n$ and $h_x(y) = \phi(x-y)$ for all $y \in \partial\bar{\mathbb{H}}^n$. These remarks give rise to the following theorem.

Theorem 43.31. *For $x, y \in \bar{\mathbb{H}}^n$, let*

$$G(x,y) := \phi(y-x) - \phi(y - Rx) = \phi(y-x) - \phi(Ry - x).$$

Then G is the Greens function for Δ on $\mathbb{H}^n$ and

$$K(x,y) := -\frac{\partial G}{\partial \mathbf{n}}(x,y) = \frac{2x_n}{\sigma(S^{n-1})}\frac{1}{|x-y|^n} \text{ for } x \in \mathbb{H}^n \text{ and } y \in \partial\bar{\mathbb{H}}^n$$

is the Poisson kernel for $\mathbb{H}^n$. Furthermore if $\rho \in C^2(\mathbb{H}^n) \cap L^1(\mathbb{H}^n)$ and $f \in BC\left(\partial\bar{\mathbb{H}}^n\right)$, then

$$u(x) = \int_{\mathbb{H}^n} G(x,y)\rho(y)dy + \int_{\partial\bar{\mathbb{H}}^n} K(x,y)f(y)d\sigma(y)$$

solves the equation

$$\Delta u = -\rho \text{ on } \mathbb{H}^n \text{ with } u = f \text{ on } \partial\bar{\mathbb{H}}^n.$$

Proof. First notice that

$$G(y,x) = \phi(x-y) - \phi(x - Ry) = \phi(x-y) - \phi(Rx - RRy) = G(x,y)$$

since ϕ is a function of $|\cdot|$. Therefore, if

$$u(x) = \int_{\mathbb{H}^n} G(x,y)\rho(y)dy = \int_{\mathbb{H}^n} \phi(x-y)\rho(y)dy - \int_{\mathbb{H}^n} \phi(x - Ry)\rho(y)dy,$$

we have from Theorem 43.9 that

$$\Delta u(x) = -\rho(x) - \int_{\mathbb{H}^n} \Delta_x\phi(x - Ry)\rho(y)dy = -\rho(x).$$

Since $G(x,y) = 0$ for $x \in \partial\bar{\mathbb{H}}^n$ and so $u(x) = 0$ for $x \in \partial\bar{\mathbb{H}}^n$. It is left to the reader to show u is continuous on $\overline{\mathbb{H}^n}$.

For $x \in \mathbb{H}^n$ and $y \in \partial\bar{\mathbb{H}}^n$, we find form Eq. (43.7),

$$\begin{aligned}
K(x,y) &:= -\frac{\partial G}{\partial \mathbf{n}_y}(x,y) = \frac{\partial}{\partial y_n}G(x,y) \\
&= \frac{\partial}{\partial y_n}\left[\phi(y-x) - \phi(y - Rx)\right] \\
&= -\frac{1}{\sigma(S^{n-1})}\frac{1}{|y-x|^n}(y-x)\cdot e_n \\
&\quad + \frac{1}{\sigma(S^{n-1})}\frac{1}{|y-Rx|^n}(y-Rx)\cdot e_n \\
&= \frac{1}{\sigma(S^{n-1})}\frac{2x_n}{|y-x|^n}.
\end{aligned}$$

Claim: For all $x \in \mathbb{H}^n$,

$$\int_{\partial \bar{\mathbb{H}}^n} K(x,y)dy = 1.$$

It is possible to prove this by direct computation, since (writing $x = (x', x_n)$ as above)

$$\begin{aligned}
\int_{\partial \bar{\mathbb{H}}^n} K(x,y)dy &= \frac{2}{\sigma(S^{n-1})} \int_{\mathbb{R}^{n-1}} \frac{x_n}{\left(|x'-y|^2 + x_n^2\right)^{n/2}} dy \\
&= \frac{2}{\sigma(S^{n-1})} \int_{\mathbb{R}^{n-1}} \frac{1}{\left(|y|^2 + 1\right)^{n/2}} dy \\
&= \frac{2}{\sigma(S^{n-1})} \sigma(S^{n-2}) \int_0^\infty r^{n-2} \frac{1}{\left(r^2+1\right)^{n/2}} dr
\end{aligned}$$

where in the second equality we have made the change of variables $y \to x_n y$ and in the last we passed to polar coordinates. When $n = 2$ we find

$$\int_0^\infty r^{n-2} \frac{1}{\left(r^2+1\right)^{n/2}} dr = \int_0^\infty \frac{1}{r^2+1} dr = \pi/2$$

and for $n = 3$ we may let $u = r^2$ to find

$$\int_0^\infty r^{n-2} \frac{1}{\left(r^2+1\right)^{n/2}} dr = \int_0^\infty r \frac{1}{\left(r^2+1\right)^{3/2}} dr = \frac{1}{2} \int_0^\infty \frac{1}{\left(u+1\right)^{3/2}} du = 1.$$

These results along with

$$\begin{aligned}
\int_0^\infty r^{n-2} \frac{1}{\left(r^2+1\right)^{n/2}} dr &= \int_0^\infty \left(r^2+1\right)^{-n/2} d\frac{r^{n-1}}{n-1} \\
&= \frac{n/2}{n-1} \int_0^\infty \left(r^2+1\right)^{-n/2-1} 2r\, r^{n-1} dr \\
&= \frac{n}{n-1} \int_0^\infty r^n \frac{1}{\left(r^2+1\right)^{\frac{n+2}{2}}} dr
\end{aligned}$$

allows one to compute $\int_0^\infty r^{n-2} \frac{1}{(r^2+1)^{n/2}} dr$ inductively. I will not carry out the details of this method here. Rather, it is more instructive to use Corollary 43.6 to prove the claim. In order to do this let $u \in C_c^\infty(B(0,1),[0,1])$ such that $u(0) = 1$, $u(x) = U(|x|)$ and $U(r)$ is decreasing as r decreases. Then by Corollary 43.6, with $u(x) = u_M(x) := u(x/M)$,

$$u_M(x) = \int_{\partial \bar{\mathbb{H}}^n} K(x,y) u(y/M) d\sigma(y) - M^{-2} \int_{\mathbb{H}^n} G(x,y) \left(\triangle u\right)(y/M) dy. \quad (43.35)$$

By the monotone convergence theorem,

$$\lim_{M \uparrow \infty} \int_{\partial \bar{\mathbb{H}}^n} K(x,y) u(y/M) d\sigma(y) = \int_{\partial \bar{\mathbb{H}}^n} K(x,y) d\sigma(y)$$

and therefore passing the limit in Eq. (43.35) gives

$$1 = \int_{\partial \bar{\mathbb{H}}^n} K(x,y) d\sigma(y) - \lim_{M \uparrow \infty} \left[M^{-2} \int_{\mathbb{H}^n} G(x,y) \triangle u(y/M) dy \right].$$

This latter limit is zero, since

$$\begin{aligned}
M^{-2} \int_{\mathbb{H}^n} &G(x,y) \triangle u(y/M) dy \\
&= c_n M^{-2} \int_{\mathbb{H}^n} \left[\frac{1}{|x-y|^{n-2}} - \frac{1}{|Rx-y|^{n-2}} \right] \left(\triangle u\right)(y/M) dy \\
&= c_n M^{-2} M^n \int_{\mathbb{H}^n} \left[\frac{1}{|x-My|^{n-2}} - \frac{1}{|Rx-My|^{n-2}} \right] \triangle u(y) dy \\
&= c_n \int_{\mathbb{H}^n} \left[\frac{1}{|x/M-y|^{n-2}} - \frac{1}{|Rx/M-y|^{n-2}} \right] \triangle u(y) dy.
\end{aligned}$$

This latter expression tends to zero and $M \to \infty$ by the dominated convergence and this proves the claim. (Alternatively, for y large,

$$\begin{aligned}
\frac{1}{|x-y|^{n-2}} &- \frac{1}{|Rx-y|^{n-2}} \\
&= \frac{1}{|y|^{n-2}} \left[\frac{1}{\left|\frac{x}{|y|} - \hat{y}\right|^{n-2}} - \frac{1}{\left|R\frac{x}{|y|} - \hat{y}\right|^{n-2}} \right] \\
&= \frac{1}{|y|^{n-2}} \left[\left(1 + 2\frac{x}{|y|} \cdot \hat{y} + \dots\right) - \left(1 + 2\frac{Rx}{|y|} \cdot \hat{y} + \dots\right) \right] \\
&= O\left(\frac{1}{|y|^{n-1}}\right)
\end{aligned}$$

and therefore

$$\begin{aligned}
M^{-2} \int_{\mathbb{H}^n} &\left[\frac{1}{|x-y|^{n-2}} - \frac{1}{|Rx-y|^{n-2}} \right] \left(\triangle u\right)(y/M) dy \\
&= O\left(M^{-2} \frac{1}{M^{n-1}} M^n \right) = O\left(1/M\right) \to 0
\end{aligned}$$

as $M \to \infty$.

Since $G(x,y)$ is harmonic in x, it follows that $K(x,y) = -\frac{\partial}{\partial \mathbf{n}_y} G(x,y)$ is still Harmonic in x. and therefore

$$u(x) := \int_{\partial \bar{\mathbb{H}}^n} K(x,y) f(y) d\sigma(y) = \frac{2}{\sigma(S^{n-1})} \int_{\partial \bar{\mathbb{H}}^n} \frac{x_n}{|x-y|^n} f(y) d\sigma(y)$$

is harmonic as well. Since

$$\begin{aligned} u(x) &= \frac{2}{\sigma(S^{n-1})} \int_{\partial \bar{\mathbb{H}}^n} \frac{x_n}{(|x'-y|^2 + x_n^2)^{n/2}} f(y) dy \qquad (43.36) \\ &= \frac{2}{\sigma(S^{n-1})} \frac{1}{x_n^{n-1}} \int_{\partial \bar{\mathbb{H}}^n} \frac{1}{\left(|\frac{x'-y}{x_n}|^2 + 1\right)^{n/2}} f(y) dy \end{aligned}$$

it follows from Theorem 42.13 that $u((x', x_n)) \to f(x')$ as $x_n \downarrow 0$ uniformly for x' in compact subsets of $\partial \bar{\mathbb{H}}^n$. ■

43.4 Green's function for Ball

Let $r > 0$ be fixed, we will construct the Green's function for the ball $B(0,r)$. The idea for a given $x \in B(0,r)$, we should find a mirror location, say $\rho \hat{x}$ and a charge q so that

$$\phi(x-y) = q\phi(\rho\hat{x} - y) \text{ for all } |y| = r.$$

Assuming for the moment that $n \geq 3$ and writing $q = \beta^{(2-n)}$, this leads to the equations

$$|x-y|^2 = |\beta\rho\hat{x} - \beta y|^2 = \beta^2 |\rho\hat{x} - y|^2$$

or equivalently squaring out both sides and using $|y| = r$,

$$|x|^2 - 2x \cdot y + r^2 = \beta^2 \left(\rho^2 - 2\rho\hat{x} \cdot y + r^2\right).$$

Choosing $y \perp x$ and $y = r\hat{x}$ leads to the conditions

$$\begin{aligned} |x|^2 + r^2 &= \beta^2 \left(\rho^2 + r^2\right) \text{ and} \\ |x|^2 - 2r|x| + r^2 &= \beta^2 \left(\rho^2 - 2\rho r + r^2\right). \end{aligned}$$

Subtracting these two equations implies $-2r|x| = -2\rho\beta^2 r$ or equivalently that $\rho = |x|/\beta^2$. Putting this into the first equation above then implies

$$|x|^2 + r^2 = \frac{|x|^2}{\beta^2} + \beta^2 r^2$$

or equivalently that

$$0 = r^2\beta^4 - \left(|x|^2 + r^2\right)\beta^2 + |x|^2.$$

By the quadratic formula, this implies

$$\begin{aligned} \beta^2 &= \frac{\left(|x|^2 + r^2\right) \pm \sqrt{\left(|x|^2 + r^2\right)^2 - 4r^2|x|^2}}{2r^2} \\ &= \frac{\left(|x|^2 + r^2\right) \pm \sqrt{\left(|x|^2 - r^2\right)^2}}{2r^2} = \frac{\left(|x|^2 + r^2\right) \pm \left(r^2 - |x|^2\right)}{2r^2} \\ &= 1 \text{ or } \frac{|x|^2}{r^2}. \end{aligned}$$

Clearly the charge $\beta = 1$ will not work so we must take $\beta = |x|/r$ in which case, $\rho = r^2/|x|$ and hence

$$q\phi(\rho\hat{x} - y) = (|x|/r)^{(2-n)} \phi\left(r^2 \frac{\hat{x}}{|x|} - y\right) = \phi\left(r\hat{x} - \frac{|x|}{r} y\right).$$

Let us now verify that our guess has worked. Let us begin by noting the following identities for $x, y \in \mathbb{R}^n$,

$$\left|r\hat{x} - r^{-1}|x|y\right|^2 = \left(r^2 - 2x \cdot y + r^{-2}|x|^2|y|^2\right) \qquad (43.37)$$

and in particular when $|y| = r$ this implies

$$|\hat{x}r - |x|\hat{y}|^2 = \left(r^2 - 2x \cdot y + |x|^2\right) = |x-y|^2$$

so that

$$|x-y| = |\hat{x}r - |x|\hat{y}| = \left|\frac{x}{|x|} r - |x| \frac{y}{r}\right|. \qquad (43.38)$$

Now the function

$$h_x(y) = \phi\left(r\hat{x} - \frac{|x|}{r} y\right) = \phi\left(\frac{|x|}{r}\left(y - r^2 \frac{x}{|x|^2}\right)\right)$$

is harmonic in y and by Eq. (43.38),

$$h_x(y) = \phi\left(\hat{x}r - |x| \frac{y}{r}\right) = \phi\left(\hat{x}r - |x|\,\hat{y}\right) = \phi(x-y) \text{ when } |y| = r.$$

Hence we should define the Green's function for the ball to be given by

$$\begin{aligned} G(x,y) &= \phi(x-y) - h_x(y) = \phi(x-y) - \phi\left(\hat{x}r - |x|\frac{y}{r}\right) \\ &= \phi(x-y) - \phi\left(\hat{x}r - r^{-1}|x|\,|y|\,\hat{y}\right) \\ &= \phi(x-y) - \phi\left(\frac{|x|}{r}\left(y - r^2 \frac{\hat{x}}{|x|}\right)\right). \end{aligned}$$

From Eq. (43.37), it follows that $h_x(y) = h_y(x)$ and therefore $G(x,y)$ is again symmetric under the interchange of x and y.

For $y \in \partial\overline{B(0,r)}$, using Eq. (43.7) we find

$$\begin{aligned}
-K(x,y) &= \frac{\partial G}{\partial \mathbf{n}_y}(x,y) = \partial_{\hat{y}} G(x,y) = \nabla_y G(x,y)\cdot\hat{y} \\
&= \nabla_y \left[\phi(x-y) - \phi\left(\frac{|x|}{r}\left(y - r^2\frac{\hat{x}}{|x|}\right)\right)\right]\cdot\hat{y} \\
&= -\frac{1}{\sigma(S^{n-1})}\left[\frac{1}{|y-x|^n}(y-x) - \frac{|x|}{r}\frac{1}{\left(\frac{|x|}{r}\right)^n\left|y-r^2\frac{\hat{x}}{|x|}\right|^n}\frac{|x|}{r}\left(y-r^2\frac{\hat{x}}{|x|}\right)\right]\cdot\hat{y} \\
&= -\frac{1}{\sigma(S^{n-1})}\left[\frac{1}{|y-x|^n}(y-x) - \frac{|x|}{r}\frac{1}{\left|\frac{|x|}{r}y - r\hat{x}\right|^n}\left(\frac{|x|}{r}y - r\hat{x}\right)\right]\cdot\hat{y} \\
&= -\frac{1}{\sigma(S^{n-1})}\left[\frac{1}{|y-x|^n}(y-x) - \frac{|x|}{r}\frac{1}{|x-y|^n}(|x|\hat{y} - r\hat{x})\right]\cdot\hat{y} \\
&= -\frac{1}{\sigma(S^{n-1})\,|y-x|^n}\left[(y-x) - \left(\frac{|x|^2}{r}\hat{y} - x\right)\right]\cdot\hat{y} \\
&= -\frac{1}{\sigma(S^{n-1})\,|y-x|^n}\left[y - \frac{|x|^2}{r}\hat{y}\right]\cdot\hat{y} \\
&= -\frac{1}{\sigma(S^{n-1})r\,|y-x|^n}\left[r^2 - |x|^2\right].
\end{aligned}$$

These computations lead to the following theorem.

Theorem 43.32. *For $x, y \in B(0,r)$, let*

$$G(x,y) := \phi(x-y) - \phi\left(\hat{x}r - |x|\frac{y}{r}\right)$$

and if $y \in \partial\overline{B(0,r)}$, let

$$K(x,y) := -\frac{\partial G}{\partial \mathbf{n}}(x,y) = \frac{r^2 - |x|^2}{\sigma(S^{n-1})r}\,|x-y|^{-n}.$$

Then $\rho \in C^2(B(0,r)) \cap L^1(B(0,r))$ and $f \in C\left(\partial\overline{B(0,r)}\right)$, then

$$u(x) = \int_{B(0,r)} G(x,y)\rho(y)dy + \int_{\partial\overline{B(0,r)}} K(x,y)f(y)d\sigma(y) \tag{43.39}$$

solves the equation

$$\Delta u = -\rho \text{ on } B(0,r) \text{ with } u = f \text{ on } \partial\overline{B(0,r)}.$$

Remark 43.33. Letting $\rho = |x|$, we may write $K(x,y)$ as

$$K(x,y) = \frac{r^2-\rho^2}{\sigma(S^{n-1})r}\frac{1}{(\rho^2 + r^2 - 2\rho r\hat{x}\cdot\hat{y})^{n/2}}. \tag{43.40}$$

In particular when $r = 1$, $n = 2$, $x = \rho e^{i\theta}$ and $y = re^{i\alpha}$, this gives

$$K(x,y) = \frac{1-\rho^2}{2\pi\left(\rho^2 + 1 - 2\rho\cos(\theta-\alpha)\right)}$$

which agrees with the Poisson kernel $P_\rho(\theta-\alpha)$ of Eq. (22.40) which we derived earlier by Fourier series methods.

Proof. The proof is essentially the same as Theorem 43.31 but a bit easier. For these reasons, we will only prove here the assertion

$$\lim_{x\to x_0}\int_{\partial\overline{B(0,r)}} K(x,y)f(y)d\sigma(y) = f(x_0) \text{ for all } x_0 \in \partial B(0,r). \tag{43.41}$$

From Theorem 43.5 with $u = 1$ it follows again that

$$\int_{\partial\overline{B(0,r)}} K(x,y)d\sigma(y) = 1.$$

For any $\alpha \in (0,1)$ let

$$\epsilon(\alpha) := \sup\left\{|f(ry) - f(rx)| : y, x \in S^{n-1} \text{ with } \hat{y}\cdot\hat{x} \le \alpha\right\}.$$

Then by uniform continuity of f on $\partial B(0,r)$, it follows that $\epsilon(\alpha) \to 0$ as $\alpha \uparrow 1$ and hence

$$\begin{aligned}
&|\int_{\partial\overline{B(0,r)}} K(x,y)f(y)d\sigma(y) - f(r\hat{x})| \\
&\le \int_{\partial\overline{B(0,r)}} K(x,y)\left|f(y) - f(r\hat{x})\right|d\sigma(y) \\
&= \int_{\partial\overline{B(0,r)}} 1_{\hat{y}\cdot\hat{x}>\alpha}K(x,y)\left|f(y) - f(r\hat{x})\right|d\sigma(y) \\
&+ \int_{\partial\overline{B(0,r)}} 1_{\hat{y}\cdot\hat{x}\le\alpha}K(x,y)\left|f(y) - f(r\hat{x})\right|d\sigma(y) \\
&\le 2\left\|f\right\|_u \int_{\partial\overline{B(0,r)}} 1_{\hat{y}\cdot\hat{x}>\alpha}K(x,y)d\sigma(y) + \epsilon(\alpha) \\
&\le C(\alpha)\left\|f\right\|_u\left(r^2 - |x|^2\right) + \epsilon(\alpha)
\end{aligned}$$

where $C(\alpha)$ is some constant only depending on α, see Eq. (43.40). Therefore,

$$\sup_{|x|=\rho}\left|\int_{\partial\overline{B(0,r)}} K(x,y)f(y)d\sigma(y) - f(r\hat{x})\right| \le C(\alpha)\left\|f\right\|_u(r^2-\rho^2) + \epsilon(\alpha)$$

and hence

$$\limsup_{\rho\uparrow r}\sup_{|x|=\rho}\left|\int_{\partial B(0,r)}K(x,y)f(y)d\sigma(y)-f(r\hat{x})\right|\le\epsilon(\alpha)\to 0\text{ as }\alpha\downarrow 0$$

which implies Eq. (43.41). ■

Corollary 43.34. *Suppose that u is a harmonic function on Ω, then u is real analytic on Ω.*

Proof. The condition of being real analytic is local and invariant under translations as is the notion of being harmonic. Hence we may assume $0\in\overline{B(0,r)}\subset\Omega$ for some $r>0$, in which case we have, for $|x|<r$ and $f=u|_{\partial B(0,r)}$, that

$$\begin{aligned}u(x)&=\int_{\partial B(0,r)}K(x,y)f(y)d\sigma(y)=\frac{r^2-|x|^2}{\sigma(S^{n-1})r}\int_{\partial B(0,r)}|x-y|^{-n}f(y)d\sigma(y)\\&=\frac{r^2-|x|^2}{\sigma(S^{n-1})r}\int_{\partial B(0,r)}|x-\hat{y}r|^{-n}f(y)d\sigma(y).\end{aligned}\tag{43.42}$$

Now

$$\begin{aligned}|x-y|^{-n}&=|x-\hat{y}r|^{-n}=r^{-n}\left|\hat{y}-r^{-1}x\right|^{-n}\\&=r^{-n}\left(1-2r^{-1}\hat{y}\cdot x+\frac{|x|^2}{r^2}\right)^{-n/2}\\&=:r^{-n}\left(1-\alpha(x,y)\right)^{-n/2}\end{aligned}$$

where

$$\alpha(x,y):=2r^{-1}\hat{y}\cdot x-\frac{|x|^2}{r^2}.$$

Since

$$|\alpha(x,y)|\le 2r^{-1}|x|+\frac{|x|^2}{r^2}\le 2\alpha_0+\alpha_0^2<1$$

if $|x|\le\alpha_0 r$ and $\alpha_0<\sqrt{2}-1$, we find that $|x-y|^{-n}$ has a convergent power series expansion,

$$|x-y|^{-n}=r^{-n}\sum_{m=0}^{\infty}a_m\alpha(x,y)^m\text{ for }|x|\le\alpha_0 r.$$

Plugging this into Eq. (43.42) shows $u(x)$ has a convergent power series expansion in x for $|x|\le\left(\sqrt{2}-1\right)r$. ■

43.5 Perron's Method for solving the Dirichlet Problem

For this section let $\Omega\subset_o\mathbb{R}^n$ be a bounded open set and $f\in C(\mathrm{bd}(\Omega),\mathbb{R})$ be a given function. We are going to investigate the solvability of the Dirichlet problem:

$$\Delta u=0\text{ on }\Omega\text{ with }u=f\text{ on }\mathrm{bd}(\Omega).\tag{43.43}$$

Let $\mathcal{S}(\Omega)$ denote those $w\in C\left(\bar{\Omega}\right)$ such that w is subharmonic on Ω and let $\mathcal{S}_f(\Omega)$ denote those $w\in\mathcal{S}(\Omega)$ such that $w\le f$ on $\mathrm{bd}(\Omega)$. As we have seen in Corollary 43.26, if there is a solution to $u\in C^2(\Omega)\cap C\left(\bar{\Omega}\right)$, then $w\le u$ for all $w\in\mathcal{S}_f(\Omega)$. This suggests we try to define

$$u(x):=u_f(x):=\sup\{w(x):w\in\mathcal{S}_f(\Omega)\}\text{ for all }x\in\bar{\Omega}.\tag{43.44}$$

Notation 43.35 *Given $w\in\mathcal{S}(\Omega)$, $\xi\in\Omega$ and $r>0$ such that $\overline{B(\xi,r)}\subset\Omega$, let (see Figure 43.6)*

$$w_{\xi,r}(y)=\begin{cases}w(y)\text{ for }y\in\Omega\setminus B(\xi,r)\\h(y)\text{ for }\ y\in\overline{B(\xi,r)}\end{cases}$$

where $h\in C\left(\overline{B(\xi,r)}\right)$ is the unique solution to

$$\Delta h=0\text{ on }B(\xi,r)\text{ with }h=w\text{ on }\partial B(\xi,r).$$

The existence of h is guaranteed by Theorem 43.32.

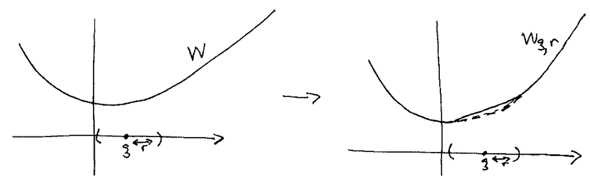


Fig. 43.6. The construction of $w_{\xi,r}$ in the one-dimensional case.

Proposition 43.36. *Let $w\in\mathcal{S}(\Omega)$ and $w_{\xi,r}$ be as above. Then*

1. *$w\le w_{\xi,r}$.*
2. *$w_{\xi,r}\in\mathcal{S}(\Omega)$, i.e. $w_{\xi,r}$ is subharmonic.*
3. *For **any** $\xi\in\Omega$ and $r>0$ such that $\overline{B(\xi,r)}\subset\Omega$, the mean value inequality is valid,*

$$w(\xi)\le\fint_{\partial B(\xi,r)}wd\sigma.$$

Proof. 1. Since $w = w_{\xi,r}$ on $\Omega \setminus B(\xi, r)$, it suffices to show $w \le h$ on $B(\xi, r)$. But this follows from Corollary 43.26.

2. Since $w_{\xi,r}$ is harmonic on $B(\xi, r)$ and subharmonic on $\Omega \setminus \overline{B(\xi, r)}$, we need only show

$$w_{\xi,r}(y) \le \fint_{\partial B(y,\rho)} w_{\xi,r} d\sigma$$

for all $y \in \partial B(\xi, r)$ and ρ sufficiently small. This is easily checked, since w is subharmonic,

$$w_{\xi,r}(y) = w(y) \le \fint_{\partial B(y,\rho)} w d\sigma \le \fint_{\partial B(y,\rho)} w_{\xi,r} d\sigma$$

wherein the last equality we made use of Item 1.

3. By item 1. and the mean value property for the harmonic function, $w_{\xi,r}$, we have

$$w(\xi) \le w_{\xi,r}(\xi) = \fint_{\partial B(\xi,r)} w_{\xi,r} d\sigma = \fint_{\partial B(\xi,r)} w d\sigma.$$

■

Theorem 43.37. *The function $u = u_f$ defined in Eq. (43.44) is harmonic on Ω and $u \le f$ on $\mathrm{bd}(\Omega)$.*

Proof. Let us begin with a couple of observations. In what follows

$$m := \min\{f(x) : x \in \mathrm{bd}(\Omega)\} \text{ and } M := \min\{f(x) : x \in \mathrm{bd}(\Omega)\}.$$

1. The function $u = u_f \ge m$ on Ω since $m \in \mathcal{S}_f(\Omega)$.
2. By the maximum principle $w \le M$ on Ω for all $w \in \mathcal{S}_f(\Omega)$ and therefore $u_f \le M$ on Ω.
3. If $w_1, \dots, w_m \in \mathcal{S}_f(\Omega)$, then $w = \max\{w_1, \dots, w_m\} \in \mathcal{S}_f(\Omega)$. Indeed for $\xi \in \Omega$ and r small,
$$\fint_{\partial B(\xi,r)} w d\sigma \ge \fint_{\partial B(\xi,r)} w_i d\sigma \ge w_i(\xi)$$
for all i.
4. Now suppose $\xi \in \Omega$ and $R > 0$ be chosen so that $\overline{B(\xi, R)} \subset \Omega$ and $D \subset B(\xi, R)$ is a countable set. Then there is a harmonic function w_D on $B(\xi, R)$ such that $w_D = u_f$ on D.

To prove this last item let $D := \{y_k\}_{k=1}^\infty$ and choose $\{w_k^m\} \subset \mathcal{S}_f(\Omega)$ such that $w_k^m(y_k) \to u(y_k)$ as $m \to \infty$ for each k. By replacing w_k^m by $\max\{w_k^1, \dots, w_k^m\}$ if necessary we may assume for each k that w_k^m is increasing in m for each k. Letting

$$W_m := \max\{w_1^m, \dots, w_m^m\}$$

we find an increasing sequence $\{W_m\} \subset \mathcal{S}_f(\Omega)$ such that $W_m(y) \uparrow u_f(y)$ for all $y \in D$. Finally define a sequence $\{w_m\} \subset \mathcal{S}_f(\Omega)$ by $w_m := (W_m)_{\xi,2R}$. By the maximum principle, w_m is still increasing and since $W_m \le w_m$ and we still have $w_m(y) \uparrow u_f(y)$ for all $y \in D$. We now define $w_D := \lim_{m\to\infty} w_m|_{B(\xi,R)}$ which exists because w_m is increasing and w. We have $w_D = u_f$ on D and because $\{w_m\}$ is a bounded and convergent sequence of harmonic functions on $B(\xi, R)$, it follows from Corollary 43.18 that w is harmonic on $B(\xi, R)$. This completes the proof of item 4.

We now use item 4. to prove u_f is continuous at $\xi \in \Omega$. To do this let $\{y_k\}_{k=1}^\infty \subset B(\xi, R)$ be any sequence such that $y_k \to \xi$ as $k \to \infty$ and let $D = \{\xi\} \cup \{y_k\}_{k=1}^\infty \subset B(\xi, R)$. Since w_D is harmonic and hence continuous,

$$\lim_{k\to\infty} u_f(y_k) = \lim_{k\to\infty} w_D(y_k) = w_D(\xi) = u_f(\xi)$$

showing u_f is continuous.

To show u_f is harmonic on $B(\xi, R)$, let D be a countable dense subset of $B(\xi, R)$. Then the continuity of u_f and the fact that $u_f = w_D$ on D, it follows that $u_f = w_D$ on $B(\xi, R)$. In particular u_f is harmonic on $B(\xi, R)$. Since ξ is arbitrary, we have shown u_f is harmonic. ■

To complete our program, we want to show that u_f extends to a function in $C(\bar{\Omega})$ and that $u_f = f$ on $\mathrm{bd}(\Omega)$. For this we will need some assumption on $\mathrm{bd}(\Omega)$.

Definition 43.38. *A function $Q \in C(\bar{\Omega})$ is a **barrier function** for $\eta \in \mathrm{bd}(\Omega)$ if Q is subharmonic on Ω, $Q(\eta) = 0$ and $Q(x) < 0$ for all $x \in \mathrm{bd}(\Omega) \setminus \{\eta\}$.*

Example 43.39. Suppose that $\eta \in \mathrm{bd}(\Omega)$ and there exists $\xi \in \mathbb{R}^n$ such that $(x - \eta)\cdot\xi < 0$ for all $x \in \mathrm{bd}(\Omega)\setminus\{\eta\}$ (see Figure 43.7 below), then the function $Q(x) := (x - \eta) \cdot \xi$ is a barrier function of η.

Example 43.40. Suppose that $\eta \in \mathrm{bd}(\Omega)$ and there exists a ball $\overline{B(\xi, r)} \cap \bar{\Omega} = \{\eta\}$ (see Figure 43.8), then $Q(x) := \alpha(r) - \alpha(|x - \xi|)$ is a barrier function for η, where α is defined in Eq. 43.4.

Theorem 43.41. *Suppose $f \in C(\mathrm{bd}(\Omega))$ and $u = u_f$ is the harmonic function defined by Eq. (43.44) and there exists a barrier function Q for $\eta \in \mathrm{bd}(\Omega)$. Then $\lim_{x\to\eta} u_f(x) = f(\eta)$. In particular if every point $\eta \in \mathrm{bd}(\Omega)$ admits a barrier function, then there is a unique solution $u \in C(\bar{\Omega}) \cap C^2(\Omega)$ to $\Delta u = 0$ with $u = f$ on $\mathrm{bd}(\Omega)$.*

Proof. Given $\epsilon > 0$ and $K > 0$, let $w(x) := f(\eta) - \epsilon - KQ(x)$ for all $x \in \bar{\Omega}$. For any $\epsilon > 0$ we may choose (using continuity of f and compactness

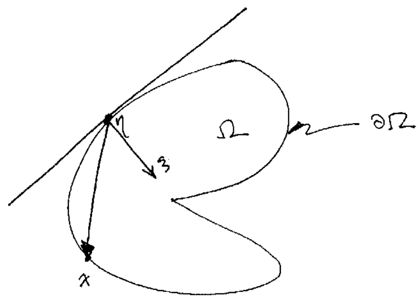


Fig. 43.7. Constructing a barrier function at point where η where $\partial\Omega$ lies in a half plane.

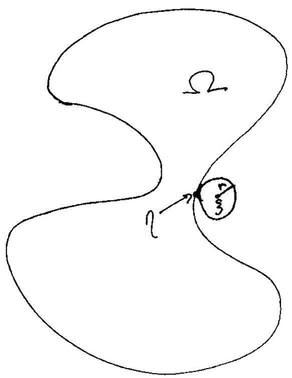


Fig. 43.8. Another η for which there exists a barrier function.

of $\mathrm{bd}(\Omega)$) K sufficiently large so that $w \le f$ on $\mathrm{bd}(\Omega)$, i.e. $w \in \mathcal{S}_f(\Omega)$. Therefore $w \le u_f$ and hence

$$f(\eta) - \epsilon = w(\eta) = \lim_{x\to\eta} w(x) \le \liminf_{x\to\eta} u_f(x).$$

Since $\epsilon > 0$ is arbitrary, this shows

$$\liminf_{x\to\eta} u_f(x) \ge f(\eta). \tag{43.45}$$

We now consider the function

$$\begin{aligned} -u_{-f}(x) &= -\sup\{w(x) : w \in \mathcal{S}_{-f}(\Omega)\} = \inf\{-w(x) : w \in \mathcal{S}_{-f}(\Omega)\} \\ &= \inf\{W(x) : -W \in \mathcal{S}_{-f}(\Omega)\}. \end{aligned} \tag{43.46}$$

If $w \in \mathcal{S}_f(\Omega)$ and $-W \in \mathcal{S}_{-f}(\Omega)$, then $w - W$ is sub harmonic and $w - W \le f - f = 0$ on $\mathrm{bd}(\Omega)$, therefore by the maximum principle it follows that $w \le W$ on $\bar{\Omega}$. Coupling this fact with Eq. (43.46) shows $-u_{-f}(x) \ge w(x)$ for all $w \in \mathcal{S}_f(\Omega)$ and then taking the supremum over w shows $u_f(x) \le -u_{-f}(x)$. Therefore using Eq. (43.45) with f replaced by $-f$ shows

$$\limsup_{x\to\eta} u_f(x) \le \limsup_{x\to\eta} (-u_{-f}(x)) = -\liminf_{x\to\eta} (u_{-f}(x)) \le -(-f(\eta)) = f(\eta) \tag{43.47}$$

which combined with Eq. (43.45) shows

$$\lim_{x\to\eta} u_f(x) = f(\eta).$$

■

Exercise 43.42. Suppose that R is an $n \times n$ orthogonal matrix ($R^{\mathrm{tr}} R = I = RR^{\mathrm{tr}}$) viewed as a linear transformation on $\mathbb{R}^n$. Show for $f \in C^2(\mathbb{R}^n)$ that $\Delta(f \circ R) = \Delta f \circ R$, i.e. Δ is invariant under rotations.

Exercise 43.43. Show that every point $\eta \in \mathrm{bd}(\Omega)$ has a barrier function when $\mathrm{bd}(\Omega)$ is C^2. **Hint:** By making a change of coordinated involving rotations and translations change of coordinates, it suffices to assume $\eta = 0 \in \mathrm{bd}(\Omega)$ and that $B(0,r) \cap \mathrm{bd}(\Omega)$ is the graph of a C^2 – function $g : B(0,r) \cap \mathbb{R}^{n-1} \to \mathbb{R}^n$ such that $g(0) = 0$ and $\nabla g(0) = 0$. Show for $\delta > 0$ sufficiently small that

$$d_\delta(x) := |\delta e_n - x|^2 \text{ for } x \in \mathrm{bd}(\Omega)$$

has a unique global minimum at $x = 0$. Use this fact and Example 43.40 to complete the proof.

Remark 43.44. To make Barrier functions for cones C, let $D := C \cap S^{n-1}$ and let $u(\omega)$ for $\omega \in D$ denote the Dirichlet eigenfunction on D for the spherical Laplacian with smallest eigenvalue $\lambda > 0$, i.e. $-\Delta_{S^{n-1}} u = \lambda u$. This function is positive on D^o and vanishes on the boundary. If C has sperical symmetry, the function should be describable explicitly. At any rate, we can now considier the function $U(r\omega) = r^\alpha u$, then

$$\begin{aligned} \Delta U &= \frac{1}{r^{n-1}} \partial_r \left(r^{n-1} \partial_r [r^\alpha u]\right) + \frac{r^\alpha}{r^2} \Delta_{S^{n-1}} u \\ &= \alpha(\alpha + n - 2) r^{\alpha-2} u - \lambda r^{\alpha-2} u \end{aligned}$$

which will be zero if $\lambda = \alpha(\alpha + n - 2)$, i.e.

$$\alpha^2 + (n-2)\alpha - \lambda = 0$$

where

$$\alpha = \frac{-(n-2) \pm \sqrt{(n-2)^2 + 4\lambda}}{2}.$$

Hence taking

$$\alpha = \frac{\sqrt{(n-2)^2 + 4\lambda} - (n-2)}{2},$$

the function $U(r\omega) = r^{\alpha}u(\omega)$ is harmonic in C and 0 on ∂C. (Probably should be doing these considerations for the exterior of C.)

43.6 Solving the Dirichlet Problem by Integral Equations

Another method for solving the Dirichlet problem to reduce it to a question of solvability of a certain integral equation in $\text{bd}(\Omega)$. For a nice sketch of how this goes the reader is referred to Reed and Simon [11], included below. For a more detailed account the reader may consult Sobolev [16] or Guenther and Lee [6].

The following text is taken from Reed and Simon Volume 1.

USES RESULTS FOR THE BANACH SPACE C

Example (Dirichlet problem) √ The main impetus for the study of compact operators arose from the use of integral equations in attempting to solve the classical boundary value problems of mathematical physics. We briefly describe this method. Let D be an open bounded region in $\mathbb{R}^3$ with a smooth boundary surface ∂D. The Dirichlet problem for Laplace's equation is: given a continuous function f on ∂D, find a function u, twice differentiable in D and continuous on $\bar{D}$, which satisfies

$$\Delta u(x) = 0 \qquad x \in D$$
$$u(x) = f(x) \qquad x \in \partial D$$

Let $K(x, y) = (x - y, n_y)/2\pi|x - y|^3$ where n_y is the outer normal to ∂D at the point $y \in \partial D$. Then, as a function of x, $K(x, y)$ satisfies $\Delta_x K(x, y) = 0$ in the interior which suggests that we try to write u as a superposition

$$u(x) = \int_{\partial D} K(x, y)\varphi(y)\, dS(y) \qquad \text{(VI.6a)}$$

where $\varphi(y)$ is some continuous function on ∂D and dS is the usual surface measure. Indeed, for $x \in D$, the integral makes perfectly good sense and $\Delta u(x) = 0$ in D. Furthermore, if x_0 is any point in ∂D and $x \to x_0$ from *inside* D, it can be proven that

$$u(x) \to -\varphi(x_0) + \int_{\partial D} K(x_0, y)\varphi(y)\, dS(y) \qquad \text{(VI.6b)}$$

If $x \to x_0$ from outside D, the minus is replaced by a plus. Also,

$$\int_{\partial D} K(x_0, y)\varphi(y)\, dS(y)$$

ists and is a continuous function on ∂D if φ is a continuous function on ∂D. The proof depends on the fact that the boundary of D is smooth which implies that for $x, y \in \partial D$, $(x - y, n_y) \approx c|x - y|^2$ as $x \to y$.

Since we wish $u(x) = f(x)$ on ∂D, the whole question reduces to whether we can find φ so that

$$f(x) = -\varphi(x) + \int_{\partial D} K(x, y)\varphi(y)\, dS(y), \qquad x \in \partial D$$

Let T: $C(\partial D) \to C(\partial D)$ be defined by

$$T\varphi = \int_{\partial D} K(x, y)\varphi(y)\, dS(y)$$

Not only is T bounded but (as we will shortly see) T is also compact. Thus, by the Fredholm alternative, either $\lambda = 1$ is in the point spectrum of T in which case there is a $\psi \in C(\partial D)$ such that $(I - T)\psi = 0$, or $-f = (I - T)\varphi$ has a unique solution for each $f \in C(\partial D)$. If u is defined by (VI.6a) with ψ replacing φ, then $u \equiv 0$ in D by the maximum principle. Further, $\partial u/\partial n$ is continuous across ∂D and therefore equals zero on ∂D. By an integration by parts this implies that $u \equiv 0$ outside ∂D. Therefore, by (VI.6b), $2\psi(x) \equiv 0$ on ∂D, so the first alternative does not hold.

The idea of the compactness proof is the following. Let

$$K_\delta(x, z) = \frac{(x - z, n_z)}{|x - z|^3 + \delta}$$

If $\delta > 0$, the kernel K_δ is continuous, so, by the discussion at the beginning of this section, the corresponding integral operators T_δ, are compact. To prove that T is compact, we need only show that $\|T - T_\delta\| \to 0$ as $\delta \to 0$. By the estimate

$$|(T_\delta f)(x) - (Tf)(x)| \le \|f\|_\infty \int_{\partial D} |K(x, z) - K_\delta(x, z)|\, dS(z)$$

we must only show that the integral converges to zero uniformly in x as $\delta \to 0$. To prove this, divide the integration region into the set where $|x - z| \ge \varepsilon$ and its complement. For fixed ε, the kernels converge uniformly on the first region. By using the fact that K is integrable, the contribution from the second region can be made arbitrarily small for ε sufficiently small.

44

Introduction to the Spectral Theorem

The following spectral theorem is a minor variant of the usual spectral theorem for matrices. This reformulation has the virtue of carrying over to general (unbounded) self adjoint operators on infinite dimensional Hilbert spaces.

Theorem 44.1. *Suppose A is an $n \times n$ complex self adjoint matrix, i.e. $A^* = A$ or equivalently $A_{ji} = \bar{A}_{ij}$ and let μ be counting measure on $\{1, 2, \ldots, n\}$. Then there exists a unitary map $U : \mathbb{C}^n \to L^2(\{1, 2, \ldots, n\}, d\mu)$ and a real function $\lambda : \{1, 2, \ldots, n\} \to \mathbb{R}$ such that $UA\xi = \lambda \cdot U\xi$ for all $\xi \in \mathbb{C}^n$. We summarize this equation by writing $UAU^{-1} = M_\lambda$ where*

$$M_\lambda : L^2(\{1, 2, \ldots, n\}, d\mu) \to L^2(\{1, 2, \ldots, n\}, d\mu)$$

is the linear operator, $g \in L^2(\{1, 2, \ldots, n\}, d\mu) \to \lambda \cdot g \in L^2(\{1, 2, \ldots, n\}, d\mu)$.

Proof. By the usual form of the spectral theorem for self-adjoint matrices, there exists an orthonormal basis $\{e_i\}_{i=1}^n$ of eigenvectors of A, say $Ae_i = \lambda_i e_i$ with $\lambda_i \in \mathbb{R}$. Define $U : \mathbb{C}^n \to L^2(\{1, 2, \ldots, n\}, d\mu)$ to be the unique (unitary) map determined by $Ue_i = \delta_i$ where

$$\delta_i(j) = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases}$$

and let $\lambda : \{1, 2, \ldots, n\} \to \mathbb{R}$ be defined by $\lambda(i) := \lambda_i$. ■

Definition 44.2. *Let $A : H \to H$ be a possibly unbounded operator on H. We let*

$$D(A^*) = \{y \in H : \ \exists\, z \in H \ \ni \ (Ax, y) = (x, z)\ \forall\, x \in D(A)\}$$

and for $y \in D(A^*)$ set $A^*y = z$.

Definition 44.3. *An operator A on H is* ***symmetric*** *if $A \subset A^*$ and is* ***self-adjoint*** *iff $A = A^*$.*

The reader should check that $A : H \to H$ is symmetric iff $(Ax, y) = (x, Ay)$ for all $x, y \in D(A)$.

Proposition 44.4. *Let (X, μ) be σ – finite measure space, $H = L^2(X, d\mu)$ and $f : X \to \mathbb{C}$ be a measurable function. Set $Ag = fg = M_f g$ for all*

$$g \in D(M_f) = \{g \in H : fg \in H\}.$$

Then $D(M_f)$ is a dense subspace of H and $M_f^* = M_{\bar{f}}$.

Proof. For any $g \in H = L^2(X, d\mu)$ and $m \in \mathbb{N}$, let $g_m := g1_{|f| \le m}$. Since $|fg_m| \le m\,|g|$ it follows that $fg_m \in H$ and hence $g_m \in D(M_f)$. By the dominated convergence theorem, it follows that $g_m \to g$ in H as $m \to \infty$, hence $D(M_f)$ is dense in H.

Suppose $h \in \mathcal{D}(M_f^*)$ then there exists $k \in L^2$ such that $(M_f g, h) = (g, k)$ for all $g \in D(M_f)$, i.e.

$$\int_X fg\bar{h}\, d\mu = \int_X g\bar{k}\, d\mu \text{ for all } g \in D(M_f)$$

or equivalently

$$\int_X g(\overline{\bar{f}h - k})d\mu = 0 \text{ for all } g \in D(M_f). \tag{44.1}$$

Choose $X_n \subset X$ such that $X_n \uparrow X$ and $\mu(X_n) < \infty$ for all n. It is easily checked that

$$g_n := 1_{X_n} \frac{\bar{f}h - k}{\left|\bar{f}h - k\right|} 1_{|f| \le n}$$

is in $D(M_f)$ and putting this function into Eq. (44.1) shows

$$\int_X \left|\bar{f}h - k\right| 1_{|f| \le n} d\mu = 0 \text{ for all } n.$$

Using the monotone convergence theorem, we may let $n \to \infty$ in this equation to find $\int_X \left|\bar{f}h - k\right| d\mu = 0$ and hence that $\bar{f}h = k \in L^2$. This shows $h \in D(M_{\bar{f}})$ and $M_f^* h = \bar{f}h$. ■

Theorem 44.5 (Spectral Theorem). *Suppose $A^* = A$ then there exists (X, μ) a σ – finite measure space, $f : X \to \mathbb{R}$ measurable, and $U : H \to L^2(x, \mu)$ unitary such that $UAU^{-1} = M_f$. Note this is a statement about domains as well, i.e. $UD(M_f) = D(A)$.*

I would like to give some examples of computing A^* and Theorem 44.5 as well. We will consider here the case of constant coefficient differential operators on $L^2(\mathbb{R}^n)$. First we need the following definition.

Definition 44.6. *Let $a_\alpha \in C^\infty(U)$, $L = \sum_{|\alpha| \le m} a_\alpha \partial^\alpha$ – a m^{th} order linear differential operator on $\mathcal{D}(U)$ and*

$$L^\dagger \phi = \sum_{|\alpha| \le m} (-1)^{|\alpha|}\, \partial^\alpha \left[a_\alpha \phi\right]$$

denote the ***formal adjoint*** *of L as in Lemma 40.4 above. For* $f \in L^p(U)$ *we say* $Lf \in L^p(U)$ *or* $L^p_{loc}(U)$ *if the generalized function* Lf *may be represented by an element of* $L^p(U)$ *or* $L^p_{loc}(U)$ *respectively, i.e.* $Lf = g \in L^p_{loc}(U)$ *iff*

$$\int_U f \cdot L^\dagger \phi \, dm = \int_U g\phi dm \text{ for all } \phi \in C_c^\infty(U). \tag{44.2}$$

In terms of the complex inner product,

$$(f, g) := \int_U f(x)\bar{g}(x)dm(x)$$

Eq. (44.2) is equivalent to

$$\left(f \cdot L^{\circledast}\phi\right) = (g, \phi) \text{ for all } \phi \in C_c^\infty(U)$$

where

$$L^{\circledast}\phi := \sum_{|\alpha|\le m} (-1)^{|\alpha|} \partial^\alpha \left[\bar{a}_\alpha \phi\right].$$

Notice that $L^{\circledast}$ *satisfies* $L^{\circledast}\bar{\phi} = \overline{L^\dagger \phi}$. *(We do not write* L^* *here since* $L^{\circledast}$ *is to be considered an operator on the space on* $\mathcal{D}'(U)$*.)*

Remark 44.7. Recall that if $f, h \in L^2(\mathbb{R}^n)$, then the following are equivalent

1. $\hat{f} = h$.
2. $(h, g) = \left(f, \mathcal{F}^{-1}g\right)$ for all $g \in C_c^\infty(\mathbb{R}^n)$.
3. $(h, g) = \left(f, \mathcal{F}^{-1}g\right)$ for all $g \in \mathcal{S}(\mathbb{R}^n)$.
4. $(h, g) = \left(f, \mathcal{F}^{-1}g\right)$ for all $g \in L^2(\mathbb{R}^n)$.

Indeed if $\hat{f} = h$ and $g \in L^2(\mathbb{R}^n)$, the unitarity of $\mathcal{F}$ implies

$$(h, g) = \left(\hat{f}, g\right) = (\mathcal{F}f, g) = \left(f, \mathcal{F}^{-1}g\right).$$

Hence $1 \implies 4$ and it is clear that $4 \implies 3 \implies 2$. If 2 holds, then again since $\mathcal{F}$ is unitary we have

$$(h, g) = \left(f, \mathcal{F}^{-1}g\right) = \left(\hat{f}, g\right) \text{ for all } g \in C_c^\infty(\mathbb{R}^n)$$

which implies $h = \hat{f}$ a.e., i.e. $h = \hat{f}$ in $L^2(\mathbb{R}^n)$.

Proposition 44.8. *Let* $p(x) = \sum_{|\alpha|\le m} a_\alpha x^\alpha$ *be a polynomial on* $\mathbb{C}^n$,

$$L := p(\partial) := \sum_{|\alpha|\le m} a_\alpha \partial^\alpha \tag{44.3}$$

and $f \in L^2(\mathbb{R}^n)$. *Then* $Lf \in L^2(\mathbb{R}^n)$ *iff* $p(i\xi)\hat{f}(\xi) \in L^2(\mathbb{R}^n)$ *and in which case*

$$(Lf)\hat{}\ (\xi) = p(i\xi)\hat{f}(\xi). \tag{44.4}$$

Put more concisely, letting

$$D(B) = \left\{f \in L^2(\mathbb{R}^n) : Lf \in L^2(\mathbb{R}^n)\right\}$$

with $Bf = Lf$ *for all* $f \in D(B)$, *we have*

$$\mathcal{F}B\mathcal{F}^{-1} = M_{p(i\xi)}.$$

Proof. As above, let

$$L^\dagger := \sum_{|\alpha|\le m} a_\alpha (-\partial)^\alpha \text{ and } L^{\circledast} := \sum_{|\alpha|\le m} \bar{a}_\alpha (-\partial)^\alpha. \tag{44.5}$$

For $\phi \in C_c^\infty(\mathbb{R}^n)$,

$$L^{\circledast}\phi^\vee(x) = L^{\circledast}\int \phi(\xi) e^{ix\cdot\xi} d\lambda(\xi) = \sum_{|\alpha|\le m} \bar{a}_\alpha (-\partial_x)^\alpha \int \phi(\xi) e^{ix\cdot\xi} d\lambda(\xi)$$
$$= \int \overline{p(i\xi)}\phi(\xi) e^{ix\cdot\xi} d\lambda(\xi) = \mathcal{F}^{-1}\left[\overline{p(i\xi)}\phi(\xi)\right](x)$$

So if $f \in L^2(\mathbb{R}^n)$ such that $Lf \in L^2(\mathbb{R}^n)$. Then by Remark 44.7,

$$(\widehat{Lf}, \phi) = (Lf, \phi^\vee) = \langle f, L^{\circledast}\phi^\vee\rangle = \langle f(x), \mathcal{F}^{-1}\left[\overline{p(i\xi)}\phi(\xi)\right](x)\rangle$$
$$= \langle \hat{f}(\xi), \left[\overline{p(i\xi)}\phi(\xi)\right]\rangle = \langle p(i\xi)\hat{f}(\xi), \phi(\xi)\rangle \text{ for all } \phi \in C_c^\infty(\mathbb{R}^n)$$

from which it follows that Eq. (44.4) holds and that $p(i\xi)\hat{f}(\xi) \in L^2(\mathbb{R}^n)$.

Conversely, if $f \in L^2(\mathbb{R}^n)$ is such that $p(i\xi)\hat{f}(\xi) \in L^2(\mathbb{R}^n)$ then for $\phi \in C_c^\infty(\mathbb{R}^n)$,

$$\left(f, L^{\circledast}\phi\right) = \left(\hat{f}, \mathcal{F}L^{\circledast}\phi\right). \tag{44.6}$$

Since

$$\mathcal{F}\left(L^{\circledast}\phi\right)(\xi) = \int L^{\circledast}\phi(x) e^{-ix\cdot\xi} d\lambda(x) = \int \phi(x) \overline{L_x} e^{-ix\cdot\xi} d\lambda(x)$$
$$= \int \phi(x) \overline{a_\alpha} \partial_x^\alpha e^{-ix\cdot\xi} d\lambda(x) = \int \phi(x) \overline{a_\alpha} (-i\xi)^\alpha e^{-ix\cdot\xi} d\lambda(x)$$
$$= \overline{p(i\xi)}\hat{\phi}(\xi),$$

Eq. (44.6) becomes

$$\left(f, L^{\circledast}\phi\right) = \left(\hat{f}(\xi), \overline{p(i\xi)}\hat{\phi}(\xi)\right) = \left(p(i\xi)\hat{f}(\xi), \hat{\phi}(\xi)\right)$$
$$= \left(\mathcal{F}^{-1}\left[p(i\xi)\hat{f}(\xi)\right](x), \phi(x)\right).$$

This shows $Lf = \mathcal{F}^{-1}\left[p(i\xi)\hat{f}(\xi)\right] \in L^2(\mathbb{R}^n)$. ∎

Lemma 44.9. *Suppose $p(x) = \sum_{|\alpha| \le m} a_\alpha x^\alpha$ is a polynomial on $\mathbb{R}^n$ and $L = p(\partial)$ is the constant coefficient differential operator $B = \sum_{|\alpha| \le m} a_\alpha \partial^\alpha$ with $D(B) := \mathcal{S}(\mathbb{R}^n) \subset L^2(\mathbb{R}^n)$. Then*

$$\mathcal{F} B \mathcal{F}^{-1} = M_{p(i\xi)}|_{\mathcal{S}(\mathbb{R}^n)}.$$

Proof. This is result of the fact that $\mathcal{F}(\mathcal{S}(\mathbb{R}^n)) = \mathcal{S}(\mathbb{R}^n)$ and for $f \in \mathcal{S}(\mathbb{R}^n)$ we have

$$f(x) = \int_{\mathbb{R}^n} \hat{f}(\xi) e^{i\xi \cdot x} d\lambda(\xi)$$

so that

$$Bf(x) = \int_{\mathbb{R}^n} \hat{f}(\xi) L_x e^{i\xi \cdot x} d\lambda(\xi) = \int_{\mathbb{R}^n} \hat{f}(\xi) p(i\xi) e^{i\xi \cdot x} d\lambda(\xi)$$

so that

$$(Bf)\hat{}\,(\xi) = p(i\xi)\hat{f}(\xi) \text{ for all } f \in \mathcal{S}(\mathbb{R}^n).$$

■

Lemma 44.10. *Suppose $g : \mathbb{R}^n \to \mathbb{C}$ is a measurable function such that $|g(x)| \le C\left(1 + |x|^M\right)$ for some constants C and M. Let A be the unbounded operator on $L^2(\mathbb{R}^n)$ defined by $D(A) = \mathcal{S}(\mathbb{R}^n)$ and for $f \in \mathcal{S}(\mathbb{R}^n)$, $Af = gf$. Then $A^* = M_{\bar{g}}$.*

Proof. If $h \in D(M_{\bar{g}})$ and $f \in D(A)$, we have

$$(Af, h) = \int_{\mathbb{R}^n} gf\bar{h} dm = \int_{\mathbb{R}^n} f\overline{\bar{g}h} dm = (f, M_{\bar{g}} h)$$

which shows $M_{\bar{g}} \subset A^*$, i.e. $h \in D(A^*)$ and $A^* h = M_{\bar{g}} h$. Now suppose $h \in D(A^*)$ and $A^* h = k$, i.e.

$$\int_{\mathbb{R}^n} gf\bar{h} dm = (Af, h) = (f, k) = \int_{\mathbb{R}^n} f\bar{k} dm \text{ for all } f \in \mathcal{S}(\mathbb{R}^n)$$

or equivalently that

$$\int_{\mathbb{R}^n} \left(g\bar{h} - \bar{k}\right) f dm = 0 \text{ for all } f \in \mathcal{S}(\mathbb{R}^n).$$

Since the last equality (even just for $f \in C_c^\infty(\mathbb{R}^n)$) implies $g\bar{h} - \bar{k} = 0$ a.e. we may conclude that $h \in D(M_{\bar{g}})$ and $k = M_{\bar{g}} h$, i.e. $A^* \subset M_{\bar{g}}$. ■

Theorem 44.11. *Suppose $p(x) = \sum_{|\alpha| \le m} a_\alpha x^\alpha$ is a polynomial on $\mathbb{R}^n$ and $A = p(\partial)$ is the constant coefficient differential operator with $D(A) := C_c^\infty(\mathbb{R}^n) \subset L^2(\mathbb{R}^n)$ such that $A = L = p(\partial)$ on $D(A)$, see Eq. (44.3). Then A^* is the operator described by*

$$D(A^*) = \left\{f \in L^2(\mathbb{R}^n) : L^\dagger f \in L^2(\mathbb{R}^n)\right\} \\ = \left\{f \in L^2(\mathbb{R}^n) : p(i\xi)\hat{f}(\xi) \in L^2(\mathbb{R}^n)\right\}$$

and $A^* f = L^\dagger f$ for $f \in D(A^*)$ where $L^\dagger$ is defined in Eq. (44.5) above. Moreover we have $\mathcal{F} A^* \mathcal{F}^{-1} = M_{\overline{p(i\xi)}}$.

Proof. Let $D(B) = \mathcal{S}(\mathbb{R}^n)$ and $B := L$ on $D(B)$ so that $A \subset B$. We are first going to show $A^* = B^*$. As is easily verified, in general if $A \subset B$ then $B^* \subset A^*$. So we need only show $A^* \subset B^*$. Now by definition, if $g \in D(A^*)$ with $k = A^* g$, then

$$(Af, g) = (f, k) \text{ for all } f \in D(A) := C_c^\infty(\mathbb{R}^n).$$

Suppose that $f \in \mathcal{S}(\mathbb{R}^n)$ and $\phi \in C_c^\infty(\mathbb{R}^n)$ such that $\phi = 1$ in a neighborhood of 0. Then $f_n(x) := \phi(x/n) f(x)$ is in $\mathcal{S}(\mathbb{R}^n)$ and hence

$$(f_n, k) = (Lf_n, g). \tag{44.7}$$

An exercise in the product rule and the dominated convergence theorem shows $f_n \to f$ and $Lf_n \to Lf$ in $L^2(\mathbb{R}^n)$ as $n \to \infty$. Therefore we may pass to the limit in Eq. (44.7) to learn

$$(f, k) = (Bf, g) \text{ for all } f \in \mathcal{S}(\mathbb{R}^n)$$

which shows $g \in D(B^*)$ and $B^* g = k$.

By Lemma 44.10, we may conclude that $A^* = B^* = M_{\overline{p(i\xi)}}$ and by Proposition 44.8 we then conclude that

$$D(A^*) = \left\{f \in L^2(\mathbb{R}^n) : p(i\xi)\hat{f}(\xi) \in L^2(\mathbb{R}^n)\right\} \\ = \left\{f \in L^2(\mathbb{R}^n) : L^\dagger f \in L^2(\mathbb{R}^n)\right\}$$

and for $f \in D(A^*)$ we have $A^* f = L^\dagger f$. ■

Example 44.12. If we take $L = \Delta$ with $D(L) := C_c^\infty(\mathbb{R}^n)$, then

$$L^* = \bar{\Delta} = \mathcal{F} M_{-|\xi|^2} \mathcal{F}^{-1}$$

where $D(\bar{\Delta}) = \left\{f \in L^2(\mathbb{R}^n) : \Delta f \in L^2(\mathbb{R}^n)\right\}$ and $\bar{\Delta} f = \Delta f$.

Theorem 44.13. *Suppose $A = A^*$ and $A \le 0$. Then for all $u_0 \in D(A)$ there exists a unique solution $u \in C^1([0, \infty))$ such that $u(t) \in D(A)$ for all t and*

$$\dot{u}(t) = Au(t) \text{ with } u(0) = u_0. \tag{44.8}$$

Writing $u(t) = e^{tA} u_0$, the map $u_0 \to e^{tA} u_0$ is a linear contraction semi-group, i.e.

$$\|e^{tA} u_0\| \le \|u_0\| \text{ for all } t \ge 0. \tag{44.9}$$

So e^{tA} extends uniquely to H by continuity. This extension satisfies:

1. ***Strong Continuity:*** *the map* $t \in [0,\infty) \to e^{tA}u_0$ *is continuous for all* $u_0 \in H$.
2. ***Smoothing property:*** $t > 0$

$$e^{tA}u_0 \in \bigcap_{n=0}^{\infty} D(A^n) =: C^\infty(A)$$

and

$$\|A^k e^{tA}\| \le \left(\frac{k}{t}\right)^k e^{-k} \text{ for all } k \in \mathbb{N}. \tag{44.10}$$

Proof. Uniqueness. Suppose u solves Eq. (44.8), then

$$\frac{d}{dt}(u(t),u(t)) = 2\mathrm{Re}(\dot{u},u) = 2\mathrm{Re}(Au,u) \le 0.$$

Hence $\|u(t)\|$ is decreasing so that $\|u(t)\| \le \|u_0\|$.This implies the uniqueness assertion in the theorem and the norm estimate in Eq. (44.9).

Existence: By the spectral theorem we may assume $A = M_f$ acting on $L^2(X,\mu)$ for some σ – finite measure space (X,μ) and some measurable function $f: X \to (-\infty,0]$. We wish to show $u(t) = e^{tf}u_0 \in L^2$ solves

$$\dot{u}(t) = fu(t) \text{ with } u(0) = u_0 \in D(M_f) \subset L^2.$$

Let $t > 0$ and $|\Delta| < t$. Then by the mean value inequality

$$\left|\frac{e^{(t+\triangle)f} - e^{tf}}{\triangle}u_0\right| = \max\left\{|fe^{(t+\tilde{\triangle})f}u_0| : \tilde{\Delta} \text{ between } 0 \text{ and } \Delta\right\} \le |fu_0| \in L^2.$$

This estimated along with the fact that

$$\frac{u(t+\Delta)-u(t)}{\triangle} = \frac{e^{(t+\triangle)f} - e^{tf}}{\triangle}u_0 \overset{\text{point wise}}{\to} fe^{tf}u_0 \text{ as } \Delta \to 0$$

enables us to use the dominated convergence theorem to conclude

$$\dot{u}(t) = L^2\text{–}\lim_{\Delta\to 0}\frac{u(t+\triangle)-u(t)}{\triangle} = e^{tf}fu_0 = fu(t)$$

as desired. i.e. $\dot{u}(t) = fu(t)$.

The extension of e^{tA} to H is given by $M_{e^{tf}}$. For $g \in L^2$, $\left|e^{tf}g\right| \le |g| \in L^2$ and $e^{tf}g \to e^{\tau g}f$ pointwise as $t \to \tau$, so the Dominated convergence theorem shows $t \in [0,\infty) \to e^{tA}g \in H$ is continuous. For the last two assertions, let $t > 0$ and $f(x) = x^k e^{tx}$. Then $(\ln f)'(x) = \frac{k}{x} + t$ which is zero when $x = -k/t$ and therefore

$$\max_{x\le 0}\left|x^k e^{tx}\right| = |f(-k/t)| = \left(\frac{k}{t}\right)^k e^{-k}.$$

Hence

$$\|A^k e^{tA}\|_{op} \le \max_{x\le 0}\left|x^k e^{tx}\right| \le \left(\frac{k}{t}\right)^k e^{-k} < \infty.$$

■

Theorem 44.14. *Take* $A = \mathcal{F}M_{-|\xi|^2}\mathcal{F}^{-1}$ *so* $A|_{\mathcal{S}} = \Delta$ *then*

$$C^\infty(A) := \bigcap_{n=1}^{\infty} D(A^n) \subset C^\infty\left(\mathbb{R}^d\right)$$

i.e. for all $f \in C^\infty(A)$ *there exists a version* $\tilde{f}$ *of* f *such that* $\tilde{f} \in C^\infty(\mathbb{R}^d)$.

Proof. By assumption $|\xi|^{2n}\hat{f}(\xi) \in L^2$ for all n. Therefore $\hat{f}(\xi) = \frac{g_n(\xi)}{1+|\xi|^{2n}}$ for some $g_n \in L^2$ for all n. Therefore for n chosen so that $2n > m + d$, we have

$$\int_{\mathbb{R}^d} |\xi|^m|\hat{f}(\xi)|d\xi \le \|g_n\|_{L^2}\left\|\frac{|\xi|^m}{1+|\xi|^{2n}}\right\|_2 < \infty$$

which shows $|\xi|^m|\hat{f}(\xi)| \in L^1$ for all $m = 0,1,2,\ldots$ We may now differentiate the inversion formula, $f(x) = \int \hat{f}(\xi)e^{ix\cdot\xi}d\xi$ to find

$$D^\alpha f(x) = \int (i\xi)^\alpha \hat{f}(\xi)e^{ix\cdot\xi}d\xi \text{ for any } \alpha$$

and thus conclude $f \in C^\infty$. ■

Exercise 44.15. Some Exercises: Section 2.5 4, 5, 6, 8, 9, 11, 12, 17.

44.1 Du Hammel's principle again

Lemma 44.16. *Suppose* A *is an operator on* H *such that* A^* *is densely defined then* A^* *is closed.*

Proof. If $f_n \in D(A^*) \to f \in H$ and $A^*f_n \to g$ then for all $h \in D(A)$

$$(g,h) = \lim_{n\to\infty}(A^*f_n,h)$$

while

$$\lim_{n\to\infty}(A^*f_n,h) = \lim_{n\to\infty}(f_n,Ah) = (f,Ah),$$

i.e. $(Ah,f) = (h,g)$ for all $h \in D(A)$. Thus $f \in D(A^*)$ and $A^*f = g$. ■

Corollary 44.17. *If* $A^* = A$ *then* A *is closed.*

Corollary 44.18. *Suppose* A *is closed and* $u(t) \in D(A)$ *is a path such that* $u(t)$ *and* $Au(t)$ *are continuous in* t*. Then*

$$A\int_0^T u(\tau)d\tau = \int_0^T Au(\tau)d\tau.$$

Proof. Let π_n be a sequence of partitions of $[0,T]$ such that $mesh(\pi_n) \to 0$ as $n \to \infty$ and set

$$f_n = \sum_{\pi_n} u(\tau_i)(\tau_{i+1} - \tau_i) \in D(A).$$

Then $f_n \to \int_0^T u(\tau)d\tau$ and

$$A f_n = \sum_{\pi_n} Au(\tau_i)(\tau_{i+1} - \tau_i) \to \int_0^T Au(\tau)d\tau.$$

Therefore $\int_0^T u(\tau)d\tau \in D(A)$ and $A\int_0^T u(\tau)d\tau = \int_0^T Au(\tau)d\tau$. ■

Lemma 44.19. *Suppose $A = A^*$, $A \le 0$, and $h : [0,\infty] \to H$ is continuous. Then*

$$(s,t) \in [0,\infty) \times [0,\infty) \to e^{sA}h(t)$$
$$(s,t) \in (0,\infty) \times [0,\infty) \to A^k e^{sA}h(t)$$

are continuous maps into H.

Proof. Let $k \ge 0$, then if $s \ge \sigma$,

$$\begin{aligned}
&\|A^k\left(e^{sA}h(t) - e^{\sigma A}h(\tau)\right)\| \\
&\quad = \left\|A^k e^{\sigma A}\left(e^{(s-\sigma)A}h(t) - h(\tau)\right)\right\| \\
&\quad \le \left\|A^k e^{\sigma A}\right\| \left\|e^{(s-\sigma)A}\left[h(t) - h(\tau)\right] + e^{(s-\sigma)A}h(\tau) - h(\tau)\right\| \\
&\quad \le \left(\frac{k}{\sigma}\right)^k e^{-k} \cdot \left[\|h(t) - h(\tau)\| + \left\|e^{(s-\sigma)A}h(\tau) - h(\tau)\right\|\right].
\end{aligned}$$

So

$$\lim_{s\downarrow\sigma \text{ and } t\to\tau} \left\|A^k\left(e^{sA}h(t) - e^{\sigma A}h(\tau)\right)\right\| = 0$$

and we may take $\sigma = 0$ if $k = 0$. Similarly, if $s \le \sigma$,

$$\begin{aligned}
&\|A^k\left(e^{sA}h(t) - e^{\sigma A}h(\tau)\right)\| \\
&\quad = \left\|A^k e^{sA}\left(h(t) - e^{(\sigma-s)A}h(\tau)\right)\right\| \\
&\quad \le \left\|A^k e^{sA}\right\| \left[\|h(t) - h(\tau)\| + \left\|h(\tau) - e^{(\sigma-s)A}h(\tau)\right\|\right] \\
&\quad \le \left(\frac{k}{s}\right)^k e^{-k} \left[\|h(t) - h(\tau)\| + \left\|h(\tau) - e^{(\sigma-s)A}h(\tau)\right\|\right]
\end{aligned}$$

and the latter expression tends to zero as $s \uparrow \sigma$ and $t \to \tau$. ■

Lemma 44.20. *Let $h \in C([0,\infty), H)$, $D := \left\{(s,t) \in \mathbb{R}^2 : s > t \ge 0\right\}$ and $F(s,t) := \int_0^t e^{(s-\tau)A}h(\tau)d\tau$ for $(s,t) \in D$. Then*

1. *$F \in C^1(D,H)$ (in fact $F \in C^\infty(D,H)$),*
$$\frac{\partial}{\partial t}F(s,t) = e^{(s-t)A}h(t) \tag{44.11}$$
and
$$\frac{\partial F(s,t)}{\partial s} = \int_0^t Ae^{(s-\tau)A}h(\tau)d\tau. \tag{44.12}$$
2. *Given $\epsilon > 0$ let*
$$u_\epsilon(t) := F(t+\epsilon, t) = \int_0^t e^{(t+\epsilon-\tau)A}h(\tau)d\tau.$$
Then $u_\epsilon \in C^1((-\epsilon,\infty), H)$, $u_\epsilon(t) \in D(A)$ for all $t > -\epsilon$ and
$$\dot{u}_\epsilon(t) = e^{\epsilon A}h(t) + Au_\epsilon(t). \tag{44.13}$$

Proof. We claim the function

$$(s,t) \in D \to F(s,t) := \int_0^t e^{(s-\tau)A}h(\tau)d\tau$$

is continuous. Indeed if $(s',t') \in D$ and $(s,t) \in D$ is sufficiently close to (s',t') so that $s > t'$, we have

$$\begin{aligned}
F(s,t) - F(s',t') &= \int_0^t e^{(s-\tau)A}h(\tau)d\tau - \int_0^{t'} e^{(s'-\tau)A}h(\tau)d\tau \\
&= \int_0^t e^{(s-\tau)A}h(\tau)d\tau - \int_0^{t'} e^{(s-\tau)A}h(\tau)d\tau \\
&\quad + \int_0^{t'} \left[e^{(s-\tau)A} - e^{(s'-\tau)A}\right]h(\tau)d\tau
\end{aligned}$$

so that

$$\begin{aligned}
\|F(s,t) - F(s',t')\| &\le \left|\int_{t'}^t \left\|e^{(s-\tau)A}h(\tau)\right\| d\tau\right| \\
&\quad + \int_0^{t'} \left\|\left[e^{(s-\tau)A} - e^{(s'-\tau)A}\right]h(\tau)\right\| d\tau \\
&\le \left|\int_{t'}^t \|h(\tau)\| d\tau\right| \\
&\quad + \int_0^{t'} \left\|\left[e^{(s-\tau)A} - e^{(s'-\tau)A}\right]h(\tau)\right\| d\tau.
\end{aligned} \tag{44.14}$$

By the dominated convergence theorem,

$$\lim_{(s,t)\to(s',t')} \left|\int_{t'}^t \|h(\tau)\| d\tau\right| = 0$$

and

$$\lim_{(s,t)\to(s',t')} \int_0^{t'} \left\| \left[e^{(s-\tau)A} - e^{(s'-\tau)A} \right] h(\tau) \right\| d\tau = 0$$

which along with Eq. (44.14) shows F is continuous.

By the fundamental theorem of calculus,

$$\frac{\partial}{\partial t} F(s,t) = e^{(s-t)A} h(t)$$

and as we have seen this expression is continuous on D. Moreover, since

$$\frac{\partial}{\partial s} e^{(s-\tau)A} h(\tau) = A e^{(s-\tau)A} h(\tau)$$

is continuous and bounded for on $s > t > \tau$, we may differentiate under the integral to find

$$\frac{\partial F(s,t)}{\partial s} = \int_0^t A e^{(s-\tau)A} h(\tau) d\tau \text{ for } s > t.$$

A similar argument (making use of Eq. (44.10) with $k = 1$) shows $\frac{\partial F(s,t)}{\partial s}$ is continuous for $(s,t) \in D$.

By the chain rule, $u_\epsilon(t) := F(t+\epsilon, t)$ is C^1 for $t > -\epsilon$ and

$$\begin{aligned} \dot{u}_\epsilon(t) &= \frac{\partial F(t+\epsilon,t)}{\partial s} + \frac{\partial F(t+\epsilon,t)}{\partial t} \\ &= e^{\epsilon A} h(t) + \int_0^t A e^{(s-\tau+\epsilon)A} h(\tau) d\tau = e^{\epsilon A} h(t) + u_\epsilon(t). \end{aligned}$$

■

Theorem 44.21. *Suppose $A = A^*$, $A \le 0$, $u_0 \in H$ and $h : [0,\infty) \to H$ is continuous. Assume further that $h(t) \in D(A)$ for all $t \in [0,\infty)$ and $t \to Ah(t)$ is continuous, then*

$$u(t) := e^{tA} u_0 + \int_0^t e^{(t-\tau)A} h(\tau) d\tau \tag{44.15}$$

is the unique function $u \in C^1((0,\infty), H) \cap C([0,\infty), H)$ such that $u(t) \in D(A)$ for all $t > 0$ satisfying the differential equation

$$\dot{u}(t) = Au(t) + h(t) \textit{ for } t > 0 \textit{ and } u(0+) = u_0.$$

Proof. Uniqueness: If $v(t)$ is another such solution then $w(t) := u(t) - v(t)$ satisfies,

$$\dot{w}(t) = Aw(t) \text{ with } w(0+) = 0$$

which we have already seen implies $w = 0$.

Existence: By linearity and Theorem 44.13 we may assume with out loss of generality that $u_0 = 0$ in which case

$$u(t) = \int_0^t e^{(t-\tau)A} h(\tau) d\tau.$$

By Lemma 44.19, we know $\tau \in [0,t] \to e^{(t-\tau)A} h(\tau) \in H$ is continuous, so the integral in Eq. (44.15) is well defined. Similarly by Lemma 44.19,

$$\tau \in [0,t] \to e^{(t-\tau)A} Ah(\tau) = Ae^{(t-\tau)A} h(\tau) \in H$$

and so by Corollary 44.18, $u(t) \in D(A)$ for all $t \ge 0$ and

$$Au(t) = \int_0^t Ae^{(t-\tau)A} h(\tau) d\tau = \int_0^t e^{(t-\tau)A} Ah(\tau) d\tau.$$

Let

$$u_\epsilon(t) = \int_0^t e^{(t+\epsilon-\tau)A} h(\tau) d\tau$$

be defined as in Lemma 44.20. Then using the dominated convergence theorem,

$$\begin{aligned} \sup_{t\le T} \|u_\epsilon(t) - u(t)\| &\le \sup_{t\le T} \int_0^t \left\| \left(e^{(t+\epsilon-\tau)A} - e^{(t-\tau)A} \right) h(\tau) \right\| d\tau \\ &\le \int_0^T \left\| \left(e^{\epsilon A} - I \right) h(\tau) \right\| d\tau \to 0 \text{ as } \epsilon \downarrow 0, \end{aligned}$$

$$\sup_{t\le T} \|Au_\epsilon(t) - Au(t)\| \le \int_0^T \left\| \left(e^{\epsilon A} - I \right) Ah(\tau) \right\| d\tau \to 0 \text{ as } \epsilon \downarrow 0$$

and

$$\left\| \int_0^t e^{\epsilon A} h(\tau) d\tau - \int_0^t h(\tau) d\tau \right\| \le \int_0^t \left\| \left(e^{\epsilon A} - I \right) h(\tau) \right\| d\tau \to 0 \text{ as } \epsilon \downarrow 0.$$

Integrating Eq. (44.13) shows

$$u_\epsilon(t) = \int_0^t e^{\epsilon A} h(\tau) d\tau + \int_0^t Au_\epsilon(\tau) d\tau \tag{44.16}$$

and then passing to the limit as $\epsilon \downarrow 0$ in this equations shows

$$u(t) = \int_0^t h(\tau) d\tau + \int_0^t Au(\tau) d\tau.$$

This shows u is differentiable and $\dot{u}(t) = h(t) + Au(t)$ for all $t > 0$. ■

Theorem 44.22. *Let $\alpha > 0$, $h : [0,\infty) \to H$ be a locally α – Holder continuous function, $A = A^*$, $A \le 0$ and $u_0 \in H$. The function*

$$u(t) := e^{tA}u_0 + \int_0^t e^{(t-\tau)A}h(\tau)d\tau$$

is the unique function $u \in C^1((0,\infty),H) \cap C([0,\infty),H)$ such that $u(t) \in D(A)$ for all $t > 0$ satisfying the differential equation

$$\dot{u}(t) = Au(t) + h(t) \text{ for } t > 0 \text{ and } u(0+) = u_0.$$

(For more details see Pazy [9, §5.7].)

Proof. The proof of uniqueness is the same as in Theorem 44.21 and for existence we may assume $u_0 = 0$.

With out loss of generality we may assume $u_0 = 0$ so that

$$u(t) = \int_0^t e^{(t-\tau)A}h(\tau)d\tau.$$

By Lemma 44.19, we know $\tau \in [0,t] \to e^{(t-\tau)A}h(\tau) \in H$ is continuous, so the integral defining u is well defined. For $\epsilon > 0$, let

$$u_\epsilon(t) := \int_0^t e^{(t+\epsilon-\tau)A}h(\tau)d\tau = \int_0^t e^{(t-\tau)A}e^{\epsilon A}h(\tau)d\tau.$$

Notice that $v(\tau) := e^{\epsilon A}h(\tau) \in C^\infty(A)$ for all τ and moreover since $Ae^{\epsilon A}$ is a bounded operator, it follows that $\tau \to Av(\tau)$ is continuous. So by Lemma 44.19, it follows that $\tau \in [0,t] \to Ae^{(t-\tau)A}v(\tau) \in H$ is continuous as well. Hence we know $u_\epsilon(t) \in D(A)$ and

$$Au_\epsilon(t) = \int_0^t Ae^{(t-\tau)A}e^{\epsilon A}h(\tau)d\tau.$$

Now

$$Au_\epsilon(t) = \int_0^t Ae^{(t+\epsilon-\tau)A}h(t)d\tau + \int_0^t Ae^{(t+\epsilon-\tau)A}\left[h(\tau) - h(t)\right]d\tau,$$

$$\int_0^t Ae^{(t+\epsilon-\tau)A}h(t)d\tau = -e^{(t+\epsilon-\tau)A}h(t)|_{\tau=0}^{\tau=t} = e^{(t+\epsilon)A}h(t) - e^{\epsilon A}h(t)$$

and

$$\begin{aligned}\left\|Ae^{(t+\epsilon-\tau)A}\left[h(\tau)-h(t)\right]\right\| &\le e^{-1}\frac{1}{(t+\epsilon-\tau)}\left\|h(\tau)-h(t)\right\| \\ &\le Ce^{-1}\frac{1}{(t+\epsilon-\tau)}\left|t-\tau\right|^\alpha \le Ce^{-1}\left|t-\tau\right|^{\alpha-1}.\end{aligned}$$

These results along with the dominated convergence theorem shows $\lim_{\epsilon\downarrow 0} Au_\epsilon(t)$ exists and is given by

$$\begin{aligned}\lim_{\epsilon\downarrow 0} Au_\epsilon(t) &= \lim_{\epsilon\downarrow 0}\left[e^{(t+\epsilon)A}h(t) - e^{\epsilon A}h(t)\right] \\ &\quad + \lim_{\epsilon\downarrow 0}\int_0^t Ae^{(t+\epsilon-\tau)A}\left[h(\tau)-h(t)\right]d\tau \\ &= e^{tA}h(t) - h(t) + \int_0^t Ae^{(t-\tau)A}\left[h(\tau)-h(t)\right]d\tau.\end{aligned}$$

Because A is a closed operator, it follows that $u(t) \in D(A)$ and

$$Au(t) = e^{tA}h(t) - h(t) + \int_0^t Ae^{(t-\tau)A}\left[h(\tau)-h(t)\right]d\tau.$$

Claim: $t \to Au(t)$ is continuous. To prove this it suffices to show

$$v(t) := A\int_0^t e^{(t-\tau)A}(h(\tau)-h(t))d\tau$$

is continuous and for this we have

$$\begin{aligned}v(t+\triangle) - v(t) &= \int_0^{t+\triangle} Ae^{(t+\triangle-\tau)A}(h(\tau)-h(t+\triangle))d\tau \\ &\quad - \int_0^t Ae^{(t-\tau)A}(h(\tau)-h(t))d\tau \\ &= I + II\end{aligned}$$

where

$$\begin{aligned}I &= \int_t^{t+\triangle} Ae^{(t+\triangle-\tau)A}(h(\tau)-h(t+\triangle))d\tau \text{ and} \\ II &= \int_0^t \left[Ae^{(t+\triangle-\tau)A}(h(\tau)-h(t+\triangle)) - Ae^{(t-\tau)A}(h(\tau)-h(t))\right]d\tau \\ &= \int_0^t \left[Ae^{(t+\triangle-\tau)A}(h(\tau)-h(t)) - Ae^{(t-\tau)A}(h(\tau)-h(t))\right]d\tau \\ &\quad + \int_0^t \left[Ae^{(t+\triangle-\tau)A}(h(t)-h(t+\triangle))\right]d\tau \\ &= II_1 + II_2\end{aligned}$$

and

$$\begin{aligned}II_1 &= \int_0^t A\left[e^{(t+\triangle-\tau)A} - e^{(t-\tau)A}\right](h(\tau)-h(t))d\tau \text{ and} \\ II_2 &= \left[e^{(t+\Delta)A} - e^{\Delta A}\right](h(t)-h(t+\triangle)).\end{aligned}$$

We estimate I as

$$\begin{aligned}\|I\| &\leq \left|\int_t^{t+\triangle} \left\|Ae^{(t+\triangle-\tau)A}(h(\tau)-h(t+\triangle))\right\| d\tau\right| \\ &\leq C\left|\int_t^{t+\triangle} \frac{1}{t+\triangle-\tau}\,|t+\triangle-\tau|^{\alpha}\, d\tau\right| \\ &= C\int_0^{|\Delta|} x^{\alpha-1}dx = C\alpha^{-1}\,|\Delta|^{\alpha} \to 0 \text{ as } \Delta\to 0.\end{aligned}$$

It is easily seen that $\|II_2\| \leq 2C\,|\Delta|^{\alpha} \to 0$ as $\Delta \to 0$ and

$$\left\|A\left[e^{(t+\triangle-\tau)A}-e^{(t-\tau)A}\right](h(\tau)-h(t))\right\| \leq C\,|t-\tau|^{\alpha-1}$$

which is integrable, so by the dominated convergence theorem,

$$\|II_1\| \leq \int_0^t \left\|A\left[e^{(t+\triangle-\tau)A}-e^{(t-\tau)A}\right](h(\tau)-h(t))\right\| d\tau \to 0 \text{ as } \Delta\to 0.$$

This completes the proof of the claim.

Moreover,

$$\begin{aligned}Au_\epsilon(t)-Au(t) &= e^{(t+\epsilon)A}h(t)-e^{tA}h(t)+h(t)-e^{\epsilon A}h(t) \\ &\quad+\int_0^t A\left(e^{(t+\epsilon-\tau)A}-e^{(t-\tau)A}\right)[h(\tau)-h(t)]\,d\tau\end{aligned}$$

so that

$$\begin{aligned}\|Au_\epsilon(t)-Au(t)\| \leq& 2\left\|h(t)-e^{\epsilon A}h(t)\right\| \\ &+\int_0^t \left\|Ae^{(t-\tau)A}\left(e^{\epsilon A}-I\right)[h(\tau)-h(t)]\right\| d\tau \\ \leq& 2\left\|h(t)-e^{\epsilon A}h(t)\right\| \\ &+e^{-1}\int_0^t \frac{1}{|t-\tau|}\left\|\left(e^{\epsilon A}-I\right)[h(\tau)-h(t)]\right\| d\tau\end{aligned}$$

from which it follows $Au_\epsilon(t) \to Au(t)$ boundedly. We may now pass to the limit in Eq. (44.16) to find

$$\begin{aligned}u(t) = \lim_{\epsilon\downarrow 0} u_\epsilon(t) &= \lim_{\epsilon\downarrow 0}\left[\int_0^t e^{\epsilon A}h(\tau)d\tau+\int_0^t Au_\epsilon(\tau)d\tau\right] \\ &= \int_0^t h(\tau)d\tau+\int_0^t Au(\tau)d\tau\end{aligned}$$

from which it follows that $u \in C^1((0,\infty),H)$ and $\dot{u}(t) = h(t)+Au(t)$. ■

45

Heat Equation

The heat equation for a function $u:\mathbb{R}_+\times\mathbb{R}^n\to\mathbb{C}$ is the partial differential equation

$$\left(\partial_t-\frac{1}{2}\Delta\right)u=0 \text{ with } u(0,x)=f(x), \tag{45.1}$$

where f is a given function on $\mathbb{R}^n$. By Fourier transforming Eq. (45.1) in the x – variables only, one finds that (45.1) implies that

$$\left(\partial_t+\frac{1}{2}|\xi|^2\right)\hat{u}(t,\xi)=0 \text{ with } \hat{u}(0,\xi)=\hat{f}(\xi). \tag{45.2}$$

and hence that $\hat{u}(t,\xi)=e^{-t|\xi|^2/2}\hat{f}(\xi)$. Inverting the Fourier transform then shows that

$$u(t,x)=\mathcal{F}^{-1}\left(e^{-t|\xi|^2/2}\hat{f}(\xi)\right)(x)=\left(\mathcal{F}^{-1}\left(e^{-t|\xi|^2/2}\right)\bigstar f\right)(x)=:e^{t\Delta/2}f(x).$$

From Example 32.4,

$$\mathcal{F}^{-1}\left(e^{-t|\xi|^2/2}\right)(x)=p_t(x)=t^{-n/2}e^{-\frac{1}{2t}|x|^2}$$

and therefore,

$$u(t,x)=\int_{\mathbb{R}^n}p_t(x-y)f(y)\mathbf{d}y.$$

This suggests the following theorem.

Theorem 45.1. *Let*

$$p_t(x-y):=(2\pi t)^{-n/2}e^{-|x-y|^2/2t} \tag{45.3}$$

be the ***heat kernel*** *on* $\mathbb{R}^n$. *Then*

$$\left(\partial_t-\frac{1}{2}\Delta_x\right)p_t(x-y)=0 \text{ and } \lim_{t\downarrow 0}p_t(x-y)=\delta_x(y), \tag{45.4}$$

where δ_x is the δ – function at x in $\mathbb{R}^n$. More precisely, if f is a continuous bounded function on $\mathbb{R}^n$, then

$$u(t,x)=\int_{\mathbb{R}^n}p_t(x-y)f(y)dy$$

is a solution to Eq. (45.1) where $u(0,x):=\lim_{t\downarrow 0}u(t,x)$.

Proof. Direct computations show that $\left(\partial_t-\frac{1}{2}\Delta_x\right)p_t(x-y)=0$ and an application of Theorem 11.21 shows $\lim_{t\downarrow 0}p_t(x-y)=\delta_x(y)$ or equivalently that $\lim_{t\downarrow 0}\int_{\mathbb{R}^n}p_t(x-y)f(y)dy=f(x)$ uniformly on compact subsets of $\mathbb{R}^n$. This shows that $\lim_{t\downarrow 0}u(t,x)=f(x)$ uniformly on compact subsets of $\mathbb{R}^n$. ∎

Proposition 45.2 (Properties of $e^{t\Delta/2}$).

1. *For $f\in L^2(\mathbb{R}^n,dx)$, the function*
$$\left(e^{t\Delta/2}f\right)(x)=(P_tf)(x)=\int_{\mathbb{R}^n}f(y)\frac{e^{-\frac{1}{2t}|x-y|^2}}{(2\pi t)^{n/2}}dy$$
is smooth in (t,x) for $t>0$ and $x\in\mathbb{R}^n$ and is in fact real analytic.
2. *$e^{t\Delta/2}$ acts as a contraction on $L^p(\mathbb{R}^n,dx)$ for all $p\in[0,\infty]$ and $t>0$. Indeed,*
3. *Moreover, $p_t*f\to f$ in L^p as $t\to 0$.*

Proof. Item 1. is fairly easy to check and is left the reader. One just notices that $p_t(x-y)$ analytically continues to $\operatorname{Re}t>0$ and $x\in\mathbb{C}^n$ and then shows that it is permissible to differentiate under the integral.

Item 2.

$$|(p_t*f)(x)|\le\int_{\mathbb{R}^n}|f(y)|p_t(x-y)dy$$

and hence with the aid of Jensen's inequality we have,

$$\|p_t*f\|_{L^p}^p\le\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}|f(y)|^pp_t(x-y)dydx=\|f\|_{L^p}^p$$

So P_t is a contraction $\forall t>0$.

Item 3. It suffices to show, because of the contractive properties of p_t*, that $p_t*f\to f$ as $t\downarrow 0$ for $f\in C_c(\mathbb{R}^n)$. Notice that if f has support in the ball of radius R centered at zero, then

$$\begin{aligned}|(p_t*f)(x)|&\le\int_{\mathbb{R}^n}|f(y)|P_t(x-y)dy\le\|f\|_\infty\int_{|y|\le R}P_t(x-y)dy\\&=\|f\|_\infty CR^ne^{-\frac{1}{2t}(|x|-R)^2}\end{aligned}$$

and hence

$$\|p_t*f-f\|_{L^p}^p=\int_{|y|\le R}|p_t*f-f|^pdy+\|f\|_\infty CR^ne^{-\frac{1}{2t}(|x|-R)^2}.$$

Therefore $p_t*f\to f$ in L^p as $t\downarrow 0$ $\forall f\in C_c(\mathbb{R}^n)$. ∎

Theorem 45.3 (Forced Heat Equation). *Suppose* $g \in C_b(\mathbb{R}^d)$ *and* $f \in C_b^{1,2}([0,\infty) \times \mathbb{R}^d)$ *then*

$$u(t,x) := p_t * g(x) + \int_0^t p_{t-\tau} * f(\tau, x)d\tau$$

solves

$$\frac{\partial u}{\partial t} = \frac{1}{2}\triangle u + f \text{ with } u(0,\cdot) = g.$$

Proof. Because of Theorem 45.1, we may with out loss of generality assume $g = 0$ in which case

$$u(t,x) = \int_0^t p_t * f(t-\tau, x)d\tau.$$

Therefore

$$\begin{aligned}\frac{\partial u}{\partial t}(t,x) &= p_t * f(0,x) + \int_0^t p_\tau * \frac{\partial}{\partial t} f(t-\tau,x)d\tau \\ &= p_t * f_0(x) - \int_0^t p_\tau * \frac{\partial}{\partial \tau} f(t-\tau,x)d\tau\end{aligned}$$

and

$$\frac{\triangle}{2}u(t,x) = \int_0^t p_t * \frac{\triangle}{2} f(t-\tau,x)d\tau.$$

Hence we find, using integration by parts and approximate δ – function arguments, that

$$\begin{aligned}\left(\frac{\partial}{\partial t} - \frac{\triangle}{2}\right)u(t,x) &= p_t * f_0(x) + \int_0^t p_\tau * \left(-\frac{\partial}{\partial \tau} - \frac{1}{2}\triangle\right) f(t-\tau,x)d\tau \\ &= p_t * f_0(x) \\ &\quad + \lim_{\epsilon \downarrow 0} \int_\epsilon^t p_\tau * \left(-\frac{\partial}{\partial \tau} - \frac{1}{2}\triangle\right) f(t-\tau,x)d\tau \\ &= p_t * f_0(x) - \lim_{\epsilon \downarrow 0} p_\tau * f(t-\tau,x)\Big|_\epsilon^t \\ &\quad + \lim_{\epsilon \downarrow 0} \int_\epsilon^t \left(\frac{\partial}{\partial \tau} - \frac{1}{2}\triangle\right) p_\tau * f(t-\tau,x)d\tau \\ &= p_t * f_0(x) - p_t * f_0(x) + \lim_{\epsilon \downarrow 0} p_\epsilon * f(t-\epsilon,x) \\ &= f(t,x).\end{aligned}$$

■

45.1 Extensions of Theorem 45.1

Proposition 45.4. *Suppose* $f : \mathbb{R}^n \to \mathbb{R}$ *is a measurable function and there exists constants* $c, C < \infty$ *such that*

$$|f(x)| \le Ce^{\frac{c}{2}|x|^2}.$$

Then $u(t,x) := p_t * f(x)$ *is smooth for* $(t,x) \in (0, c^{-1}) \times \mathbb{R}^n$ *and for all* $k \in \mathbb{N}$ *and all multi-indices* α,

$$D^\alpha \left(\frac{\partial}{\partial t}\right)^k u(t,x) = \left(D^\alpha \left(\frac{\partial}{\partial t}\right)^k p_t\right) * f(x). \tag{45.5}$$

In particular u *satisfies the heat equation* $u_t = \Delta u/2$ *on* $(0, c^{-1}) \times \mathbb{R}^n$.

Proof. The reader may check that

$$D^\alpha \left(\frac{\partial}{\partial t}\right)^k p_t(x) = q(t^{-1}, x)p_t(x)$$

where q is a polynomial in its variables. Let $x_0 \in \mathbb{R}^n$ and $\epsilon > 0$ be small, then for $x \in B(x_0, \epsilon)$ and any $\beta > 0$,

$$\begin{aligned}|x-y|^2 &= |x|^2 - 2|x||y| + |y|^2 \ge |y|^2 + |x|^2 - \left(\beta^{-2}|x|^2 + \beta^2|y|^2\right) \\ &\ge \left(1-\beta^2\right)|y|^2 - \left(\beta^{-2} - 1\right)\left(|x_0|^2 + \epsilon\right).\end{aligned}$$

Hence

$$\begin{aligned}g(y) &:= \sup\left\{\left|D^\alpha \left(\frac{\partial}{\partial t}\right)^k p_t(x-y)f(y)\right| : \epsilon \le t \le c-\epsilon \;\&\; x \in B(x_0,\epsilon)\right\} \\ &\le \sup\left\{\left|q(t^{-1}, x-y)\frac{e^{-\frac{1}{2t}|x-y|^2}}{(2\pi t)^{n/2}} Ce^{\frac{c}{2}|y|^2}\right| : \epsilon \le t \le c-\epsilon \;\&\; x \in B(x_0,\epsilon)\right\} \\ &\le C(\beta, x_0, \epsilon) \sup\left\{\left|q(t^{-1}, x-y)\frac{e^{\left[-\frac{1}{2t}(1-\beta^2)+\frac{c}{2}\right]|y|^2}}{(2\pi t)^{n/2}}\right| : \begin{array}{c}\epsilon \le t \le c-\epsilon \text{ and} \\ x \in B(x_0,\epsilon)\end{array}\right\}.\end{aligned}$$

By choosing β close to 0, the reader should check using the above expression that for any $0 < \delta < (1/t - c)/2$ there is a $\tilde{C} < \infty$ such that $g(y) \le \tilde{C}e^{-\delta|y|^2}$. In particular $g \in L^1(\mathbb{R}^n)$. Hence one is justified in differentiating past the integrals in $p_t * f$ and this proves Eq. (45.5). ■

Lemma 45.5. *There exists a polynomial* $q_n(x)$ *such that for any* $\beta > 0$ *and* $\delta > 0$,

$$\int_{\mathbb{R}^n} 1_{|y| \ge \delta} e^{-\beta|y|^2} dy \le \delta^n q_n\left(\frac{1}{\beta\delta^2}\right) e^{-\beta\delta^2}$$

Proof. Making the change of variables $y \to \delta y$ and then passing to polar coordinates shows

$$\int_{\mathbb{R}^n} 1_{|y|\ge\delta} e^{-\beta|y|^2} dy = \delta^n \int_{\mathbb{R}^n} 1_{|y|\ge 1} e^{-\beta\delta^2|y|^2} dy = \sigma\left(S^{n-1}\right)\delta^n \int_1^\infty e^{-\beta\delta^2 r^2} r^{n-1} dr.$$

Letting $\lambda = \beta\delta^2$ and $\phi_n(\lambda) := \int_{r=1}^\infty e^{-\lambda r^2} r^n dr$, integration by parts shows

$$\begin{aligned}\phi_n(\lambda) &= \int_{r=1}^\infty r^{n-1} d\left(\frac{e^{-\lambda r^2}}{-2\lambda}\right) = \frac{1}{2\lambda}e^{-\lambda} + \frac{1}{2}\int_{r=1}^\infty (n-1) r^{(n-2)} \frac{e^{-\lambda r^2}}{\lambda} dr \\ &= \frac{1}{2\lambda}e^{-\lambda} + \frac{n-1}{2\lambda}\phi_{n-2}(\lambda).\end{aligned}$$

Iterating this equation implies

$$\phi_n(\lambda) = \frac{1}{2\lambda}e^{-\lambda} + \frac{n-1}{2\lambda}\left(\frac{1}{2\lambda}e^{-\lambda} + \frac{n-3}{2\lambda}\phi_{n-4}(\lambda)\right)$$

and continuing in this way shows

$$\phi_n(\lambda) = e^{-\lambda} r_n(\lambda^{-1}) + \frac{(n-1)!!}{2^\delta \lambda^\delta}\phi_i(\lambda)$$

where δ is the integer part of $n/2$, $i = 0$ if n is even and $i = 1$ if n is odd and r_n is a polynomial. Since

$$\phi_0(\lambda) = \int_{r=1}^\infty e^{-\lambda r^2} dr \le \phi_1(\lambda) = \int_{r=1}^\infty r e^{-\lambda r^2} dr = \frac{e^{-\lambda}}{2\lambda},$$

it follows that

$$\phi_n(\lambda) \le e^{-\lambda} q_n(\lambda^{-1})$$

for some polynomial q_n. ■

Proposition 45.6. *Suppose $f \in C(\mathbb{R}^n, \mathbb{R})$ such that $|f(x)| \le Ce^{\frac{c}{2}|x|^2}$ then $p_t * f \to f$ uniformly on compact subsets as $t \downarrow 0$. In particular in view of Proposition 45.4, $u(t,x) := p_t * f(x)$ is a solution to the heat equation with $u(0,x) = f(x)$.*

Proof. Let $M > 0$ be fixed and assume $|x| \le M$ throughout. By uniform continuity of f on compact set, given $\epsilon > 0$ there exists $\delta = \delta(t) > 0$ such that $|f(x) - f(y)| \le \epsilon$ if $|x - y| \le \delta$ and $|x| \le M$. Therefore, choosing $a > c/2$ sufficiently small,

$$\begin{aligned}|p_t * f(x) - f(x)| &= \left|\int p_t(y)\left[f(x-y) - f(x)\right] dy\right| \\ &\le \int p_t(y)\left|f(x-y) - f(x)\right| dy \\ &\le \epsilon \int_{|y|\le\delta} p_t(y) dy + \frac{C}{(2\pi t)^{n/2}} \int_{|y|\ge\delta} [e^{\frac{c}{2}|x-y|^2} + e^{\frac{c}{2}|x|^2}] e^{-\frac{1}{2t}|y|^2} dy \\ &\le \epsilon + \tilde{C}(2\pi t)^{-n/2} \int_{|y|\ge\delta} e^{-\left(\frac{1}{2t}-a\right)|y|^2} dy.\end{aligned}$$

So by Lemma 45.5, it follows that

$$|p_t * f(x) - f(x)| \le \epsilon + \tilde{C}(2\pi t)^{-n/2} \delta^n q_n\left(\frac{1}{\beta\left(\frac{1}{2t} - a\right)^2}\right) e^{-\left(\frac{1}{2t}-a\right)\delta^2}$$

and therefore

$$\limsup_{t\downarrow 0} \sup_{|x|\le M} |p_t * f(x) - f(x)| \le \epsilon \to 0 \text{ as } \epsilon \downarrow 0.$$

■

Lemma 45.7. *If $q(x)$ is a polynomial on $\mathbb{R}^n$, then*

$$\int_{\mathbb{R}^n} p_t(x-y) q(y) dy = \sum_{n=0}^\infty \frac{t^n}{n!}\frac{\Delta^n}{2^n} q(x).$$

Proof. Since

$$f(t,x) := \int_{\mathbb{R}^n} p_t(x-y) q(y) dy = \int_{\mathbb{R}^n} p_t(y) \sum a_{\alpha\beta} x^\alpha y^\beta dy = \sum C_\alpha(t) x^\alpha,$$

$f(t,x)$ is a polynomial in x of degree no larger than that of q. Moreover $f(t,x)$ solves the heat equation and $f(t,x) \to q(x)$ as $t \downarrow 0$. Since $g(t,x) := \sum_{n=0}^\infty \frac{t^n}{n!}\frac{\Delta^n}{2^n} q(x)$ has the same properties of f and Δ is a bounded operator when acting on polynomials of a fixed degree we conclude $f(t,x) = g(t,x)$. ■

Example 45.8. Suppose $q(x) = x_1 x_2 + x_3^4$, then

$$\begin{aligned}e^{t\Delta/2} q(x) &= x_1 x_2 + x_3^4 + \frac{t}{2}\Delta\left(x_1 x_2 + x_3^4\right) + \frac{t^2}{2!\cdot 4}\Delta^2\left(x_1 x_2 + x_3^4\right) \\ &= x_1 x_2 + x_3^4 + \frac{t}{2} 12 x_3^2 + \frac{t^2}{2!\cdot 4} 4! \\ &= x_1 x_2 + x_3^4 + 6 t x_3^2 + 3t^2.\end{aligned}$$

Proposition 45.9. *Suppose $f \in C^\infty(\mathbb{R}^n)$ and there exists a constant $C < \infty$ such that*

$$\sum_{|\alpha| = 2N+2} |D^\alpha f(x)| \le Ce^{C|x|^2},$$

then

$$(p_t * f)(x) = \text{“}e^{t\Delta/2} f(x)\text{”} = \sum_{k=0}^N \frac{t^k}{k!}\Delta^k f(x) + O(t^{N+1}) \text{ as } t \downarrow 0$$

Proof. Fix $x \in \mathbb{R}^n$ and let

$$f_N(y) := \sum_{|\alpha|\le 2N+1} \frac{1}{\alpha!} D^\alpha f(x) y^\alpha.$$

Then by Taylor's theorem with remainder

$$|f(x+y)-f_N(y)| \le C\,|y|^{2N+2} \sup_{t\in[0,1]} e^{C|x+ty|^2} \le C\,|y|^{2N+2} e^{2C\left[|x|^2+|y|^2\right]} \le \tilde{C}\,|y|^{2N+2} e^{2C|y|^2}$$

and thus

$$\left|\int_{\mathbb{R}^n} p_t(y)f(x+y)dy - \int_{\mathbb{R}^n} p_t(y)f_N(y)dy\right| \le \tilde{C}\int_{\mathbb{R}^n} p_t(y)\,|y|^{2N+2} e^{2C|y|^2}dy = \tilde{C}t^{N+1}\int_{\mathbb{R}^n} p_1(y)\,|y|^{2N+2} e^{2t^2C|y|^2}dy = O(t^{N+1}).$$

Since $f(x+y)$ and $f_N(y)$ agree to order $2N+1$ for y near zero, it follows that

$$\int_{\mathbb{R}^n} p_t(y)f_N(y)dy = \sum_{k=0}^{N}\frac{t^k}{k!}\Delta^k f_N(0) = \sum_{k=0}^{N}\frac{t^k}{k!}\Delta_y^k f(x+y)|_{y=0} = \sum_{k=0}^{N}\frac{t^k}{k!}\Delta^k f(x)$$

which completes the proof. ■

45.2 Representation Theorem and Regularity

In this section, suppose that Ω is a bounded domain such that $\bar{\Omega}$ is a C^2 – submanifold with C^2 boundary and for $T>0$ let $\Omega_T := (0,T)\times\Omega$, and

$$\Gamma_T := ([0,T]\times\partial\Omega)\cup(\{0\}\times\Omega)\subset \mathrm{bd}(\Omega_T) = ([0,T]\times\partial\Omega)\cup(\{0,T\}\times\Omega)$$

as in Figure 45.1 below.

Theorem 45.10 (Representation Theorem). *Suppose* $u \in C^{2,1}(\bar{\Omega}_T)$ $(\bar{\Omega}_T = \bar{\Omega}_T = [0,T]\times\overline{\Omega})$ *solves* $u_t = \frac{1}{2}\triangle u + f$ *on* $\bar{\Omega}_T$. *Then*

$$u(T,x) = \int_{\Omega} p_T(x-y)u(0,y)dy + \int_{[0,T]\times\Omega} p_{T-t}(x-y)f(t,y)dydt + \frac{1}{2}\int_{[0,T]\times\partial\Omega}\left[\frac{\partial p_{T-t}}{\partial n_y}(x-y)u(t,y) - p_{T-t}(x-y)\frac{\partial u}{\partial n}(y)\right]d\sigma(y)dt \tag{45.6}$$

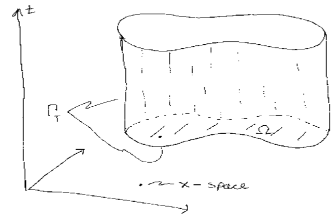


Fig. 45.1. A cylindrical region Ω_T and the parabolic boundary Γ_T.

Proof. For $v\in C^{2,1}([0,T]\times\overline{\Omega})$, integration by parts shows

$$\begin{aligned}\int_{\Omega_T} fvdydt &= \int_{\Omega_T} v(u_t - \frac{1}{2}\triangle v)dydt \\ &= \int_{\Omega_T}(-v_t + \frac{1}{2}\nabla v\cdot\nabla u)dydt + \int_{\Omega} vu\Big|_{t=0}^{t=T}dy + \frac{1}{2}\int_{[0,T]\times\partial\Omega} v\frac{\partial v}{\partial n}dtd\sigma \\ &= \int_{\Omega_T}(-v_t - \frac{1}{2}\triangle v)udydt + \int_{\Omega} vu\Big|_0^T dy + \frac{1}{2}\int_{[0,T]\times\partial\Omega}\left(\frac{\partial u}{\partial n}u - v\frac{\partial u}{\partial n}\right)d\sigma\; dt.\end{aligned}$$

Given $\epsilon>0$, taking $v(t,y) := p_{T+\epsilon-t}(x-y)$ (note that $v_t + \frac{1}{2}\triangle v = 0$ and $v\in C^{2,1}([0,T]\times\Omega)$) implies

$$\int_{[0,T]\times\Omega} f(t,y)p_{T+\epsilon-t}(x-y)dydt = 0 + \int_{\Omega} p_\epsilon(x-y)u(t,y)dy - \int_{\Omega} p_{T+\epsilon}(x-y)u(t,y)dy + \frac{1}{2}\int_{[0,T]\times\partial\Omega}\left[\begin{array}{c}\frac{\partial p_{T+\epsilon-t}(x-y)}{\partial n_y}u(t,y) \\ -p_{T+\epsilon-t}(x-y)\frac{\partial u}{\partial n}(y)\end{array}\right]d\sigma(y)dt$$

Let $\epsilon\downarrow 0$ above to complete the proof. ■

Corollary 45.11. *Suppose* $f := 0$ *so* $u_t(t,x) = \frac{1}{2}\Delta u(t,x)$. *Then* $u\in C^\infty((0,T)\times\Omega)$.

Proof. Extend $p_t(x)$ for $t \le 0$ by setting $p_t(x) := 0$ if $t \le 0$. It is not to hard to check that this extension is C^∞ on $\mathbb{R} \times \mathbb{R}^n \backslash \{0\}$. Using this notation we may write Eq. (45.6) as

$$u(t,x) = \int_\Omega p_t(x-y)u(0,y)dy$$
$$+ \frac{1}{2} \int_{[0,\infty)\times\partial\Omega} \left[\frac{\partial p_{t-\tau}}{\partial n_y}(x-y)u(t,y) - p_{T-t}(x-y)\frac{\partial u}{\partial n}(y)\right] d\sigma(y)d\tau.$$

The result follows since now it permissible to differentiate under the integral to show $u \in C^\infty((0,T) \times \Omega)$. ■

Remark 45.12. Since $x \to p_t(x)$ is analytic one may show that $x \to u(t,x)$ is analytic for all $x \in \Omega$.

45.3 Weak Max Principles

Notation 45.13 *Let $a_{ij}, b_j \in C\left(\bar{\Omega}_T\right)$ satisfy $a_{ij} = a_{ji}$ and for $u \in C^2(\Omega)$ let*

$$Lu(t,x) = \sum_{i,j=1}^n a_{ij}(t,x)u_{x_i x_j}(x) + \sum_{i=1}^n b_i(t,x)u_{x_i}(x). \tag{45.7}$$

*We say L is **elliptic** if there exists $\theta > 0$ such that*

$$\sum a_{ij}(t,x)\xi_i\xi_j \ge \theta|\xi|^2 \text{ for all } \xi \in \mathbb{R}^n \text{ and } (t,x) \in \bar{\Omega}_T.$$

Assumption 3 *In this section we assume L is elliptic. As an example $L = \frac{1}{2}\Delta$ is elliptic.*

Lemma 45.14. *Let L be an elliptic operator as above and suppose $u \in C^2(\Omega)$ and $x_0 \in \Omega$ is a point where $u(x)$ has a local maximum. Then $Lu(t,x_0) \le 0$ for all $t \in [0,T]$.*

Proof. Fix $t \in [0,T]$ and set $B_{ij} = u_{x_i x_j}(x_0)$, $A_{ij} := a_{ij}(t,x_0)$ and let $\{e_i\}_{i=1}^n$ be an orthonormal basis for $\mathbb{R}^n$ such that $Ae_i = \lambda_i e_i$. Notice that $\lambda_i \ge \theta > 0$ for all i. By the first derivative test, $u_{x_i}(x_0) = 0$ for all i and hence

$$\begin{aligned} Lu(t,x_0) &= \sum A_{ij}B_{ij} = \sum A_{ji}B_{ij} = \operatorname{tr}(AB) \\ &= \sum e_i \cdot ABe_i = \sum Ae_i \cdot Be_i = \sum_i \lambda_i e_i \cdot Be_i \\ &= \sum_i \lambda_i \partial_{e_i}^2 u(t,x_0) \le 0. \end{aligned}$$

The last inequality if a consequence of the second derivative test which asserts $\partial_v^2 u(t,x_0) \le 0$ for all $v \in \mathbb{R}^n$. ■

Theorem 45.15 (Elliptic weak maximum principle). *Let Ω be a bounded domain and L be an elliptic operator as in Eq. (45.7). We now assume that a_{ij} and b_j are functions of x alone. For each $u \in C\left(\bar{\Omega}\right) \cap C^2(\Omega)$ such that $Lu \ge 0$ on Ω (i.e. u is L – subharmonic) we have*

$$\max_{\bar{\Omega}} u \le \max_{\mathrm{bd}(\Omega)} u. \tag{45.8}$$

Proof. Let us first assume $Lu > 0$ on Ω. If u and had an interior local maximum at $x_0 \in \Omega$ then by Lemma 45.14, $Lu(x_0) \le 0$ which contradicts the assumption that $Lu(x_0) > 0$. So if $Lu > 0$ on Ω we conclude that Eq. (45.8) holds.

Now suppose that $Lu \ge 0$ on Ω. Let $\phi(x) := e^{\lambda x_1}$ with $\lambda > 0$, then

$$L\phi(x) = \left(\lambda^2 a_{11}(x) + b_1(x)\lambda\right)e^{\lambda x_1} \ge \lambda\left(\lambda\theta + b_1(x)\right)e^{\lambda x_1}.$$

By continuity of $b(x)$ we may choose λ sufficiently large so that $\lambda\theta + b_1(x) > 0$ on $\bar{\Omega}$ in which case $L\phi > 0$ on Ω. The results in the first paragraph may now be applied to $u_\epsilon(x) := u(x) + \epsilon\phi(x)$ (for any $\epsilon > 0$) to learn

$$u(x) + \epsilon\phi(x) = u_\epsilon(x) \le \max_{\mathrm{bd}(\Omega)} u_\epsilon \le \max_{\mathrm{bd}(\Omega)} u + \epsilon \max_{\mathrm{bd}(\Omega)} \phi \text{ for all } x \in \bar{\Omega}.$$

Letting $\epsilon \downarrow 0$ in this expression then implies

$$u(x) \le \max_{\mathrm{bd}(\Omega)} u \text{ for all } x \in \bar{\Omega}$$

which is equivalent to Eq. (45.8). ■

Theorem 45.16 (Parabolic weak maximum principle). *Assume $u \in C^{1,2}(\overline{\Omega}_T \backslash \Gamma_T) \cap C(\overline{\Omega}_T)$.*

1. If $u_t - Lu \le 0$ in Ω_T then

$$\max_{\overline{\Omega}_T} u = \max_{\Gamma_T} u. \tag{45.9}$$

2. If $u_t - Lu \ge 0$ in Ω_T then $\min_{\overline{\Omega}_T} u = \min_{\Gamma_T} u$.

Proof. Item 1. follows from Item 2. by replacing $u \to -u$, so it suffices to prove item 1. We begin by assuming $u_t - Lu < 0$ on $\bar{\Omega}_T$ and suppose for the sake of contradiction that there exists a point $(t_0, x_0) \in \overline{\Omega}_T \backslash \Gamma_T$ such that $u(t_0,x_0) = \max_{\overline{\Omega}_T} u$.

1. If $(t_0,x_0) \in \Omega_T$ (i.e. $0 < t_0 < T$) then by the first derivative test $\frac{\partial u}{\partial t}(t_0,x_0) = 0$ and by Lemma 45.14 $Lu(t_0,x_0) \le 0$. Therefore,
$$\left(u_t - Lu\right)(t_0,x_0) = -Lu(t_0,x_0) \ge 0$$
which contradicts the assumption that $u_t - Lu < 0$ in Ω_T.

2. If $(t_0, x_0) \in \overline{\Omega}_T \backslash \Gamma_T$ with $t_0 = T$, then by the first derivative test, $\frac{\partial u}{\partial t}(T, x_0) \geq 0$ and by Lemma 45.14 $Lu(t_0, x_0) \leq 0$. So again

$$(u_t - Lu)(t_0, x_0) \geq 0$$

which contradicts the assumption that $u_t - Lu < 0$ in Ω_T.

Thus we have proved Eq. (45.9) holds if $u_t - Lu < 0$ on $\bar{\Omega}_T$. Finally if $u_t - Lu \leq 0$ on $\bar{\Omega}_T$ and $\epsilon > 0$, the function $u^\epsilon(t,x) := u(t,x) - \epsilon t$ satisfies $u^\epsilon_t - Lu^\epsilon \leq -\epsilon < 0$. Therefore by what we have just proved

$$u(t,x) - \epsilon t \leq \max_{\bar{\Omega}_T} u^\epsilon = \max_{\Gamma_T} u^\epsilon \leq \max_{\Gamma_T} u \text{ for all } (t,x) \in \bar{\Omega}_T.$$

Letting $\epsilon \downarrow 0$ in the last equation shows that Eq. (45.9) holds. ■

Corollary 45.17. *There is at most one solution* $u \in C^{1,2}(\overline{\Omega}_T \backslash \Gamma_T) \cap C(\overline{\Omega}_T)$ *to the partial differential equation*

$$\frac{\partial u}{\partial t} = Lu \text{ with } u = f \text{ on } \Gamma_T.$$

Proof. If there were another solution v, then $w := u - v$ would solve $\frac{\partial w}{\partial t} = Lw$ with $w = 0$ on Γ_T. So by the maximum principle in Theorem 45.16, $w = 0$ on $\bar{\Omega}_T$. ■

We now restrict back to $L = \frac{1}{2}\triangle$ and we wish to see what can be said when $\Omega = \mathbb{R}^n$ – an unbounded set.

Theorem 45.18. *Suppose* $u \in C([0,T] \times \mathbb{R}^n) \cap C^{2,1}((0,T) \times \mathbb{R}^n)$,

$$u_t - \frac{1}{2}\triangle u \leq 0 \text{ on } [0,T] \times \mathbb{R}^n$$

and there exists constants $A, a < \infty$ *such that*

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

Then

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

Proof. Recall that

$$p_t(x) = \left(\frac{1}{t}\right)^{n/2} e^{-\frac{1}{2t}|x|^2} = \left(\frac{1}{t}\right)^{n/2} e^{-\frac{1}{2t}x \cdot x}$$

solves the heat equation

$$\partial_t p_t(x) = \frac{1}{2}\triangle p_t(x). \tag{45.10}$$

Since both sides of Eq. (45.10) are analytic as functions in x, so[1]

$$\frac{\partial p_t}{\partial t}(ix) = \frac{1}{2}(\triangle p_t)(ix) = -\frac{1}{2}\triangle_x p_t(ix)$$

and therefore for all $\tau > 0$ and $t < \tau$

$$\frac{\partial p_{\tau - t}}{\partial t}(ix) = -\dot{p}_{\tau-t}(ix) = \frac{1}{2}\triangle_x p_{\tau-t}(ix).$$

That is to say the function

$$\rho(t,x) := p_{\tau-t}(ix) = \left(\frac{1}{\tau - t}\right)^{n/2} e^{\frac{1}{2(\tau-t)}|x|^2} \text{ for } 0 \leq t < \tau$$

solves the heat equation. (This can be checked directly as well.)

Let $\epsilon, \tau > 0$ (to be chosen later) and set

$$v(t,x) = u(t,x) - \epsilon\rho(t,x) \text{ for } 0 \leq t \leq \tau/2.$$

Since $\rho(t,x)$ is increasing in t,

$$v(t,x) \leq Ae^{a|x|^2} - \epsilon\left(\frac{1}{\tau}\right)^{n/2} e^{\frac{1}{2\tau}|x|^2} \text{ for } 0 \leq t \leq \tau/2.$$

Hence if we require $\frac{1}{2\tau} > a$ or $\tau < \frac{1}{2a}$ it will follows that

$$\lim_{|x| \to \infty}\left[\sup_{0 \leq t \leq \tau/2} v(t,x)\right] = -\infty.$$

Therefore we may choose M sufficiently large so that

$$v(t,x) \leq K := \sup_z u(0,z) \text{ for all } |x| \geq M \text{ and } 0 \leq t \leq \tau/2.$$

Since

$$\left(\partial_t - \frac{\triangle}{2}\right) v = \left(\partial_t - \frac{\triangle}{2}\right) u \leq 0$$

we may apply the maximum principle with $\Omega = B(0,M)$ and $T = \tau/2$ to conclude for $(t,x) \in \Omega_T$ that

$$u(t,x) - \epsilon\rho(t,x) = v(t,x) \leq \sup_{z \in \Omega} v(0,z) \leq K \text{ if } 0 \leq t \leq \tau/2.$$

[1] Similarly since both sides of Eq. (45.10) are analytic functions in t, it follows that

$$\frac{\partial}{\partial t}p_{-t}(x) = -\dot{p}_t(x) = -\frac{1}{2}\triangle p_{-t}.$$

We may now let $\epsilon \downarrow 0$ in this equation to conclude that

$$u(t,x) \le K \text{ if } 0 \le t \le \tau/2. \tag{45.11}$$

By applying Eq. (45.11) to $u(t+\tau/2, x)$ we may also conclude

$$u(t,x) \le K \text{ if } 0 \le t \le \tau.$$

Repeating this argument then enables us to show $u(t,x) \le K$ for all $0 \le t \le T$. ■

Corollary 45.19. *The heat equation*

$$u_t - \frac{1}{2}\triangle u = 0 \text{ on } [0,T] \times \mathbb{R}^n \text{ with } u(0,\cdot) = f(\cdot) \in C(\mathbb{R}^n)$$

has at most one solution in the class of functions $u \in C([0,T] \times \mathbb{R}^n) \cap C^{2,1}((0,T) \times \mathbb{R}^n)$ *which satisfy*

$$u(t,x) \le Ae^{a|x|^2} \text{ for } (t,x) \in (0,T) \times \mathbb{R}^n$$

for some constants A *and* a.

Theorem 45.20 (Max Principle a la Hamilton). *Suppose* $u \in C^{1,2}\left([0,T] \times \mathbb{R}\right.$

1. $u(t,x) \le Ae^{a|x|^2}$ *for some* A, a *(for all* $t \le T$*)*
2. $u(0,x) \le 0$ *for all* x
3. $\frac{\partial u}{\partial t} \le \Delta u$ *i.e.* $(\partial_t - \triangle)u \le 0$.

Then $u(t,x) \le 0$ *for all* $(t,x) \in [0,T] \times \mathbb{R}^d$.

Proof. Special Case. Assume $\frac{\partial u}{\partial t} < \Delta u$ on $[0,T] \times \mathbb{R}^d$, $u(0,x) < 0$ for all $x \in \mathbb{R}^d$ and there exists $M > 0$ such that $u(t,x) < 0$ if $|x| \ge M$ and $t \in [0,T]$.For the sake of contradiction suppose there is some point $(t,x) \in [0,T] \times \mathbb{R}^d$ such that $u(t,x) > 0$.

By the intermediate value theorem there exists $\tau \in [0,t]$ such that $u(\tau,x) = 0$. In particular the set $\{u = 0\}$ is a non-empty closed compact subset of $(0,T] \times B(0,M)$. Let

$$\pi : (0,T] \times B(0,M) \to (0,T]$$

be projection onto the first factor, since $\{u \neq 0\}$ is a compact subset of $(0,T] \times B(0,M)$ if follows that

$$t_0 := \min\{t \in \pi\left(\{u = 0\}\right)\} > 0.$$

Choose a point $x_0 \in B(0,M)$ such that $(t_0, x_0) \in \{u = 0\}$, i.e. $u(t_0,x_0) = 0$, see Figure 45.2 below. Since $u(t,x) < 0$ for all $0 \le t < t_0$ and $x \in \mathbb{R}^d$,

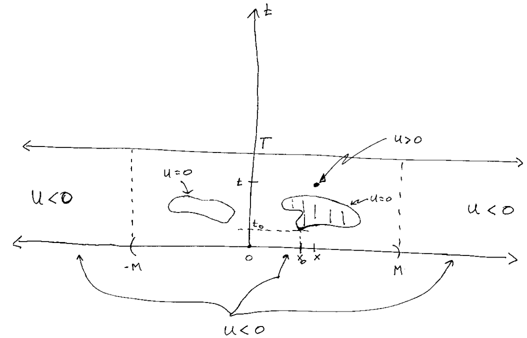


Fig. 45.2. Finding a point (t_0, x_0) such that t_0 is as small as possible and $u(t_0, x_0) = 0$.

$u(t_0, x) \le 0$ for all $x \in \mathbb{R}^d$ with $u(t_0, x_0) = 0$. This information along with the first and second derivative tests allows us to conclude

$$\nabla u(t_0, x_0) = 0, \; \triangle u(t_0, x_0) \le 0 \text{ and } \frac{\partial u}{\partial t}(t_0, x_0) \ge 0.$$

This then implies that

$$0 \le \frac{\partial u}{\partial t}(t_0, x_0) < \triangle u(t_0, x_0) \le 0$$

which is absurd. Hence we conclude that $u \le 0$ on $[0,T] \times \mathbb{R}^d$.

General Case: Let $p_t(x) = \frac{1}{t^{d/2}} e^{-\frac{1}{4t}|x|^2}$ be the fundamental solution to the heat equation

$$\partial_t p_t = \triangle p_t.$$

Let $\tau > 0$ to be determined later. As in the proof of Theorem 45.18, the function

$$\rho(t,x) := p_{\tau - t}(ix) = \left(\frac{1}{\tau - t}\right)^{d/2} e^{\frac{1}{4(\tau - t)}|x|^2} \text{ for } 0 \le t < \tau$$

is still a solution to the heat equation. Given $\epsilon > 0$, define, for $t \le \tau/2$,

$$u_\epsilon(t,x) = u(t,x) - \epsilon - \epsilon t - \epsilon\rho(t,x).$$

Then

$$(\partial_t - \triangle)u_\epsilon = (\partial_t - \triangle)u - \epsilon \le -\epsilon < 0,$$
$$u_\epsilon(0,x) = u(0,x) - \epsilon \le 0 - \epsilon \le -\epsilon < 0$$

and for $t \le \tau/2$

$$u_\epsilon(t,x) \le Ae^{a|x|^2} - \epsilon - \epsilon\frac{1}{\tau^{d/2}}e^{\frac{1}{4\tau}|x|^2}.$$

Hence if we choose τ such that $\frac{1}{4\tau} > a$, we will have $u_\epsilon(t,x) < 0$ for $|x|$ sufficiently large. Hence by the special case already proved, $u_\epsilon(t,x) \le 0$ for all $0 \le t \le \frac{\tau}{2}$ and $\epsilon > 0$. Letting $\epsilon \downarrow 0$ implies that $u(t,x) \le 0$ for all $0 \le t \le \tau/2$. As in the proof of Theorem 45.18 we may step our way up by applying the previous argument to $u(t+\tau/2, x)$ and then to $u(t+\tau, x)$, etc. to learn $u(t,x) \le 0$ for all $0 \le t \le T$. ■

45.4 Non-Uniqueness of solutions to the Heat Equation

Theorem 45.21 (See Fritz John §7). *For any $\alpha > 1$, let*

$$g(t) := \begin{cases} e^{-t^{-\alpha}} & t > 0 \\ 0 & t \le 0 \end{cases} \tag{45.12}$$

and define

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

Then $u \in C^\infty(\mathbb{R}^2)$ and

$$u_t = u_{xx} \text{ and } u(0,x) := 0. \tag{45.13}$$

In particular, the heat equation does not have unique solutions.

Proof. We are going to look for a solution to Eq. (45.13) of the form

$$u(t,x) = \sum_{n=0}^\infty g_n(t)x^n$$

in which case we have (formally) that

$$\begin{aligned} 0 = u_t - u_{xx} &= \sum_{n=0}^\infty (\dot{g}_n(t)x^n - g_n(t)n(n-1)x^{n-2}) \\ &= \sum_{n=0}^\infty \left[\dot{g}_n(t) - (n+2)(n+1)g_{n+2}(t)\right]x^n. \end{aligned}$$

This implies

$$g_{n+2} = \frac{\dot{g}_n}{(n+2)(n+1)}. \tag{45.14}$$

To simplify the final answer, we will now assume $u_x(0,x) = 0$, i.e. $g_1 \equiv 0$ in which case Eq. (45.14) implies $g_n \equiv 0$ for all n odd. We also have with $g := g_0$,

$$g_2 = \frac{\dot{g}_0}{2\cdot 1} = \frac{\dot{g}}{2!},\ g_4 = \frac{\dot{g}_2 0}{4\cdot 3} = \frac{\ddot{g}}{4!},\ g_6 = \frac{g^{(3)}}{6!} \ \dots\ g_{2k} = \frac{g^{(k)}}{(2k)!}$$

and hence

$$u(t,x) = \sum_{k=0}^\infty \frac{g^{(k)}(t)x^{2k}}{(2k)!}. \tag{45.15}$$

The function $u(t,x)$ will solve $u_t = u_{xx}$ for $(t,x) \in \mathbb{R}^2$ with $u(0,x) = 0$ provided the convergence in the sum is adequate to justify the above computations.

Now let $g(t)$ be given by Eq. (45.12) and extend g to $\mathbb{C}\setminus(-\infty,0]$ via $g(z) = e^{-z^{-\alpha}}$ where

$$z^{-\alpha} = e^{-\alpha\log(z)} = e^{-\alpha(\ln r + i\theta)} \text{ for } z = re^{i\theta} \text{ with } -\pi < \theta < \pi.$$

In order to estimate $g^{(k)}(t)$ we will use of the Cauchy estimates on the contour $|z-t| = \gamma t$ where γ is going to be chosen sufficiently close to 0. Now

$$\mathrm{Re}(z^{-\alpha}) = e^{-\alpha\ln r}\cos(\alpha\theta) = |z|^{-\alpha}\cos(\alpha\theta)$$

and hence

$$|g(z)| = e^{-\mathrm{Re}(z^{-\alpha})} = e^{-|z|^{-\alpha}\cos(\alpha\theta)}.$$

From Figure 45.3, we see

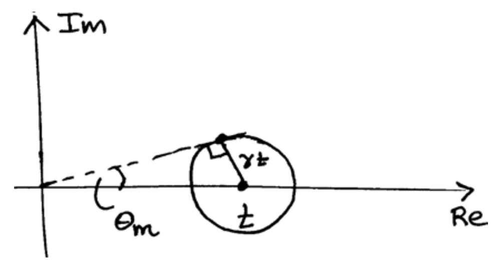


Fig. 45.3. Here is a picture of the maximum argument θ_m that a point z on $\partial B(t,\gamma t)$ may attain. Notice that $\sin\theta_m = \gamma t/t = \gamma$ is independent of t and $\theta_m \to 0$ as $\gamma \to 0$.

$$\beta(\gamma) := \min\left\{\cos(\alpha\theta) : -\pi < \theta < \pi \text{ and } |re^{i\theta} - t| = \gamma t\right\}$$

is independent of t and $\beta(\gamma) \to 1$ as $\gamma \to 0$. Therefore for $|z-t| = \gamma t$ we have

$$|g(z)| \le e^{-|z|^{-\alpha}\beta(\gamma)} \le e^{-([\gamma+1]t)^{-\alpha}\beta(\gamma)} = e^{-\frac{\beta(\gamma)}{1+\gamma}t^{-\alpha}} \le e^{-\frac{1}{2}t^{-\alpha}}$$

provided γ is chosen so small that $\frac{\beta(\gamma)}{1+\gamma} \ge \frac{1}{2}$.

By for $w \in B(t, t\gamma)$, the Cauchy integral formula and its derivative give

$$g(w) = \frac{1}{2\pi i}\oint_{|z-t|=\gamma t} \frac{g(z)}{z-w}\,dz \text{ and}$$
$$g^{(k)}(w) = \frac{k!}{2\pi i}\oint_{|z-t|=\gamma t} \frac{g(z)}{(z-w)^{k+1}}\,dz$$

and in particular

$$\begin{aligned}\left|g^{(k)}(t)\right| &\le \frac{k!}{2\pi}\oint_{|z-t|=\gamma t} \frac{|g(z)|}{|z-w|^{k+1}}\,|dz| \\ &\le \frac{k!}{2\pi} e^{-\frac{1}{2}t^{-\alpha}} \frac{2\pi\gamma t}{|\gamma t|^{k+1}} = \frac{k!}{|\gamma t|^{k}} e^{-\frac{1}{2}t^{-\alpha}}.\end{aligned} \tag{45.16}$$

We now use this to estimate the sum in Eq. (45.15) as

$$\begin{aligned}|u(t,x)| &\le \sum_{k=0}^{\infty}\left|\frac{g^{(k)}(t)x^{2k}}{(2k)!}\right| \le e^{-\frac{1}{2}t^{-\alpha}}\sum_{k=0}^{\infty}\frac{k!}{(\gamma t)^k}\frac{|x|^{2k}}{(2k)!} \\ &\le e^{-\frac{1}{2}t^{-\alpha}}\sum_{k=0}^{\infty}\frac{1}{k!}\left(\frac{x^2}{\gamma t}\right)^k = \exp\left(\frac{x^2}{\gamma t} - \frac{1}{2}t^{-\alpha}\right) < \infty.\end{aligned}$$

Therefore $\lim_{t\downarrow 0} u(t,x) = 0$ uniformly for x in compact subsets of $\mathbb{R}$. Similarly one may use the estimate in Eq. (45.16) to show u is smooth and

$$\begin{aligned}u_{xx} &= \sum_{k=0}^{\infty}\frac{g^{(k)}(t)(2k)(2k-1)x^{2k-2}}{(2k)!} = \sum_{k=1}^{\infty}\frac{g^{(k)}(t)x^{2(k-1)}}{(2(k-1))!} \\ &= \sum_{k=0}^{\infty}\frac{g^{(k+1)}(t)x^{2k}}{(2k)!} = u_t.\end{aligned}$$

■

45.5 The Heat Equation on the Circle and $\mathbb{R}$

In this subsection, let $S_L := \{Lz : z \in S\}$ – be the circle of radius L. As usual we will identify functions on S_L with $2\pi L$ – periodic functions on $\mathbb{R}$. Given two $2\pi L$ periodic functions f, g, let

$$(f,g)_L := \frac{1}{2\pi L}\int_{-\pi L}^{\pi L} f(x)\bar{g}(x)dx$$

and denote $H_L := L^2_{2\pi L}$ to be the $2\pi L$ – periodic functions f on $\mathbb{R}$ such that $(f,f)_L < \infty$. By Fourier's theorem we know that the functions $\chi_k^L(x) := e^{ikx/L}$ with $k \in \mathbb{Z}$ form an orthonormal basis for H_L and this basis satisfies

$$\frac{d^2}{dx^2}\chi_k^L = -\left(\frac{k}{L}\right)^2\chi_k^L.$$

Therefore the solution to the heat equation on S_L,

$$u_t = \frac{1}{2}u_{xx} \text{ with } u(0,\cdot) = f \in H_L$$

is given by

$$\begin{aligned}u(t,x) &= \sum_{k\in\mathbb{Z}}(f,\chi_k^L)e^{-\frac{1}{2}\left(\frac{k}{L}\right)^2 t}e^{ikx/L} \\ &= \sum_{k\in\mathbb{Z}}\left(\frac{1}{2\pi L}\int_{-\pi L}^{\pi L} f(y)e^{-iky/L}dy\right)e^{-\frac{1}{2}\left(\frac{k}{L}\right)^2 t}e^{ikx/L} \\ &= \int_{-\pi L}^{\pi L} p_t^L(x-y)f(y)dy\end{aligned}$$

where

$$p_t^L(x) = \frac{1}{2\pi L}\sum_{k\in\mathbb{Z}} e^{-\frac{1}{2}\left(\frac{k}{L}\right)^2 t}e^{ikx/L}.$$

If f is L periodic then it is nL – periodic for all $n \in \mathbb{N}$, so we also would learn

$$u(t,x) = \int_{-\pi nL}^{\pi nL} p_t^{nL}(x-y)f(y)dy \text{ for all } n \in \mathbb{N}.$$

this suggest that we might pass to the limit as $n \to \infty$ in this equation to learn

$$u(t,x) = \int_{\mathbb{R}} p_t(x-y)f(y)dy$$

where

$$\begin{aligned}p_t(x) &:= \lim_{n\to\infty} p_t^{nL}(x) = \lim_{L\to\infty}\frac{1}{2\pi L}\sum_{k\in\mathbb{Z}} e^{-\frac{1}{2}\left(\frac{k}{L}\right)^2 t}e^{i\left(\frac{k}{L}\right)x} \\ &= \frac{1}{2\pi}\int_{\mathbb{R}} e^{-\frac{1}{2}\xi^2 t}e^{i\xi x}d\xi = \frac{1}{\sqrt{2\pi t}}e^{-\frac{x^2}{2t}}.\end{aligned}$$

From this we conclude

$$u(t,x) = \int_{\mathbb{R}} p_t(x-y)f(y)dy = \int_{-\pi L}^{\pi L}\sum_{n\in\mathbb{Z}} p_t(x-y+2\pi nL)f(y)dy$$

and we arrive at the identity

$$\sum_{n\in\mathbb{Z}}\frac{1}{\sqrt{2\pi t}}e^{-\frac{(x+2\pi nL)^2}{2t}} = \sum_{n\in\mathbb{Z}} p_t(x+2\pi nL) = \frac{1}{2\pi L}\sum_{k\in\mathbb{Z}} e^{-\frac{1}{2}\left(\frac{k}{L}\right)^2 t}e^{ikx/L}$$

which is a special case of Poisson's summation formula.

46

Abstract Wave Equation

In the next section we consider

$$u_{tt} - \triangle u = 0 \text{ with } u(x,0) = f(x) \text{ and } u_t(x,0) = g(x) \text{ for } x \in \mathbb{R}^n. \quad (46.1)$$

Before working with this explicit equation we will work out an abstract Hilbert space theory first.

Theorem 46.1 (Existence). *Suppose $A : H \to H$ is a self-adjoint non-positive operator, i.e. $A^* = A$ and $A \le 0$ and $f \in D(A)$ and $g \in g \in D\left(\sqrt{-A}\right)$ are given. Then*

$$u(t) = \cos(t\sqrt{-A})f + \frac{\sin(t\sqrt{-A})}{\sqrt{-A}}g \quad (46.2)$$

satisfies:

1. *$\dot{u}(t) = \cos(t\sqrt{-A})\sqrt{-A}\ f + \sin(t\sqrt{-A})g$ exists and is continuous.*
2. *$\ddot{u}(t)$ exists and is continuous*

$$\ddot{u}(t) = Au \text{ with } u(0) = f \text{ and } \dot{u}(0) = g. \quad (46.3)$$

3. *$\frac{d}{dt}\sqrt{-A}\ u(t) = -\cos(t\sqrt{-A})A\ f + \sin(t\sqrt{-A})\sqrt{-A}g$ exists and is continuous.*

Eq. (46.3) is Newton's equation of motion for an infinite dimensional harmonic oscillation. Given any solution u to Eq. (46.3) it is natural to define its energy by

$$E(t,u) := \frac{1}{2}\left[\|\dot{u}(t)\|^2 + \|\omega u(t)\|^2\right] = \text{K.E.} + \text{ P.E.}$$

where $\omega := \sqrt{-A}$. Notice that Eq. (46.3) becomes $\ddot{u} + \omega^2 u = 0$ with this definition of ω.

Lemma 46.2 (Conservation of Energy). *Suppose u is a solution to Eq. (46.3) such that $\frac{d}{dt}\sqrt{-A}u(t)$ exists and is continuous. Then $\dot{E}(t) = 0$.*

Proof.

$$\dot{E}(t) = \mathrm{Re}(\dot{u}, \ddot{u}) + \mathrm{Re}(\omega u, \omega \dot{u}) = \mathrm{Re}(\dot{u}, -\omega^2 u) - \mathrm{Re}(\omega^2 u, \dot{u}) = 0.$$

■

Theorem 46.3 (Uniqueness of Solutions). *The only function $u \in C^2(\mathbb{R}, H)$ satisfying 1) $u(t) \in D(A)$ for all t and 2)*

$$\ddot{u} = Au \text{ with } u(0) = 0 = \dot{u}(0)$$

is the $u(t) \equiv 0$ for all t.

Proof. Let $\chi_M(x) = 1_{|x| \le M}$ and define $P_M = \chi_M(A)$ so that P_M is orthogonal projection onto the spectral subspace of H where $-M \le A \le 0$. Then for all $f \in D(A)$ we have $P_M A f = A P_M f$ and for all $f \in H$ we have $P_M f \in D((-A)^\alpha)$ for any $\alpha \ge 0$. Let $u_M(t) := P_M u(t)$, then $u_M \in C^2(\mathbb{R}, H)$, $u_M(t) \in D((-A)^\alpha)$ for all t and α, $t \to \sqrt{-A}u_M(t)$ is continuous and

$$\ddot{u}_M = \frac{d^2}{dt^2}(P_M u) = P_M \ddot{u} = P_M A u = A P_M u = A u_M$$

with $u_M(0) = 0 = \dot{u}_M(0)$. By Lemma 46.2,

$$\frac{1}{2}\left[\|\dot{u}_M(t)\|^2 + \|\omega u_M(t)\|^2\right] = \frac{1}{2}\left[\|\dot{u}_M(0)\|^2 + \|\omega u_M(0)\|^2\right] = 0$$

for all t. In particular this implies $\dot{u}_M(t) = 0$ and hence $P_M u(t) = u_M(t) \equiv 0$. Letting $M \to \infty$ then shows $u(t) \equiv 0$. ■

Corollary 46.4. *Any solution to $\ddot{u} = Au$ with $u(0) \in D(A)$ and $\dot{u}(0) \in D(\sqrt{-A})$ must satisfy $t \to \sqrt{-A}u(t)$ is C^1.*

46.1 Corresponding first order O.D.E.

Let $v(t) = \dot{u}(t)$, and

$$x(t) = \begin{pmatrix} u(t) \\ v(t) \end{pmatrix} = \begin{pmatrix} u \\ \dot{u} \end{pmatrix}$$

then

$$\dot{x} = \begin{pmatrix} \dot{u} \\ \ddot{u} \end{pmatrix} = \begin{pmatrix} v \\ Au \end{pmatrix} = \begin{pmatrix} 0 & I \\ A & 0 \end{pmatrix} x = Bx \text{ with}$$

$$x(0) = \begin{pmatrix} f \\ g \end{pmatrix},$$

where

$$B := \begin{pmatrix} 0 & I \\ A & 0 \end{pmatrix} = \begin{pmatrix} 0 & I \\ -\omega^2 & 0 \end{pmatrix}.$$

Note formally that

$$e^{tB}\begin{pmatrix} f \\ g \end{pmatrix} = \begin{pmatrix} u(t) \\ \dot{u}(t) \end{pmatrix} = \begin{pmatrix} \cos\omega t\ f + \frac{\sin t\omega}{\omega} g \\ -\omega\sin\omega t\ f + \cos\omega t\ g \end{pmatrix}$$
$$= \begin{pmatrix} \cos\omega t & \frac{\sin t\omega}{\omega} \\ -\omega\sin\omega t & \cos\omega t \end{pmatrix}\begin{pmatrix} f \\ g \end{pmatrix} \tag{46.4}$$

and this suggests that

$$e^{tB} = \begin{pmatrix} \cos\omega t & \frac{\sin t\omega}{\omega} \\ -\omega\sin\omega t & \cos\omega t \end{pmatrix}$$

which is formally correct since

$$\frac{d}{dt}e^{tB} = \begin{pmatrix} -\omega\sin\omega t & \cos\omega t \\ -\omega^2\cos\omega t & -\omega\sin\omega t \end{pmatrix}$$
$$= \begin{pmatrix} 0 & I \\ -\omega^2 & 0 \end{pmatrix}\begin{pmatrix} \cos\omega t & \frac{\sin t\omega}{\omega} \\ -\omega\sin\omega t & \cos\omega t \end{pmatrix} = Be^{tB}.$$

Since the energy form $E(t) = \|\dot{u}\|^2 + \|\omega\ u\|^2$ is conserved, it is reasonable to let

$$K = D(\sqrt{-A}) \oplus H := \begin{pmatrix} D(\sqrt{-A}) \\ H \end{pmatrix}$$

with inner product

$$\left\langle \begin{matrix} f \\ g \end{matrix} \middle| \begin{matrix} \tilde{f} \\ \tilde{g} \end{matrix} \right\rangle = (g, \tilde{g}) + (\omega f, \omega\tilde{f}).$$

For simplicity we assume $\mathrm{Nul}(\sqrt{-A}) = \mathrm{Nul}(\omega) = \{0\}$ in which case K becomes a Hilbert space and e^{tB} is a unitary evolution on K. Indeed,

$$\|e^{tB}\begin{pmatrix} f \\ g \end{pmatrix}\|_K^2 = \|\cos\omega t g - \omega\sin\omega t f\|^2 + \|\omega(\cos + \omega f) + \sin\omega t g\|^2$$
$$= \|\cos\omega t g\|^2 + \|\omega\sin\omega t f\|^2 + \|\omega\cos\omega t f\|^2 + \|\sin\omega t g\|^2$$
$$= \|\omega f\|^2 + \|g\|^2.$$

From Eq. (46.4), it easily follows that $\frac{d}{dt}\Big|_0 e^{tB}\begin{pmatrix} f \\ g \end{pmatrix}$ exists iff $g \in D(\omega)$ and $f \in D(-\omega^2) = D(A)$. Therefore we define $D(B) := D(A) \oplus D(\omega) = D(A) \oplus D\left(\sqrt{-A}\right)$and

$$B = \begin{pmatrix} 0 & I \\ A & 0 \end{pmatrix} : D(B) \to \begin{matrix} D(\omega) \\ \oplus \\ H \end{matrix} = K$$

Since B is the infinitesimal generator of a unitary semigroup, it follows that $B^{*K} = -B$, i.e. B is skew adjoint. This may be checked directly as well as follows.

Alternate Proof that $B^{*K} = -B$. For

$$\begin{pmatrix} u \\ v \end{pmatrix}, \begin{pmatrix} \tilde{u} \\ \tilde{v} \end{pmatrix} \in D(B) = D(A) \oplus D(\omega),$$

$$\langle B\begin{pmatrix} u \\ v \end{pmatrix}, \begin{pmatrix} \tilde{u} \\ \tilde{v} \end{pmatrix}\rangle = \langle\begin{pmatrix} v \\ Au \end{pmatrix}, \begin{pmatrix} \tilde{u} \\ \tilde{v} \end{pmatrix}\rangle = (Au, \tilde{v}) + (\omega v, \omega\tilde{u})$$
$$= (Au, \tilde{v}) - (Av, \tilde{u}) = (u, A\tilde{v}) - (v, A\tilde{u})$$

and similarly

$$\langle\begin{pmatrix} u \\ v \end{pmatrix}, B\begin{pmatrix} \tilde{u} \\ \tilde{v} \end{pmatrix}\rangle = \langle\begin{pmatrix} u \\ v \end{pmatrix}, \begin{pmatrix} \tilde{v} \\ A\tilde{u} \end{pmatrix}\rangle = (\omega u, \omega\tilde{v}) + (v, A\tilde{u})$$
$$= (-Au, \tilde{v}) + (v, A\tilde{u}) = -\langle B\begin{pmatrix} u \\ v \end{pmatrix}, \begin{pmatrix} \tilde{u} \\ \tilde{v} \end{pmatrix}\rangle$$

which shows $-B \subset B^*$. Conversely if $\begin{pmatrix} \tilde{u} \\ \tilde{v} \end{pmatrix} \in D(B^*)$ and $B^*\begin{pmatrix} \tilde{u} \\ \tilde{v} \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix}$, then

$$\langle B\begin{pmatrix} u \\ v \end{pmatrix} \Big| \begin{matrix} \tilde{u} \\ \tilde{v} \end{matrix}\rangle = \langle \begin{matrix} u \\ v \end{matrix} \Big| \begin{matrix} f \\ g \end{matrix}\rangle = (v, g) + (\omega u, \omega f) \tag{46.5}$$

$(Au, \tilde{v}) + (\omega v, \omega\tilde{u})$ for all $u \in D(A), v \in D(\omega)$. Take $u = 0$ implies $(\omega v, \omega\tilde{u}) = (v, g)$ for all $v \in D(\omega)$ which then implies $\omega\tilde{u} \in D(\omega^*) = D(\omega)$ and hence $-A\tilde{u} = \omega^2\tilde{u} = g$. (Note $\tilde{u} \in D(A)$.) Taking $v = 0$ in Eq. (46.5) implies $(Au, \tilde{v}) = (\omega u, \omega f) = (-Au, f)$. Since

$$\overline{\mathrm{Ran}(A)} = \mathrm{Nul}(A)^\perp = \{0\}^\perp = H,$$

we find that $f = -\tilde{v} \in D(\omega)$ since $f \in D(\omega)$. Therefore $D(B^*) \subset D(B)$ and for $(\tilde{u}, \tilde{v}) \in D(B^*)$ we have

$$B^*\begin{pmatrix} \tilde{u} \\ \tilde{v} \end{pmatrix} = -\begin{pmatrix} -\tilde{v} \\ -A\tilde{u} \end{pmatrix} = -B\begin{pmatrix} \tilde{u} \\ \tilde{v} \end{pmatrix}.$$

46.2 Du Hamel's Principle

Consider

$$\ddot{u} = Au + f(t) \text{ with } u(0) = g \text{ and } \dot{u}(0) = h. \tag{46.6}$$

Eq. (46.6) implies, with $v = \dot{u}$, that

$$\frac{d}{dt}\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} v \\ \ddot{u} \end{pmatrix} = \begin{pmatrix} v \\ Au + f \end{pmatrix} = \begin{pmatrix} 0 & I \\ A & 0 \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix} + \begin{pmatrix} 0 \\ f \end{pmatrix}.$$

Therefore

$$\begin{pmatrix} u \\ v \end{pmatrix}(t) = e^{t\begin{pmatrix} 0 & I \\ A & 0 \end{pmatrix}} \begin{pmatrix} g \\ h \end{pmatrix} + \int_0^t e^{(t-\tau)\begin{pmatrix} 0 & I \\ A & 0 \end{pmatrix}} \begin{pmatrix} 0 \\ f(\tau) \end{pmatrix} d\tau$$

hence

$$u(t) = \cos(t\sqrt{-A})g + \frac{\sin(t\sqrt{-A})}{\sqrt{-A}}h + \int_0^t \frac{\sin((t-\tau)\sqrt{-A})}{\sqrt{-A}}f(\tau)d\tau.$$

Theorem 46.5. *Suppose $f(t) \in D(\sqrt{-A})$ for all t and that $f(t)$ is continuous relative to $\|f\|_A := \|f\| + \|\sqrt{-A}\, f\|$. Then*

$$u(t) := \int_0^t \frac{\sin((t-\tau)\sqrt{-A})}{\sqrt{-A}} f(\tau)d\tau$$

solves $\ddot{u} = Au + f$ with $u(0) = 0, \dot{u}(0) = 0$.

Proof. $\dot{u}(t) = \int_0^t \cos((t-\tau)\sqrt{-A})f(\tau)d\tau.$

$$\begin{aligned} \ddot{u}(t) &= f(t) - \int_0^t \sin((t-\tau)\sqrt{-A})\sqrt{-A}\, f(\tau)d\tau \\ &= f(t) - A\int_0^t \frac{\sin((t-\tau)\sqrt{-A})}{\sqrt{-A}} f(\tau)d\tau. \end{aligned}$$

So $\dot{u} = Au + f$. Note $u(0) = 0 = \dot{u}(0)$.

Alternate. Let $\omega := \sqrt{-A}$, then

$$\begin{aligned} u(t) &= \int_0^t \frac{\sin((t-\tau)\omega)}{\omega} f(\tau)d\tau \\ &= \int_0^t \frac{\sin\omega t\cos\omega\tau - \sin\omega\tau\cos\omega t}{\omega} f(\tau)d\tau \end{aligned}$$

and hence

$$\begin{aligned} \dot{u}(t) &= \frac{\sin\omega t\cos\omega t - \sin\omega t\cos\omega t}{\omega} f(t) \\ &\quad + \int_0^t (\cos\omega t\cos\omega\tau + \sin\omega\tau\sin\omega t)\ f(\tau)d\tau \\ &= \int_0^t (\cos\omega t\cos\omega\tau + \sin\omega\tau\sin\omega t)\ f(\tau)d\tau. \end{aligned}$$

Similarly,

$$\begin{aligned} \ddot{u}(t) &= (\cos\omega t\cos\omega t + \sin\omega t\sin\omega t)\ f(t) \\ &\quad + \int_0^t \omega\,(-\sin\omega t\cos\omega\tau + \sin\omega\tau\cos\omega t)\ f(\tau)d\tau \\ &= f(t) - \int_0^t \sin((t-\tau)\omega)\ \omega f(\tau)d\tau = f(t) - \omega^2 u(t) \\ &= Au(t) + f(t). \end{aligned}$$

■

47

Wave Equation on $\mathbb{R}^n$

(**Ref** Courant & Hilbert Vol II, Chap VI §12.)

We now consider the wave equation

$$u_{tt} - \triangle u = 0 \text{ with } u(0,x) = f(x) \text{ and } u_t(0,x) = g(x) \text{ for } x \in \mathbb{R}^n. \tag{47.1}$$

According to Section 46, the solution (in the L^2 – sense) is given by

$$u(t,\cdot) = (\cos(t\sqrt{-\triangle})f + \frac{\sin(t\sqrt{-\triangle})}{\sqrt{-\triangle}}g. \tag{47.2}$$

To work out the results in Eq. (47.2) we must diagonalize Δ. This is of course done using the Fourier transform. Let $\mathcal{F}$ denote the Fourier transform in the x – variables only. Then

$$\ddot{\hat{u}}(t,k) + |k|^2 \hat{u}(t,k) = 0 \text{ with}$$
$$\hat{u}(0,k) = \hat{f}(k) \text{ and } \dot{\hat{u}}(t,k) = \hat{g}(k).$$

Therefore

$$\hat{u}(t,k) = \cos(t|k|)\hat{f}(k) + \frac{\sin(t|k|)}{|k|}\hat{g}(k).$$

and so

$$u(t,x) = \mathcal{F}^{-1}\left[\cos(t|k|)\hat{f}(k) + \frac{\sin(t|k|)}{|k|}\hat{g}(k)\right](x),$$

i.e.

$$\frac{\sin(t\sqrt{-\triangle})}{\sqrt{-\triangle}}g = \mathcal{F}^{-1}\left[\frac{\sin(t|k|)}{|k|}\hat{g}(k)\right] \text{ and} \tag{47.3}$$

$$\cos(t\sqrt{-\triangle})f = \mathcal{F}^{-1}\left[\cos(t|k|)\hat{f}(k)\right] = \frac{d}{dt}\mathcal{F}^{-1}\left[\frac{\sin(t|k|)}{|k|}\hat{g}(k)\right]. \tag{47.4}$$

.

Our next goal is to work out these expressions in x – space alone.

47.1 $n = 1$ Case

As we see from Eq. (47.4) it suffices to compute:

$$\begin{aligned}\frac{\sin(t\sqrt{-\triangle})}{\sqrt{-\triangle}}g &= \mathcal{F}^{-1}\left(\frac{\sin(t|\xi|)}{|\xi|}\hat{g}(\xi)\right) = \lim_{M\to\infty}\mathcal{F}^{-1}\left(1_{|\xi|\le M}\frac{\sin(t|\xi|)}{|\xi|}\hat{g}(\xi)\right) \\ &= \lim_{M\to\infty}\mathcal{F}^{-1}\left(1_{|\xi|\le M}\frac{\sin(t|\xi|)}{|\xi|}\right)\star g. \end{aligned}\tag{47.5}$$

This inverse Fourier transform will be computed in Proposition 47.2 below using the following lemma.

Lemma 47.1. *Let C_M denote the contour shown in Figure 47.1, then for $\lambda \neq 0$ we have*

$$\lim_{M\to\infty}\int_{C_M}\frac{e^{i\lambda\xi}}{\xi}d\xi = 2\pi i 1_{\lambda>0}.$$

Proof. First assume that $\lambda > 0$ and let Γ_M denote the contour shown in Figure 47.1. Then

$$\left|\int_{\Gamma_M}\frac{e^{i\lambda\xi}}{\xi}d\xi\right| \le \int_0^\pi \left|e^{i\lambda M e^{i\theta}}\right| d\theta = 2\pi\int_0^\pi d\theta e^{-\lambda M\sin\theta} \to 0 \text{ as } M \to \infty.$$

Therefore

$$\lim_{M\to\infty}\int_{C_M}\frac{e^{i\lambda\xi}}{\xi}d\xi = \lim_{M\to\infty}\int_{C_M+\Gamma_M}\frac{e^{i\lambda\xi}}{\xi}d\xi = 2\pi i \operatorname{res}_{\xi=0}\left(\frac{e^{i\lambda\xi}}{\xi}\right) = 2\pi i.$$

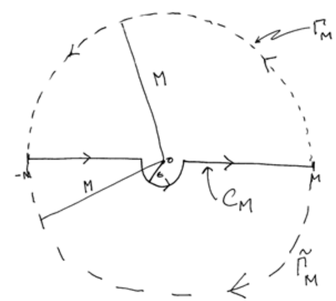


Fig. 47.1. A couple of contours in $\mathbb{C}$.

If $\lambda < 0$, the same argument shows

$$\lim_{M\to\infty}\int_{C_M}\frac{e^{i\lambda\xi}}{\xi}\,d\xi=\lim_{M\to\infty}\int_{C_M+\tilde{\Gamma}_M}\frac{e^{i\lambda\xi}}{\xi}d\xi$$

and the later integral is 0 since the integrand is holomorphic inside the contour $C_M+\tilde{\Gamma}_M$. ■

Proposition 47.2. $\lim_{M\to\infty}\mathcal{F}^{-1}\left(1_{|\xi|\le M}\,\frac{\sin(t|\xi|)}{|\xi|}\right)(x)=\operatorname{sgn}(t)\frac{\sqrt{\pi}}{\sqrt{2}}\,1_{|x|<|t|}$.

Proof. Let

$$I_M=\sqrt{2\pi}\mathcal{F}^{-1}\left(1_{|\xi|\le M}\,\frac{\sin(t|\xi|)}{|\xi|}\right)(x)=\int_{|\xi|\le M}\frac{\sin(t\xi)}{\xi}\,e^{i\xi\cdot x}d\xi.$$

Then by deforming the contour we may write

$$I_M=\int_{C_M}\frac{\sin t\xi}{\xi}\,e^{i\xi\cdot x}d\xi=\frac{1}{2i}\int_{C_M}\frac{e^{it\xi}-e^{-it\xi}}{\xi}\,e^{i\xi\cdot x}d\xi$$
$$=\frac{1}{2i}\int_{C_M}\frac{e^{i(x+t)\xi}-e^{i(x-t)\xi}}{\xi}d\xi$$

By Lemma 47.1 we conclude that

$$\lim_{M\to\infty}I_M=\frac{1}{2i}2\pi i(1_{(x+t)>0}-1_{(x-t)>0})=\pi\operatorname{sgn}(t)\,1_{|x|<|t|}.$$

(For the last equality, suppose $t>0$. Then $x-t>0$ implies $x+t>0$ so we get 0 and if $x<-t$, i.e. $x+t<0$ then $x-t<0$ and we get 0 again. If $|x|<t$ the first term is 1 while the second is zero. Similar arguments work when $t<0$ as well.) ■

Theorem 47.3. *For* $n=1$,

$$\frac{\sin(t\sqrt{-\triangle})}{\sqrt{-\triangle}}\,g(x)=\frac{1}{2}\int_{x-t}^{x+t}g(y)\,d\lambda(y)\ \text{and} \tag{47.6}$$

$$\cos(t\sqrt{-\triangle})g(x)=\frac{1}{2}\,[g(x+t)+g(x-t)]. \tag{47.7}$$

In particular

$$u(t,x)=\frac{1}{2}(f(x+t)+f(x-t))+\frac{1}{2}\int_{x-t}^{x+t}g(y)\ dy \tag{47.8}$$

is the solution to the wave equation (47.2).

Proof. From Eq. (47.5) and Proposition 47.2 we find

$$\frac{\sin(t\sqrt{-\triangle})}{\sqrt{-\triangle}}\,g(x)=\operatorname{sgn}(t)\frac{1}{2}\int_{\mathbb{R}}1_{|x-y|>|t|}\,g(y)\,dy$$
$$=\operatorname{sgn}(t)\frac{1}{2}\int_{x-|t|}^{x+|t|}g(y)\,dy=\frac{1}{2}\int_{x-t}^{x+t}g(y)\,dy.$$

Differentiating this equation in t gives Eq. (47.7). ■

If we have a forcing term, so $\ddot{u}=u_{x\ x}+h$, with $u(0,\cdot)=0$ and $u_t(0,\cdot)=0$, then

$$u(t,x)=\int_0^t\frac{\sin((t-\tau)\sqrt{-\triangle})}{\sqrt{-\triangle}}h(\tau,x)d\tau=\frac{1}{2}\int_0^t d\tau\int_{x-t+\tau}^{x+t-\tau}dyh(\tau,y)$$
$$=\frac{1}{2}\int_0^t d\tau\int_{-(t+\tau)}^{t-\tau}drh(\tau,x+r).$$

47.1.1 Factorization method for $n = 1$

Writing the wave equation as

$$0=\left(\partial_t^2-\partial_x^2\right)u=(\partial_t+\partial_x)(\partial_t-\partial_x)u=(\partial_t+\partial_x)v$$

with $v:=(\partial_t-\partial_x)u$ implies $v(t,x)=v(0,x-t)$ with

$$v(0,x)=u_t(0,x)-u_x(0,x)=g(x)-f'(x).$$

Now u solves $(\partial_t-\partial_x)u=v$, i.e. $\partial_t u=\partial_x u+v$. Therefore

$$\begin{aligned}
u(t,x)&=e^{t\partial_x}u(0,x)+\int_0^t e^{(t-\tau)\partial_x}\ v(\tau,x)d\tau\\
&=u(0,x+t)+\int_0^t v(\tau,x+t-\tau)d\tau\\
&=u(0,x+t)+\int_0^t v(0,x+\underbrace{t-2\tau}_{s})d\tau\\
&=u(0,x+t)+\frac{1}{2}\int_{-t}^t v(0,x+s)ds\\
&=f(x+t)+\frac{1}{2}\int_{-t}^t(g(x+s)-f'(x+s))ds\\
&=f(x+t)-\frac{1}{2}\ f(x+s)\Big|_{s=-t}^{s=t}+\frac{1}{2}\int_{-t}^t g(x+s)ds\\
&=\frac{f(x+t)+f(x-t)}{2}+\frac{1}{2}\int_{-t}^t g(x+s)ds
\end{aligned}$$

which is equivalent to Eq. (47.8).

47.2 Solution for $n = 3$

Given a function $f : \mathbb{R}^n \to \mathbb{R}$ and $t \in \mathbb{R}$ let

$$\bar{f}(x;t) := \fint_{S^2} f(x+t\omega)d\sigma(\omega) = \fint\int_{|y|=|t|} f(x+y)d\sigma(y).$$

Theorem 47.4. *For $f \in L^2(\mathbb{R}^3)$,*

$$\frac{\sin\left(\sqrt{-\Delta}t\right)}{\sqrt{-\Delta}} f = \mathcal{F}^{-1}\left[\frac{\sin|\xi|t}{|\xi|}\hat{f}(\xi)\right](x) = t\bar{f}(x;t)$$

and

$$\cos\left(\sqrt{-\Delta}t\right)g = \frac{d}{dt}\left[t\bar{f}(x;t)\right].$$

In particular the solution to the wave equation (47.1) for $n = 3$ is given by

$$\begin{aligned} u(t,x) &= \frac{\partial}{\partial t}(t\ \bar{f}(x;t)) + t\ \bar{g}(x;t) \\ &= \frac{1}{4\pi}\int_{|\omega|=1}(tg(x+t\omega)+f(x+t\omega)+t\nabla f(x+t\omega)\cdot\omega)d\sigma(\omega). \end{aligned}$$

Proof. Let $g_M := \mathcal{F}^{-1}\left[\frac{\sin|\xi|t}{|\xi|}1_{|\xi|\le M}\right]$, then by symmetry and passing to spherical coordinates,

$$\begin{aligned} (2\pi)^{3/2} g_M(x) &= \int_{|\xi|\le M}\frac{\sin|\xi|t}{|\xi|}e^{i\xi\cdot x}d\xi = \int_{|\xi|\le M}\frac{\sin|\xi|t}{|\xi|}\ e^{i|x|\xi_3}d\xi \\ &= \int_0^M d\rho\rho^2\int_0^{2\pi}d\theta\int_0^\pi d\phi\frac{\sin\rho t}{\rho}e^{i\rho|x|\cos\phi}\sin\phi \\ &= 2\pi\int_0^M d\rho\sin\rho t\left.\frac{e^{i\rho|x|\cos\phi}}{-i|x|}\right|_0^\pi \\ &= 2\pi\int_0^M d\rho\sin\rho t\frac{e^{i\rho|x|}-e^{-i\rho|x|}}{i|x|} = \frac{4\pi}{|x|}\int_0^M \sin\rho t\sin\rho\,|x|\ \ d\rho. \end{aligned}$$

Using

$$\sin A\sin B = \frac{1}{2}\left[\cos(A-B)-\cos(A+B)\right]$$

in this last equality, shows

$$\begin{aligned} g_M(x) &= (2\pi)^{-3/2}\frac{2\pi}{|x|}\int_0^M[\cos((t-|x|)\rho)-\cos((t+|x|)\rho)]d\rho \\ &= (2\pi)^{-3/2}\frac{\pi}{|x|}h_M(|x|) \end{aligned}$$

where

$$h_M(r) := \int_{-M}^M[\cos((t-r)\alpha)-\cos((t+r)\alpha)]d\alpha,$$

an odd function in r. Since

$$\mathcal{F}^{-1}\left[\frac{\sin|\xi|t}{|\xi|}\ \hat{f}(\xi)\right] = \lim_{M\to\infty}\mathcal{F}^{-1}(\hat{g}_M(\xi)\hat{f}(\xi)) = \lim_{M\to\infty}(g_M\star f)(x)$$

we need to compute $g_M\star f$. To this end

$$\begin{aligned} g_M\star f(x) &= \left(\frac{1}{2\pi}\right)^3\pi\int_{\mathbb{R}^3}\frac{1}{|y|}h_M(|y|)f(x-y)dy \\ &= \left(\frac{1}{2\pi}\right)^3\pi\int_0^\infty d\rho\frac{h_M(\rho)}{\rho}\int_{|y|=\rho}f(x-y)d\sigma(y) \\ &= \left(\frac{1}{2\pi}\right)^3\pi\int_0^\infty d\rho\ \frac{h_M(\rho)}{\rho}4\pi\rho^2\fint_{|y|=\rho}f(x-y)d\sigma(y) \\ &= \frac{1}{2\pi}\int_0^\infty d\rho\ h_M(\rho)\rho\bar{f}(x;\rho) = \frac{1}{4\pi}\int_{-\infty}^\infty d\rho\ h_M(\rho)\rho\bar{f}(x;\rho) \end{aligned}$$

where the last equality is a consequence of the fact that $h_M(\rho)\rho\bar{f}(x;\rho)$ is an even function of ρ. Continuing to work on this expression suing $\rho\to\rho\bar{f}(x;\rho)$ is odd implies

$$\begin{aligned} g_M\star f(x) &= \frac{1}{4\pi}\int_{-\infty}^\infty d\rho\int_{-M}^M[\cos((t-\rho)\alpha)-\cos((t+\rho)\alpha)]d\alpha\,\rho\bar{f}(x;\rho) \\ &= \frac{1}{2\pi}\int_{-\infty}^\infty d\rho\int_{-M}^M\cos((t-\rho)\alpha)\rho\bar{f}(x;\rho)d\alpha \\ &= \frac{1}{2\pi}\text{Re}\int_{-M}^M d\rho\int_{-\infty}^\infty d\alpha e^{i(t-\rho)\alpha}\rho\bar{f}(x;\rho)d\alpha \to t\bar{f}(x;t)\text{ as }M\to\infty \end{aligned}$$

using the 1 – dimensional Fourier inversion formula. ■

47.2.1 Alternate Proof of Theorem 47.4

Lemma 47.5. $\lim_{M\to\infty}\int_{-M}^M\cos(\rho\lambda)d\rho = 2\pi\delta(\lambda).$

Proof.

$$\int_{-M}^{M} \cos(\rho\lambda)d\rho = \int_{-M}^{M} e^{i\rho\lambda} d\rho$$

so that

$$\int_{\mathbb{R}} \phi(\lambda) \left[\int_{-M}^{M} e^{i\rho\lambda} d\rho \right] d\lambda \to \int_{\mathbb{R}} d\rho \int_{\mathbb{R}} d\lambda \phi(\lambda) e^{i\lambda\rho} = 2\pi\varphi(0)$$

by the Fourier inversion formula. ■

Proof. of Theorem 47.4 again.

$$\begin{aligned}
\int \frac{\sin t|\xi|}{|\xi|} e^{i\xi\cdot x} d\xi &= \int \frac{\sin t\rho}{\rho} e^{i\rho|x|\cos\theta} \sin\theta d\theta\, d\varphi\, \rho^2\, d\rho \\
&= 2\pi \int \frac{\sin t\rho}{\rho} \frac{e^{i\rho|x|\lambda}}{i\rho|x|} \Big|_{\lambda=-1}^{1} d\rho \\
&= \frac{4\pi}{|x|} \int_0^\infty \sin t\rho\, \sin\rho|x|\, d\rho \\
&= \frac{2\pi}{|x|} \int_0^\infty [\cos(\rho(t-|x|)) - \cos(\rho(t+|x|))]\, d\rho \\
&= \frac{4\pi}{|x|} \int_{-\infty}^\infty [\cos(\rho(t-|x|)) - \cos(\rho(t+|x|))]\, d\rho \\
&= \frac{8\pi^2}{|x|} (\delta(t-|x|) - \delta(t+|x|))
\end{aligned}$$

Therefore

$$\begin{aligned}
&\mathcal{F}^{-1}\left(\frac{\sin t|\xi|}{|\xi|}\right) * g(x) \\
&= \left(\frac{1}{2\pi}\right)^3 2\pi^2 \int_{\mathbb{R}^3} \frac{(\delta(t-|y|) - \delta(t+|y|))}{|y|} g(x-y)\, d\lambda(y) \\
&= \frac{1}{4\pi} \int_0^\infty (\delta(t-\rho) - \delta(t+\rho)) g(x+\rho\omega) \frac{\rho^2}{\rho}\, d\rho\, d\sigma(\omega) \\
&= 1_{t>0}\, t\, \overline{g}(x;t) - 1_{t<0}\, (-t)\, \overline{g}(x;-t) \\
&= t\overline{g}(x;t)
\end{aligned}$$

■

47.3 Du Hamel's Principle

The solution to

$$u_{tt} = \triangle u + f \text{ with } u(0,x) = 0 \text{ and } u_t(0,x) = 0$$

is given by

$$u(t,x) = \frac{1}{4\pi} \int_{B(x,t)} \frac{f(t-|y-x|, y)}{|y-x|} dy = \frac{1}{4\pi} \int_{|z|<t} \frac{f(t-|z|, x+z)}{|z|} dz. \quad (47.9)$$

Indeed, by Du Hamel's principle,

$$\begin{aligned}
u(t,x) &= \int_0^t \frac{\sin((t-\tau)\sqrt{-\triangle})}{\sqrt{-\triangle}} f(\tau,x)d\tau = \int_0^t \frac{\sin(\tau\sqrt{-\triangle})}{\sqrt{-\triangle}} f(t-\tau,x)d\tau \\
&= \int_0^t \tau\overline{f}(t-\tau,x;\tau)d\tau = \frac{1}{4\pi}\int_0^t d\tau\, t^2 \int_{|\omega|=1} \frac{f(t-\tau, x+\tau\omega)}{\tau} d\sigma(\omega) \\
&= \frac{1}{4\pi} \int_{B(t,x)} \frac{f(t-|y-x|, y)}{|y-x|} dy \text{ (let } y = x+z) \\
&= \frac{1}{4\pi} \int_{|z|<t} \frac{f(t-|z|, x+z)}{|z|} dz.
\end{aligned}$$

Thinking of $u(t,x)$ as pressure (47.9) says that the pressure at x at time t is the "average" of the disturbance at time $t-|y-x|$ at location y.

47.4 Spherical Means

Let $n \geq 2$ and suppose u solves $u_{tt} = \triangle u$. Since Δ is invariant under rotations, i.e. for $R \in O(n)$ we have $\triangle(u \circ R) = (\triangle u) \circ R$, it follows that $u \circ R$ is also a solution to the wave equation. Indeed,

$$(u(t,\cdot) \circ R)_{tt} = u_{tt}(t,\cdot) \circ R = \Delta u(t,\cdot) \circ R = \triangle(u(t,\cdot) \circ R).$$

By the linearity of the wave equation, this also implies, with dR denoting normalized Haar measure on $O(n)$, that

$$U(t,|x|) := \int_{O(n)} (u(t,Rx) \circ R)dR$$

must be a radial solution of the Wave equation. This implies

$$U_{tt} = \triangle_x U(t,|x|) = \frac{1}{r^{n-1}} \partial_r (r^{n-1}\partial_r U(t,r))_{r=|x|} = \left[\partial_r^2 U(t,r) + \frac{n-1}{r}\, \partial_r U(t,r)\right]_r$$

Now

$$U(t,|x|) = \int_{0(n)} u(t,Rx)dR = \fint_{B(0,|x|)} u(t,y)d\sigma(y).$$

Using the translation invariance of Δ the same argument as above gives the following theorem.

Theorem 47.6. *Suppose $u_{tt} = \triangle u$ and $x \in \mathbb{R}^n$ and let*

$$U(t,r) := \overline{u}(t,x;r) := \fint_{\partial B(x,r)} u(t,y)d\sigma(y)$$
$$= \fint_{\partial B(0,1)} u(t,x+r\omega)d\sigma(\omega).$$

Then U solves

$$U_{tt} = \frac{1}{r^{n-1}}\,\partial_r(r^{n-1}U_r)$$

with

$$U(0,r) = \fint_{\partial B(0,1)} u(0,x+r\omega)d\sigma(\omega) = \bar{f}(x;r)$$
$$U_t(0,r) = \overline{g}(x;r).$$

Proof. This has already been proved, nevertheless, let us give another proof which does not rely on using integration over $O(n)$. To this hence we compute

$$\begin{aligned}
\partial_r U(t,r) &= \partial_r \fint_{\partial B(0,1)} u(t,x+r\omega)d\sigma(\omega) \\
&= \fint_{\partial B(0,1)} \nabla u(t,x+r\omega)\cdot\omega d\sigma(\omega) \\
&= \frac{1}{\sigma(S^{n-1})\,r^{n-1}}\int_{|y|=r} \nabla u(t,x+y)\cdot\hat{y}d\sigma(y) \\
&= \frac{1}{\sigma(S^{n-1})\,r^{n-1}}\int_{|y|\le r} \Delta u(t,x+y)dy \\
&= \frac{1}{\sigma(S^{n-1})\,r^{n-1}}\int_0^r d\rho\int_{|y|=\rho} \Delta u(t,x+y)d\sigma(y)
\end{aligned}$$

so that

$$\begin{aligned}
\frac{1}{r^{n-1}}\partial_r(r^{n-1}U_r) &= \frac{1}{r^{n-1}}\partial_r\left[\frac{1}{\sigma(S^{n-1})}\int_0^r d\rho\int_{|y|=\rho} \Delta u(t,x+y)d\sigma(y)\right] \\
&= \frac{1}{\sigma(S^{n-1})\,r^{n-1}}\int_{|y|=r} \Delta u(t,x+y)d\sigma(y) \\
&= \fint_{|y|=r} \Delta u(t,x+y)d\sigma(y) \\
&= \fint_{|y|=r} u_{tt}(t,x+y)d\sigma(y) = U_{tt}.
\end{aligned}$$

■

We can now use the above result to solve the wave equation. For simplicity, assume $n = 3$ and let $V(t,r) = r\,\overline{u}(t,x;r) = r\,U(t,r)$. Then for $r > 0$ we have

$$\begin{aligned}
V_{rr} &= 2U_r + r\,U_{rr} = r(U_{rr} + \frac{2}{r}\,U_r) \\
&= r\,U_{tt} = V_{tt}.
\end{aligned}$$

This is also valid for $r < 0$ because $V(t,r)$ is odd in r. Indeed for $r < 0$, let $v(t,r) = V(t,-r)$, then $V_{rr}(t,r) = V_{rr}(t,-r) = V_{tt}(t,-r) = V_{tt}(t,r)$. By our solution to the one dimensional wave equation we find

$$V(t,r) = \frac{1}{2}(V(0,t+r) + V(0,r-t)) + \frac{1}{2}\int_{r-t}^{r+t} V_t(0,y)dy.$$

Now suppose that $u(0,x) = 0$ and $u_t(0,x) = g(x)$, in which case

$$V(0,r) = 0 \text{ and } V_t\,(0,r) = r\overline{g}(x,r)$$

and the previous equation becomes

Then

$$V(t,r) = \frac{1}{2}\int_{r-t}^{r+t} y\overline{g}(x,y)dy$$

and noting that

$$\frac{\partial}{\partial r}\Big|_0 V(t,r) = \overline{u}(t,x;0) = u(t,x)$$

we learn

$$u(t,x) = \frac{1}{2}\left[t\overline{g}(x;t) - (-t)\,\overline{g}(x;-t)\right] = t\overline{g}(x;t)$$

as before.

47.5 Energy methods

Theorem 47.7 (Uniqueness on Bounded Domains). *Let Ω be a bounded domain such that $\bar{\Omega}$ is a submanifold with C^2 – boundary and consider the boundary value problem*

$$\begin{aligned}
&u_{tt} - \triangle u = h \text{ on } \Omega_T \\
&u = f \qquad \text{on } (\partial\Omega \times [0,T]) \cup (\Omega \times \{t=0\}) \\
&u_t = g \qquad \text{on } \Omega \times \{t=0\}
\end{aligned}$$

If $u \in C^2(\overline{\Omega}_T)$ then u is unique.

Proof. As usual, using the linearity of the equation, it suffices to consider the special case where $f = 0$, $g = 0$ and $h = 0$ and to show this implies $u \equiv 0$. Let

$$E_\Omega(t) = \frac{1}{2}\int_\Omega \left[\dot{u}(t,x)^2 + |\nabla u(t,x)|^2\right] dx.$$

Clearly by assumption, $E_\Omega(0) = 0$ while the usual computation shows

$$\begin{aligned}\dot{E}_\Omega(t) &= (\dot{u},\ddot{u})_{L^2(\Omega)} + (\nabla u(t), \nabla\dot{u}(t))_{L^2(\Omega)}\\ &= (\dot{u},\Delta u)_{L^2(\Omega)} + (\nabla u(t), \nabla\dot{u}(t))_{L^2(\Omega)}\\ &= -(\nabla\dot{u}(t), \nabla u(t))_{L^2(\Omega)} + \int_{\partial\Omega} \dot{u}(t,x)\frac{\partial u(t,x)}{\partial n} d\sigma(x)\\ &\quad + (\nabla u(t), \nabla\dot{u}(t))_{L^2(\Omega)}\\ &= 0\end{aligned}$$

wherein we have used $u(t,x) = 0$ implies $\dot{u}(t,x) = 0$ for $x \in \partial\Omega$.

From this we conclude that $E_\Omega(t) = 0$ and therefore $\dot{u}(t,x) = 0$ and hence $u \equiv 0$. ■

The following proposition is expected to hold given the finite speed of propagation we have seen exhibited above for solutions to the wave equation.

Proposition 47.8 (Local Energy). *Let $x \in \mathbb{R}^n$, $T > 0$, $u_{tt} = \Delta u$ and define*

$$e(t) := E_{B(x,T-t)}(u;t) := \frac{1}{2}\int_{B(x,T-t)} \left[|\dot{u}(t,y)|^2 + |\nabla u(t,y)|^2\right] dy.$$

Then $e(t)$ is decreasing for $0 \le t \le T$.

Proof. First recall that

$$\frac{d}{dr}\int_{B(x,r)} f\,dx = \frac{d}{dr}\int_0^r d\rho \int_{|y-x|=\rho} f(y)d\sigma(y) = \int_{\partial B(x,r)} f\,d\sigma.$$

Hence

$$\begin{aligned}\dot{e}(t) &= \frac{d}{dt}\int_{B(x,R-t)} \{|\dot{u}(t,y)|^2 + |\nabla u(t,y)|^2\}dy\\ &= -\frac{1}{2}\int_{\partial B(x,R-t)} (|\dot{u}|^2 + |\nabla u|^2)d\sigma + \int_{B(x,R-t)} [\dot{u}\,\ddot{u} + \nabla u\cdot\nabla\dot{u}]\,dm\\ &= -\frac{1}{2}\int_{\partial B(x,R-t)} (|\dot{u}|^2 + |\nabla u|^2)d\sigma + \int_{B(x,R-t)} [\dot{u}\,\Delta u + \nabla u\cdot\nabla\dot{u}]\,dm\\ &= -\frac{1}{2}\int_{\partial B(x,R-t)} (|\dot{u}|^2 + |\nabla u|^2)d\sigma + 2\int_{\partial B(x,R-t)} \dot{u}\,\frac{\partial u}{\partial n}d\sigma\\ &= \frac{1}{2}\int_{\partial B(x,R-t)} \{2\,\dot{u}\,(\nabla u\cdot n) - (|\dot{u}|^2 + |\nabla u|^2)\}d\sigma \le 0\end{aligned}$$

wherein we have used the elementary estimate,

$$2\,(\nabla u\cdot n)\,\dot{u} \le 2|\nabla u|\,|\dot{u}| \le (|\dot{u}|^2 + |\nabla u|^2).$$

Therefore $e(t) \le e(0) = 0$ for all t i.e. $e(t) := 0$. ■

Corollary 47.9 (Uniqueness of Solutions). *Suppose that u is a classical solution to the wave equation with $u(0,\cdot) = 0 = u_t(0,\cdot)$. Then $u \equiv 0$.*

Proof. Proposition 47.8 shows

$$\frac{1}{2}\int_{B(x,T-t)} \left[|\dot{u}(t,y)|^2 + |\nabla u(t,y)|^2\right] dy = E_{B(x,T)}(0) = 0$$

for all $0 \le t < T$ and $x \in \mathbb{R}^n$. This then implies that $\dot{u}(t,y) = 0$ for all $y \in \mathbb{R}^n$ and $0 \le t \le T$ and hence $u \equiv 0$. ■

Remark 47.10. This result also applies to certain class of weak type solutions in x by first convolving u with an approximate (spatial) delta function, say $u_\epsilon(t,x) = u(t,\cdot) * \delta_\epsilon(x)$. Then u_ϵ satisfies the hypothesis of Corollary 47.9 and hence is 0. Now let $\epsilon \downarrow 0$ to find $u \equiv 0$.

Remark 47.11. Proposition 47.8 also exhibits the finite speed of propagation of the wave equation.

47.6 Wave Equation in Higher Dimensions

47.6.1 Solution derived from the heat kernel

Let

$$p_t^n(x) := \frac{1}{(2\pi t)^{n/2}} e^{-\frac{1}{2t}|x|^2}$$

and simply write p_t for p_t^1. Then

$$2\int_0^\infty \cos\omega t\; p_\lambda(t)dt = \int_{\mathbb{R}} e^{it\omega} p_\lambda(t)dt = e^{-\lambda\partial_t^2/2} e^{it\omega}|_{t=0} = e^{-\lambda\omega^2/2}.$$

Taking $\omega = \sqrt{-\Delta}$ and writing $u(t,x) := \cos\left(\sqrt{-\Delta}t\right) g(x)$ the previous identity gives

$$\begin{aligned}
2\int_0^\infty u(t,x)\,\frac{1}{\sqrt{2\pi\lambda}}e^{-\frac{1}{2\lambda}t^2}dt &= 2\int_0^\infty u(t,x)\;p_\lambda(t)dt \\
&= e^{\lambda\Delta/2}g(x) = \int_{\mathbb{R}^n} p_\lambda^n(y)g(x-y)dy \\
&= \int_{\mathbb{R}^n} \frac{1}{(2\pi\lambda)^{n/2}}e^{-\frac{1}{2\lambda}|y|^2}g(x-y)dy \\
&= \frac{1}{(2\pi\lambda)^{n/2}}\int_0^\infty d\rho e^{-\frac{1}{2\lambda}\rho^2}\int_{|y|=\rho} g(x-y)d\sigma(y) \\
&= \frac{\sigma(S^{n-1})}{(2\pi\lambda)^{n/2}}\int_0^\infty d\rho e^{-\frac{1}{2\lambda}\rho^2}\rho^{n-1}\bar{g}(x;\rho),
\end{aligned}$$

and so

$$\begin{aligned}
\int_0^\infty u(t,x)e^{-\frac{1}{2\lambda}t^2}dt &= \sqrt{\frac{\pi\lambda}{2}}\frac{\sigma(S^{n-1})}{(2\pi\lambda)^{n/2}}\int_0^\infty d\rho e^{-\frac{1}{2\lambda}\rho^2}\rho^{n-1}\bar{g}(x;\rho) \\
&= \sqrt{\frac{\pi}{2}}\frac{\sigma(S^{n-1})}{(2\pi)^{n/2}}\lambda^{-(n-1)/2}\int_0^\infty e^{-\frac{1}{2\lambda}t^2}t^{n-1}\bar{g}(x;t)dt.
\end{aligned}$$

Suppose $n=2k+1$ and let $c_n := \sqrt{\frac{\pi}{2}}\frac{\sigma(S^{n-1})}{(2\pi)^{n/2}}$, then the above equation reads

$$\begin{aligned}
\int_0^\infty u(t,x)e^{-\frac{1}{2\lambda}t^2}dt &= c_n\lambda^{-k}\int_0^\infty e^{-\frac{1}{2\lambda}t^2}t^{2k}\bar{g}(x;t)dt \\
&= c_n\int_0^\infty\left(-\frac{1}{t}\partial_t\right)^k e^{-\frac{1}{2\lambda}t^2}t^{2k}\bar{g}(x;t)dt \\
&\overset{\text{I.B.P.}}{=} c_n\int_0^\infty e^{-\frac{1}{2\lambda}t^2}\left(\partial_t M_{t^{-1}}\right)^k\left[t^{2k}\bar{g}(x;t)\right]dt.
\end{aligned}$$

By the injectivity of the Laplace transform (after making the substitution $t\to\sqrt{t}$, this implies

$$\begin{aligned}
\cos\left(\sqrt{-\Delta}t\right)g(x) &= u(t,x) = c_n\left(\partial_t M_{t^{-1}}\right)^k\left[t^{2k}\bar{g}(x;t)\right] \\
&= c_n\left(\partial_t M_{t^{-1}}\partial_t M_{t^{-1}}\dots\partial_t M_{t^{-1}}\right)\left[t^{2k}\bar{g}(x;t)\right] \\
&= c_n\partial_t\left(\overbrace{M_{t^{-1}}\partial_t M_{t^{-1}}\dots M_{t^{-1}}\partial_t}^{k-1\text{ times}}\right)\left[t^{2k-1}\bar{g}(x;t)\right] \\
&= c_n\partial_t\left(\frac{1}{t}\partial_t\right)^{k-1}\left[t^{2k-1}\bar{g}(x;t)\right].
\end{aligned}$$

Hence we have derived the following theorem.

Theorem 47.12. *Suppose $n=2k+1$ is odd and let $c_n := \sqrt{\frac{\pi}{2}}\frac{\sigma(S^{n-1})}{(2\pi)^{n/2}}$, then*

$$\cos\left(\sqrt{-\Delta}t\right)g(x) = c_n\partial_t\left(\frac{1}{t}\partial_t\right)^{k-1}\left[t^{2k-1}\bar{g}(x;t)\right]$$

and

$$\frac{\sin\left(\sqrt{-\Delta}t\right)}{\sqrt{-\Delta}}f(x) = \int_0^t\cos\left(\sqrt{-\Delta}\tau\right)f(x)d\tau = c_n\left(\frac{1}{t}\partial_t\right)^{k-1}\left[t^{2k-1}\bar{g}(x;t)\right].$$

Proof. For the last equality we have used

$$\left(\frac{1}{t}\partial_t\right)^{k-1}t^{2k-1} = \text{const.}*t^{2k-1-2(k-1)} = \text{const.}*t$$

so that $\left(\frac{1}{t}\partial_t\right)^{k-1}\left[t^{2k-1}\bar{g}(x;t)\right] = O(t)$ and in particular is 0 at $t=0$. ■

47.6.2 Solution derived from the Poisson kernel

Suppose we want to write

$$e^{-|x|} = \int_0^\infty \phi(s)p_s(x)ds.$$

Since

$$\int_{\mathbb{R}} e^{-|x|}e^{i\lambda x}dx = 2\,\mathrm{Re}\int_0^\infty e^{-x}e^{i\lambda x}dx = 2\,\mathrm{Re}\left(\frac{1}{1-i\lambda}\right) = \frac{2}{1+\lambda^2}$$

and

$$\int_{\mathbb{R}} p_s(x)e^{i\lambda x}dx = e^{s\partial_x^2/2}e^{i\lambda x}|_{x=0} = e^{-s\lambda^2/2}$$

ϕ must satisfy

$$\int_0^\infty \phi(s)e^{-s\lambda^2/2}ds = \frac{2}{1+\lambda^2} = \int_0^\infty e^{-s(1+\lambda^2)/2}ds = \int_0^\infty e^{-s/2}e^{-s\lambda^2/2}ds.$$

from which it follows that $\phi(s) = e^{-s/2}$. Thus we have derived the formula

$$e^{-|x|} = \int_0^\infty (2\pi s)^{-1/2}e^{-s/2}e^{-\frac{1}{2s}x^2}ds \tag{47.10}$$

Let : $H\to H$ such that $A=A^*$ and $A\le 0$. By the spectral theorem, we may “substitute” $x=t\sqrt{-A}$ into Eq. (47.10) to learn

$$e^{-t\sqrt{-A}} = \int_0^\infty (2\pi s)^{-1/2}e^{-s/2}e^{\frac{t^2}{2s}A}ds$$

and in particular taking $A=\Delta$ one finds

$$e^{-t\sqrt{-\Delta}} = \int_0^\infty (2\pi s)^{-1/2} e^{-s/2} e^{\frac{t^2}{2s}\Delta} ds$$

from which we conclude the convolution kernel $Q_t(x)$ for $e^{-t\sqrt{-\Delta}}$ is given by

$$\begin{aligned} Q_t(x) &= \int_0^\infty (2\pi s)^{-1/2} e^{-s/2} p^n_{t^2 s^{-1}}(x) ds = \int_0^\infty (2\pi s)^{-1/2} e^{-s/2} \frac{e^{-\frac{s}{2t^2}|x|^2}}{(2\pi t^2 s^{-1})^{n/2}} ds \\ &= (2\pi)^{-1/2} \left(2\pi t^2\right)^{-n/2} \int_0^\infty s^{\frac{n-1}{2}} e^{-s\frac{1}{2}\left(1+\frac{|x|^2}{t^2}\right)} ds \\ &= (2\pi)^{-1/2} \left(2\pi t^2\right)^{-n/2} \int_0^\infty s^{\frac{n+1}{2}} e^{-s\frac{1}{2}\left(1+\frac{|x|^2}{t^2}\right)} \frac{ds}{s}. \end{aligned}$$

Making the substitution, $u = s\frac{1}{2}\left(1 + \frac{|x|^2}{t^2}\right)$ in the previous integral shows

$$\begin{aligned} Q_t(x) &= (2\pi)^{-1/2} \left(2\pi t^2\right)^{-n/2} \left[\frac{1}{2}\left(1 + \frac{|x|^2}{t^2}\right)\right]^{-\frac{n+1}{2}} \int_0^\infty s^{\frac{n+1}{2}} e^{-s} \frac{ds}{s} \\ &= (2\pi)^{-1/2} 2^{\frac{n+1}{2}} (2\pi)^{-n/2} t \left(t^2\right)^{-\frac{n+1}{2}} \left(1 + \frac{|x|^2}{t^2}\right)^{-\frac{n+1}{2}} \Gamma\left(\frac{n+1}{2}\right) \\ &= 2^{\frac{n+1}{2}} (2\pi)^{-\frac{n+1}{2}} \Gamma\left(\frac{n+1}{2}\right) \frac{t}{\left(t^2 + |x|^2\right)^{\frac{n+1}{2}}} \\ &= \Gamma\left(\frac{n+1}{2}\right) \frac{t}{\pi^{\frac{n+1}{2}} \left(t^2 + |x|^2\right)^{\frac{n+1}{2}}}. \end{aligned}$$

Theorem 47.13. *Let*

$$c_n := \frac{\Gamma\left(\frac{n+1}{2}\right)}{\pi^{\frac{n+1}{2}}}$$

$$Q_t(x) = c_n \frac{t}{\left(t^2 + |x|^2\right)^{\frac{n+1}{2}}} \tag{47.11}$$

then

$$e^{-t\sqrt{-\Delta}} f(x) = \int_{\mathbb{R}^n} Q_t(x-y) f(y) dy. \tag{47.12}$$

Notice that if $u(t,x) := e^{-t\sqrt{-\Delta}} f(x)$, we have $\partial_t^2 u(t,x) = \left(\sqrt{-\Delta}\right)^2 u(t,x) = -\Delta u(t,x)$ with $u(0,x) = f(x)$. This explains why Q_t is the same Poisson kernel which we already saw in Eq. (43.36) of Theorem 43.31 above. To match the two results, observe Theorem 43.31 is for "spatial dimension" $n-1$ not n as in Theorem 47.13.

Integrating Eq. (47.12) from t to ∞ then implies

$$\begin{aligned} \frac{1}{\sqrt{-\Delta}} e^{-t\sqrt{-\Delta}} f(x) &= \frac{-1}{\sqrt{-\Delta}} e^{-\tau\sqrt{-\Delta}} f(x)|_{\tau=t}^\infty \\ &= \int_t^\infty e^{-\tau\sqrt{-\Delta}} f(x) d\tau \\ &= \int_{\mathbb{R}^n} \int_t^\infty d\tau Q_\tau(x-y) f(y) dy. \end{aligned}$$

Now

$$\begin{aligned} \int_t^\infty Q_\tau(x-y) d\tau &= c_n \int_t^\infty \frac{\tau}{\left(\tau^2 + |x|^2\right)^{\frac{n+1}{2}}} d\tau = \frac{c_n}{1-n} \left(\tau^2 + |x|^2\right)^{\frac{1-n}{2}} |_{\tau=t}^\infty \\ &= \frac{c_n}{n-1} \left(t^2 + |x|^2\right)^{-\frac{n-1}{2}} \end{aligned}$$

and hence

$$\frac{1}{\sqrt{-\Delta}} e^{-t\sqrt{-\Delta}} f(x) = \int_{\mathbb{R}^n} \frac{c_n}{n-1} \left(t^2 + |y|^2\right)^{-\frac{n-1}{2}} f(x-y) dy$$

and by analytic continuation,

$$\begin{aligned} \frac{1}{\sqrt{-\Delta}} e^{(it-\epsilon)\sqrt{-\Delta}} f(x) &= \frac{1}{\sqrt{-\Delta}} e^{-(\epsilon - it)\sqrt{-\Delta}} f(x) \\ &= \frac{c_n}{n-1} \int_{\mathbb{R}^n} \left((\epsilon - it)^2 + |y|^2\right)^{-\frac{n-1}{2}} f(x-y) dy \\ &= \frac{c_n}{n-1} \int_{\mathbb{R}^n} \left(|y|^2 - (t - i\epsilon)^2\right)^{-\frac{n-1}{2}} f(x-y) dy \end{aligned}$$

and hence

$$\frac{1}{\sqrt{-\Delta}} \sin\left(t\sqrt{-\Delta}\right) f(x) = c_n' \lim_{\epsilon \downarrow 0} \int_{\mathbb{R}^n} \operatorname{Im}\left(|y|^2 - (t - i\epsilon)^2\right)^{-\frac{n-1}{2}} f(x-y) dy.$$

Now if $|y| > |t|$ then

$$\lim_{\epsilon \downarrow 0} \left(|y|^2 - (t - i\epsilon)^2\right)^{-\frac{n-1}{2}} = \left(|y|^2 - t^2\right)^{-\frac{n-1}{2}}$$

is real so

$$\lim_{\epsilon \downarrow 0} \operatorname{Im}\left(|y|^2 - (t - i\epsilon)^2\right)^{-\frac{n-1}{2}} = 0 \text{ if } |y| > |t|.$$

Similarly if n is odd $\lim_{\epsilon\downarrow 0} \left(|y|^2 - (t - i\epsilon)^2\right)^{-\frac{n-1}{2}} = \left(|y|^2 - t^2\right)^{-\frac{n-1}{2}} \in \mathbb{R}$ and so

$$\lim_{\epsilon \downarrow 0} \operatorname{Im}\left(|y|^2 - (t - i\epsilon)^2\right)^{-\frac{n-1}{2}}$$

is a distribution concentrated on the sphere $|y| = |t|$ which is the sharp propagation again. See Taylor Vol. 1., p. 221– 225 for more on this approach. Let us examine here the special case $n = 3$,

$$\mathrm{Im}\left(\frac{1}{|y|^2 - (t - i\epsilon)^2}\right) = \mathrm{Im}\left(\frac{1}{|y|^2 - t^2 + \epsilon^2 + 2i\epsilon t}\right) = \frac{-2\epsilon t}{\left(|y|^2 - t^2 + \epsilon^2\right)^2 + 4\epsilon^2 t^2}$$

so

$$\begin{aligned}
I &:= \lim_{\epsilon \downarrow 0} \int_{\mathbb{R}^n} \mathrm{Im}\left(\frac{1}{|y|^2 - (t - i\epsilon)^2}\right) f(x-y)dy \\
&= \lim_{\epsilon \downarrow 0} \int_{\mathbb{R}^n} \frac{-2\epsilon t}{\left(|y|^2 - t^2 + \epsilon^2\right)^2 + 4\epsilon^2 t^2} f(x-y)dy \\
&= 4\pi \lim_{\epsilon \downarrow 0} \int_0^\infty \rho^2 \frac{-2\epsilon t}{\left(\rho^2 - t^2 + \epsilon^2\right)^2 + 4\epsilon^2 t^2} \bar{f}(x;\rho)d\rho \\
&= ct \lim_{\epsilon \downarrow 0} \int_0^\infty \rho^2 \frac{\epsilon}{\left(\rho^2 - t^2 + \epsilon^2\right)^2 + 4\epsilon^2 t^2} \bar{f}(x;\rho)d\rho.
\end{aligned}$$

Make the change of variables $\rho = t + \epsilon s$ above to find

$$\begin{aligned}
I &= ct \lim_{\epsilon \downarrow 0} \int_{-t/\epsilon}^\infty \frac{(t + \epsilon s)^2 \epsilon^2}{\left(2\epsilon s t + \epsilon^2 s^2 + \epsilon^2\right)^2 + 4\epsilon^2 t^2} \bar{f}(x; t + \epsilon s)ds \\
&= ct \lim_{\epsilon \downarrow 0} \int_{-t/\epsilon}^\infty \frac{(t + \epsilon s)^2}{\left(2st + \epsilon s^2 + \epsilon\right)^2 + 4t^2} \bar{f}(x; t + \epsilon s)ds \\
&= ct\bar{f}(x;t) \int_{-\infty}^\infty \frac{t^2}{4t^2 s^2 + 4t^2} ds = \frac{c}{4} t\bar{f}(x;t) \int_{-\infty}^\infty \frac{1}{s^2+1} ds \\
&= \frac{c}{4}\pi t\bar{f}(x;t)
\end{aligned}$$

which up to an overall constant is the result that we have seen before.

47.7 Explain Method of descent $n = 2$

$$u(t,x) = \frac{1}{2} \fint_{B(x,t)} \frac{t\, g(y) + t^2\, h(y) + t\nabla g(y) \cdot (y - x)}{(t^2 - |y - x|^2)^{1/2}}\, dy.$$

See constant coefficient PDE notes for more details on this.

Part XIV

Sobolev Theory

48

Sobolev Spaces

Definition 48.1. *For $p \in [1,\infty]$, $k \in \mathbb{N}$ and Ω an open subset of $\mathbb{R}^d$, let*

$$W^{k,p}_{loc}(\Omega) := \{f \in L^p(\Omega) : \partial^\alpha f \in L^p_{loc}(\Omega) \text{ (weakly) for all } |\alpha| \le k\},$$

$$W^{k,p}(\Omega) := \{f \in L^p(\Omega) : \partial^\alpha f \in L^p(\Omega) \text{ (weakly) for all } |\alpha| \le k\},$$

$$\|f\|_{W^{k,p}(\Omega)} := \left(\sum_{|\alpha|\le k} \|\partial^\alpha f\|^p_{L^p(\Omega)}\right)^{1/p} \quad \text{if } p < \infty \tag{48.1}$$

and

$$\|f\|_{W^{k,p}(\Omega)} = \sum_{|\alpha|\le k} \|\partial^\alpha f\|_{L^\infty(\Omega)} \quad \text{if } p = \infty. \tag{48.2}$$

In the special case of $p = 2$, we write $W^{k,2}_{loc}(\Omega) =: H^k_{loc}(\Omega)$ and $W^{k,2}(\Omega) =: H^k(\Omega)$ in which case $\|\cdot\|_{W^{k,2}(\Omega)} = \|\cdot\|_{H^k(\Omega)}$ is a Hilbertian norm associated to the inner product

$$(f,g)_{H^k(\Omega)} = \sum_{|\alpha|\le k} \int_\Omega \partial^\alpha f \cdot \overline{\partial^\alpha g}\, dm. \tag{48.3}$$

Theorem 48.2. *The function, $\|\cdot\|_{W^{k,p}(\Omega)}$, is a norm which makes $W^{k,p}(\Omega)$ into a Banach space.*

Proof. Let $f, g \in W^{k,p}(\Omega)$, then the triangle inequality for the p – norms on $L^p(\Omega)$ and $l^p(\{\alpha : |\alpha| \le k\})$ implies

$$\begin{aligned}
\|f+g\|_{W^{k,p}(\Omega)} &= \left(\sum_{|\alpha|\le k} \|\partial^\alpha f + \partial^\alpha g\|^p_{L^p(\Omega)}\right)^{1/p} \\
&\le \left(\sum_{|\alpha|\le k} \left[\|\partial^\alpha f\|_{L^p(\Omega)} + \|\partial^\alpha g\|_{L^p(\Omega)}\right]^p\right)^{1/p} \\
&\le \left(\sum_{|\alpha|\le k} \|\partial^\alpha f\|^p_{L^p(\Omega)}\right)^{1/p} + \left(\sum_{|\alpha|\le k} \|\partial^\alpha g\|^p_{L^p(\Omega)}\right)^{1/p} \\
&= \|f\|_{W^{k,p}(\Omega)} + \|g\|_{W^{k,p}(\Omega)}.
\end{aligned}$$

This shows $\|\cdot\|_{W^{k,p}(\Omega)}$ defined in Eq. (48.1) is a norm. We now show completeness.

If $\{f_n\}_{n=1}^\infty \subset W^{k,p}(\Omega)$ is a Cauchy sequence, then $\{\partial^\alpha f_n\}_{n=1}^\infty$ is a Cauchy sequence in $L^p(\Omega)$ for all $|\alpha| \le k$. By the completeness of $L^p(\Omega)$, there exists $g_\alpha \in L^p(\Omega)$ such that $g_\alpha = L^p\text{–}\lim_{n\to\infty} \partial^\alpha f_n$ for all $|\alpha| \le k$. Therefore, for all $\phi \in C_c^\infty(\Omega)$,

$$\langle f, \partial^\alpha \phi\rangle = \lim_{n\to\infty} \langle f_n, \partial^\alpha \phi\rangle = (-1)^{|\alpha|} \lim_{n\to\infty} \langle \partial^\alpha f_n, \phi\rangle = (-1)^{|\alpha|} \lim_{n\to\infty} \langle g_\alpha, \phi\rangle.$$

This shows $\partial^\alpha f$ exists weakly and $g_\alpha = \partial^\alpha f$ a.e. This shows $f \in W^{k,p}(\Omega)$ and that $f_n \to f \in W^{k,p}(\Omega)$ as $n \to \infty$. ∎

Example 48.3. Let $u(x) := |x|^{-\alpha}$ for $x \in \mathbb{R}^d$ and $\alpha \in \mathbb{R}$. Then

$$\begin{aligned}
\int_{B(0,R)} |u(x)|^p\, dx &= \sigma\left(S^{d-1}\right) \int_0^R \frac{1}{r^{\alpha p}} r^{d-1} dr = \sigma\left(S^{d-1}\right) \int_0^R r^{d-\alpha p-1} dr \\
&= \sigma\left(S^{d-1}\right) \cdot \begin{cases} \frac{R^{d-\alpha p}}{d-\alpha p} & \text{if} \quad d-\alpha p > 0 \\ \infty & \text{otherwise} \end{cases}
\end{aligned} \tag{48.4}$$

and hence $u \in L^p_{loc}(\mathbb{R}^d)$ iff $\alpha < d/p$. Now $\nabla u(x) = -\alpha\, |x|^{-\alpha-1}\, \hat{x}$ where $\hat{x} := x/|x|$. Hence if $\nabla u(x)$ is to exist in $L^p_{loc}(\mathbb{R}^d)$ it is given by $-\alpha\, |x|^{-\alpha-1}\, \hat{x}$ which is in $L^p_{loc}(\mathbb{R}^d)$ iff $\alpha + 1 < d/p$, i.e. if $\alpha < d/p - 1 = \frac{d-p}{p}$. Let us not check that $u \in W^{1,p}_{loc}(\mathbb{R}^d)$ provided $\alpha < d/p - 1$. To do this suppose $\phi \in C_c^\infty(\mathbb{R}^d)$ and $\epsilon > 0$, then

$$\begin{aligned}
-\langle u, \partial_i \phi\rangle &= -\lim_{\epsilon\downarrow 0} \int_{|x|>\epsilon} u(x)\partial_i\phi(x)dx \\
&= \lim_{\epsilon\downarrow 0} \left\{\int_{|x|>\epsilon} \partial_i u(x)\phi(x)dx + \int_{|x|=\epsilon} u(x)\phi(x)\frac{x_i}{\epsilon} d\sigma(x)\right\}.
\end{aligned}$$

Since

$$\left| \int_{|x|=\epsilon} u(x)\phi(x)\frac{x_i}{\epsilon} d\sigma(x) \right| \le \|\phi\|_\infty \sigma\left(S^{d-1}\right) \epsilon^{d-1-\alpha} \to 0 \text{ as } \epsilon \downarrow 0$$

and $\partial_i u(x) = -\alpha |x|^{-\alpha-1} \hat{x} \cdot e_i$ is locally integrable we conclude that

$$-\langle u, \partial_i \phi \rangle = \int_{\mathbb{R}^d} \partial_i u(x)\phi(x) dx$$

showing that the weak derivative $\partial_i u$ exists and is given by the usual pointwise derivative.

48.1 Mollifications

Proposition 48.4 (Mollification). *Let Ω be an open subset of $\mathbb{R}^d$, $k \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$, $p \in [1, \infty)$ and $u \in W^{k,p}_{loc}(\Omega)$. Then there exists $u_n \in C_c^\infty(\Omega)$ such that $u_n \to u$ in $W^{k,p}_{loc}(\Omega)$.*

Proof. Apply Proposition 29.12 with polynomials, $p_\alpha(\xi) = \xi^\alpha$, for $|\alpha| \le k$. ■

Proposition 48.5. *$C_c^\infty(\mathbb{R}^d)$ is dense in $W^{k,p}(\mathbb{R}^d)$ for all $1 \le p < \infty$.*

Proof. The proof is similar to the proof of Proposition 48.4 using Exercise 29.32 in place of Proposition 29.12. ■

Proposition 48.6. *Let Ω be an open subset of $\mathbb{R}^d$, $k \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$ and $p \ge 1$, then*

1. *for any α with $|\alpha| \le k$, $\partial^\alpha : W^{k,p}(\Omega) \to W^{k-|\alpha|,p}(\Omega)$ is a contraction.*
2. *For any open subset $V \subset \Omega$, the restriction map $u \to u|_V$ is bounded from $W^{k,p}(\Omega) \to W^{k,p}(V)$.*
3. *For any $f \in C^k(\Omega)$ and $u \in W^{k,p}_{loc}(\Omega)$, the $fu \in W^{k,p}_{loc}(\Omega)$ and for $|\alpha| \le k$,*

$$\partial^\alpha (fu) = \sum_{\beta \le \alpha} \binom{\alpha}{\beta} \partial^\beta f \cdot \partial^{\alpha-\beta} u \tag{48.5}$$

   where $\binom{\alpha}{\beta} := \frac{\alpha!}{\beta!(\alpha-\beta)!}$.

4. *For any $f \in BC^k(\Omega)$ and $u \in W^{k,p}_{loc}(\Omega)$, the $fu \in W^{k,p}_{loc}(\Omega)$ and for $|\alpha| \le k$ Eq. (48.5) still holds. Moreover, the linear map $u \in W^{k,p}(\Omega) \to fu \in W^{k,p}(\Omega)$ is a bounded operator.*

Proof. 1. Let $\phi \in C_c^\infty(\Omega)$ and $u \in W^{k,p}(\Omega)$, then for β with $|\beta| \le k - |\alpha|$,

$$\langle \partial^\alpha u, \partial^\beta \phi \rangle = (-1)^{|\alpha|} \langle u, \partial^\alpha \partial^\beta \phi \rangle = (-1)^{|\alpha|} \langle u, \partial^{\alpha+\beta} \phi \rangle = (-1)^{|\beta|} \langle \partial^{\alpha+\beta} u, \phi \rangle$$

from which it follows that $\partial^\beta(\partial^\alpha u)$ exists weakly and $\partial^\beta(\partial^\alpha u) = \partial^{\alpha+\beta} u$. This shows that $\partial^\alpha u \in W^{k-|\alpha|,p}(\Omega)$ and it should be clear that $\|\partial^\alpha u\|_{W^{k-|\alpha|,p}(\Omega)} \le \|u\|_{W^{k,p}(\Omega)}$.

Item 2. is trivial.

3 - 4. Given $u \in W^{k,p}_{loc}(\Omega)$, by Proposition 48.4 there exists $u_n \in C_c^\infty(\Omega)$ such that $u_n \to u$ in $W^{k,p}_{loc}(\Omega)$. From the results in Appendix A.1, $fu_n \in C_c^k(\Omega) \subset W^{k,p}(\Omega)$ and

$$\partial^\alpha (fu_n) = \sum_{\beta \le \alpha} \binom{\alpha}{\beta} \partial^\beta f \cdot \partial^{\alpha-\beta} u_n \tag{48.6}$$

holds. Given $V \subset_o \Omega$ such that $\bar{V}$ is compactly contained in Ω, we may use the above equation to find the estimate

$$\begin{aligned} \|\partial^\alpha (fu_n)\|_{L^p(V)} &\le \sum_{\beta \le \alpha} \binom{\alpha}{\beta} \|\partial^\beta f\|_{L^\infty(V)} \|\partial^{\alpha-\beta} u_n\|_{L^p(V)} \\ &\le C_\alpha(f,V) \sum_{\beta \le \alpha} \|\partial^{\alpha-\beta} u_n\|_{L^p(V)} \le C_\alpha(f,V) \|u_n\|_{W^{k,p}(V)} \end{aligned}$$

wherein the last equality we have used Exercise 48.36 below. Summing this equation on $|\alpha| \le k$ shows

$$\|fu_n\|_{W^{k,p}(V)} \le C(f,V) \|u_n\|_{W^{k,p}(V)} \text{ for all } n \tag{48.7}$$

where $C(f,V) := \sum_{|\alpha| \le k} C_\alpha(f,V)$. By replacing u_n by $u_n - u_m$ in the above inequality it follows that $\{fu_n\}_{n=1}^\infty$ is convergent in $W^{k,p}(V)$ and since V was arbitrary $fu_n \to fu$ in $W^{k,p}_{loc}(\Omega)$. Moreover, we may pass to the limit in Eq. (48.6) and in Eq. (48.7) to see that Eq. (48.5) holds and that

$$\|fu\|_{W^{k,p}(V)} \le C(f,V) \|u\|_{W^{k,p}(V)} \le C(f,V) \|u\|_{W^{k,p}(\Omega)}$$

Moreover if $f \in BC(\Omega)$ then constant $C(f,V)$ may be chosen to be independent of V and therefore, if $u \in W^{k,p}(\Omega)$ then $fu \in W^{k,p}(\Omega)$.

Alternative direct proof of 4. We will prove this by induction on $|\alpha|$. If $\alpha = e_i$ then, using Lemma 29.9,

$$\begin{aligned} -\langle fu, \partial_i \phi \rangle &= -\langle u, f\partial_i \phi \rangle = -\langle u, \partial_i [f\phi] - \partial_i f \cdot \phi \rangle \\ &= \langle \partial_i u, f\phi \rangle + \langle u, \partial_i f \cdot \phi \rangle = \langle f\partial_i u + \partial_i f \cdot u, \phi \rangle \end{aligned}$$

showing $\partial_i(fu)$ exists weakly and is equal to $\partial_i(fu) = f\partial_i u + \partial_i f \cdot u \in L^p(\Omega)$. Supposing the result has been proved for all α such that $|\alpha| \le m$ with $m \in [1,k)$. Let $\gamma = \alpha + e_i$ with $|\alpha| = m$, then by what we have just proved each summand in Eq. (48.5) satisfies $\partial_i\left[\partial^\beta f \cdot \partial^{\alpha-\beta} u\right]$ exists weakly and

$$\partial_i\left[\partial^\beta f \cdot \partial^{\alpha-\beta} u\right] = \partial^{\beta+e_i} f \cdot \partial^{\alpha-\beta} u + \partial^{\beta_i} f \cdot \partial^{\alpha-\beta+e_i} u \in L^p(\Omega).$$

Therefore $\partial^\gamma (fu) = \partial_i \partial^\alpha (fu)$ exists weakly in $L^p(\Omega)$ and

$$\partial^\gamma (fu) = \sum_{\beta \le \alpha} \binom{\alpha}{\beta} \left[\partial^{\beta + e_i} f \cdot \partial^{\alpha - \beta} u + \partial^\beta f \cdot \partial^{\alpha - \beta + e_i} u\right] = \sum_{\beta \le \gamma} \binom{\gamma}{\beta} \left[\partial^\beta f \cdot \partial^{\gamma - \beta} u\right]$$

For the last equality see the combinatorics in Appendix A.1. ■

Theorem 48.7. *Let Ω be an open subset of $\mathbb{R}^d$, $k \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$ and $p \in [1, \infty)$. Then $C^\infty(\Omega) \cap W^{k,p}(\Omega)$ is dense in $W^{k,p}(\Omega)$.*

Proof. Let $\Omega_n := \{x \in \Omega : \operatorname{dist}(x, \Omega) > 1/n\} \cap B(0, n)$, then

$$\bar{\Omega}_n \subset \{x \in \Omega : \operatorname{dist}(x, \Omega) \ge 1/n\} \cap \overline{B(0,n)} \subset \Omega_{n+1},$$

$\bar{\Omega}_n$ is compact for every n and $\Omega_n \uparrow \Omega$ as $n \to \infty$. Let $V_0 = \Omega_3$, $V_j := \Omega_{j+3} \setminus \bar{\Omega}_j$ for $j \ge 1$, $K_0 := \bar{\Omega}_2$ and $K_j := \bar{\Omega}_{j+2} \setminus \Omega_{j+1}$ for $j \ge 1$ as in figure 48.1. Then

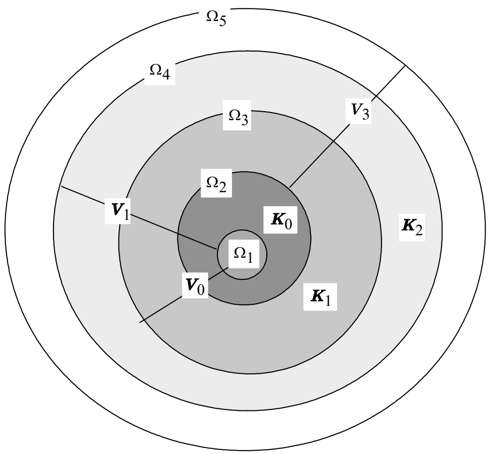


Fig. 48.1. Decomposing Ω into compact pieces. The compact sets K_0, K_1 and K_2 are the shaded annular regions while V_0, V_1 and V_2 are the indicated open annular regions.

$K_n \sqsubset\sqsubset V_n$ for all n and $\cup K_n = \Omega$. Choose $\phi_n \in C_c^\infty(V_n, [0,1])$ such that $\phi_n = 1$ on K_n and set $\psi_0 = \phi_0$ and

$$\psi_j = (1 - \psi_1 - \dots - \psi_{j-1}) \phi_j = \phi_j \prod_{k=1}^{j-1} (1 - \phi_k)$$

for $j \ge 1$. Then $\psi_j \in C_c^\infty(V_n, [0,1])$,

$$1 - \sum_{k=0}^{n} \psi_k = \prod_{k=1}^{n} (1 - \phi_k) \to 0 \text{ as } n \to \infty$$

so that $\sum_{k=0}^\infty \psi_k = 1$ on Ω with the sum being locally finite.

Let $\epsilon > 0$ be given. By Proposition 48.6, $u_n := \psi_n u \in W^{k,p}(\Omega)$ with $\operatorname{supp}(u_n) \sqsubset\sqsubset V_n$. By Proposition 48.4, we may find $v_n \in C_c^\infty(V_n)$ such that $\|u_n - v_n\|_{W^{k,p}(\Omega)} \le \epsilon/2^{n+1}$ for all n. Let $v := \sum_{n=1}^\infty v_n$, then $v \in C^\infty(\Omega)$ because the sum is locally finite. Since

$$\sum_{n=0}^\infty \|u_n - v_n\|_{W^{k,p}(\Omega)} \le \sum_{n=0}^\infty \epsilon/2^{n+1} = \epsilon < \infty,$$

the sum $\sum_{n=0}^\infty (u_n - v_n)$ converges in $W^{k,p}(\Omega)$. The sum, $\sum_{n=0}^\infty (u_n - v_n)$, also converges pointwise to $u - v$ and hence $u - v = \sum_{n=0}^\infty (u_n - v_n)$ is in $W^{k,p}(\Omega)$. Therefore $v \in W^{k,p}(\Omega) \cap C^\infty(\Omega)$ and

$$\|u - v\| \le \sum_{n=0}^\infty \|u_n - v_n\|_{W^{k,p}(\Omega)} \le \epsilon.$$

■

Notation 48.8 *Given a closed subset $F \subset \mathbb{R}^d$, let $C^\infty(F)$ denote those $u \in C(F)$ that extend to a C^∞ – function on an open neighborhood of F.*

Remark 48.9. It is easy to prove that $u \in C^\infty(F)$ iff there exists $U \in C^\infty(\mathbb{R}^d)$ such that $u = U|_F$. Indeed, suppose Ω is an open neighborhood of F, $f \in C^\infty(\Omega)$ and $u = f|_F \in C^\infty(F)$. Using a partition of unity argument (making use of the open sets V_i constructed in the proof of Theorem 48.7), one may show there exists $\phi \in C^\infty(\Omega, [0,1])$ such that $\operatorname{supp}(\phi) \sqsubset \Omega$ and $\phi = 1$ on a neighborhood of F. Then $U := \phi f$ is the desired function.

Theorem 48.10 (Density of $W^{k,p}(\Omega) \cap C^\infty(\bar{\Omega})$ in $W^{k,p}(\Omega)$). *Let $\Omega \subset \mathbb{R}^d$ be a manifold with C^0 – boundary, then for $k \in \mathbb{N}_0$ and $p \in [1, \infty)$, $W^{k,p}(\Omega^o) \cap C^\infty(\bar{\Omega})$ is dense in $W^{k,p}(\Omega^0)$. This may alternatively be stated by assuming $\Omega \subset \mathbb{R}^d$ is an open set such that $\bar{\Omega}^o = \Omega$ and $\bar{\Omega}$ is a manifold with C^0 – boundary, then $W^{k,p}(\Omega) \cap C^\infty(\bar{\Omega})$ is dense in $W^{k,p}(\Omega)$.*

Before going into the proof, let us point out that some restriction on the boundary of Ω is needed for assertion in Theorem 48.10 to be valid. For example, suppose

$$\Omega_0 := \left\{x \in \mathbb{R}^2 : 1 < |x| < 2\right\} \text{ and } \Omega := \Omega_0 \setminus \{(1,2) \times \{0\}\}$$

and $\theta : \Omega \to (0, 2\pi)$ is defined so that $x_1 = |x| \cos\theta(x)$ and $x_2 = |x| \sin\theta(x)$, see Figure 48.2. Then $\theta \in BC^\infty(\Omega) \subset W^{k,\infty}(\Omega)$ for all $k \in \mathbb{N}_0$ yet θ can not be approximated by functions from $C^\infty(\bar{\Omega}) \subset BC^\infty(\Omega_0)$ in $W^{1,p}(\Omega)$. Indeed, if this were possible, it would follows that $\theta \in W^{1,p}(\Omega_0)$. However, θ is

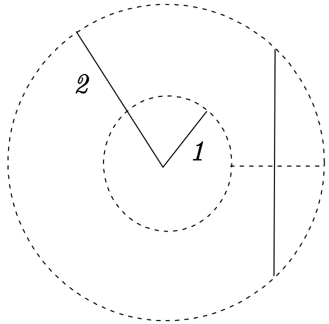


Fig. 48.2. The region Ω_0 along with a vertical in Ω.

not continuous (and hence not absolutely continuous) on the lines $\{x_1 = \rho\} \cap \Omega$ for all $\rho \in (1, 2)$ and so by Theorem 29.30, $\theta \notin W^{1,p}(\Omega_0)$.

The following is a warm-up to the proof of Theorem 48.10.

Proposition 48.11 (Warm-up). *Let $f : \mathbb{R}^{d-1} \to \mathbb{R}$ be a continuous function and $\Omega := \left\{x \in \mathbb{R}^d : x_d > f(x_1, \dots, x_{d-1})\right\}$ and $C^\infty(\bar{\Omega})$ denote those $u \in C\left(\bar{\Omega}\right)$ which are restrictions of C^∞ – functions defined on an open neighborhood of $\bar{\Omega}$. Then for $p \in [1, \infty)$, $C^\infty(\bar{\Omega}) \cap W^{k,p}(\Omega)$ is dense in $W^{k,p}(\Omega)$.*

Proof. By Theorem 48.7, it suffices to show than any $u \in C^\infty(\Omega) \cap W^{k,p}(\Omega)$ may be approximated by elements of $C^\infty\left(\bar{\Omega}\right) \cap W^{k,p}(\Omega)$. For $s > 0$ let $u_s(x) := u(x + se_d)$ which is defined for $x \in \Omega - se_d$. Since

$$\begin{aligned}\bar{\Omega} &= \left\{x \in \mathbb{R}^d : x_d \geq f(x_1, \dots, x_{d-1})\right\} \\ &\subset \left\{x \in \mathbb{R}^d : x_d + s > f(x_1, \dots, x_{d-1})\right\} = \Omega - se_d\end{aligned}$$

and $\partial^\alpha u_s = (\partial^\alpha u)_s$ for all α,

$$u_s \in W^{k,p}(\Omega - se_d) \cap C^\infty(\Omega - se_d) \subset C^\infty\left(\bar{\Omega}\right) \cap W^{k,p}(\Omega).$$

These observations along with the strong continuity of translations in L^p (see Proposition 11.13), implies $\lim_{s \downarrow 0} \|u - u_s\|_{W^{k,p}(\Omega)} = 0$. ■

48.1.1 Proof of Theorem 48.10

Proof. By Theorem 48.7, it suffices to show than any $u \in C^\infty(\Omega) \cap W^{k,p}(\Omega)$ may be approximated by elements of $C^\infty\left(\bar{\Omega}\right) \cap W^{k,p}(\Omega)$. To understand the main ideas of the proof, suppose that Ω is the triangular region in Figure 48.3 and suppose that we have used a partition of unity relative to the cover shown so that $u = u_1 + u_2 + u_3$ with $\operatorname{supp}(u_i) \subset B_i$. Now concentrating on

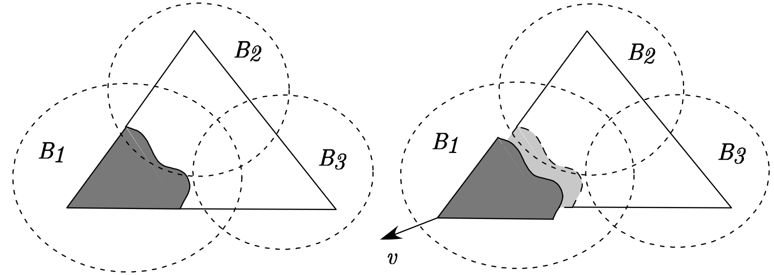


Fig. 48.3. Splitting and moving a function in $C^\infty(\Omega)$ so that the result is in $C^\infty\left(\bar{\Omega}\right)$.

u_1 whose support is depicted as the grey shaded area in Figure 48.3. We now simply translate u_1 in the direction v shown in Figure 48.3. That is for any small $s > 0$, let $w_s(x) := u_1(x + sv)$, then v_s lives on the translated grey area as seen in Figure 48.3. The function w_s extended to be zero off its domain of definition is an element of $C^\infty\left(\bar{\Omega}\right)$ moreover it is easily seen, using the same methods as in the proof of Proposition 48.11, that $w_s \to u_1$ in $W^{k,p}(\Omega)$.

The formal proof follows along these same lines. To do this choose an at most countable locally finite cover $\{V_i\}_{i=0}^\infty$ of $\bar{\Omega}$ such that $\bar{V}_0 \subset \Omega$ and for each $i \geq 1$, after making an affine change of coordinates, $V_i = (-\epsilon, \epsilon)^d$ for some $\epsilon > 0$ and

$$V_i \cap \bar{\Omega} = \{(y, z) \in V_i : \epsilon > z > f_i(y)\}$$

where $f_i : (-\epsilon, \epsilon)^{d-1} \to (-\epsilon, \epsilon)$, see Figure 48.4 below. Let $\{\eta_i\}_{i=0}^\infty$ be a partition of unity subordinated to $\{V_i\}$ and let $u_i := u\eta_i \in C^\infty(V_i \cap \Omega)$. Given $\delta > 0$, we choose s so small that $w_i(x) := u_i(x + se_d)$ (extended to be zero off its domain of definition) may be viewed as an element of $C^\infty(\bar{\Omega})$ and such

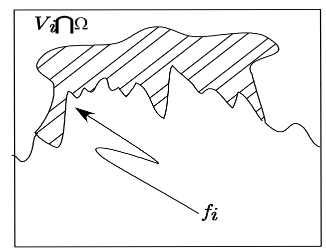


Fig. 48.4. The shaded area depicts the support of $u_i = u\eta_i$.

that $\|u_i - w_i\|_{W^{k,p}(\Omega)} < \delta/2^i$. For $i = 0$ we set $w_0 := u_0 = u\eta_0$. Then, since $\{V_i\}_{i=1}^\infty$ is a locally finite cover of $\bar{\Omega}$, it follows that $w := \sum_{i=0}^\infty w_i \in C^\infty\left(\bar{\Omega}\right)$ and further we have

$$\sum_{i=0}^\infty \|u_i - w_i\|_{W^{k,p}(\Omega)} \le \sum_{i=1}^\infty \delta/2^i = \delta.$$

This shows

$$u - w = \sum_{i=0}^\infty (u_i - w_i) \in W^{k,p}(\Omega)$$

and $\|u - w\|_{W^{k,p}(\Omega)} < \delta$. Hence $w \in C^\infty\left(\bar{\Omega}\right) \cap W^{k,p}(\Omega)$ is a δ – approximation of u and since $\delta > 0$ arbitrary the proof is complete. ■

48.2 Difference quotients

Recall from Notation 29.14 that for $h \neq 0$

$$\partial_i^h u(x) := \frac{u(x + he^i) - u(x)}{h}.$$

Remark 48.12 (Adjoints of Finite Differences). For $u \in L^p$ and $g \in L^q$,

$$\begin{aligned}\int_{\mathbb{R}^d} \partial_i^h u(x)\ g(x)\ dx &= \int_{\mathbb{R}^d} \frac{u(x + he_i) - u(x)}{h}\ g(x)\ dx \\ &= -\int_{\mathbb{R}^d} u(x)\ \frac{g(x - he_i) - g(x)}{-h}\ dx \\ &= -\int_{\mathbb{R}^d} u(x)\partial_i^{-h} g(x)\ dx.\end{aligned}$$

We summarize this identity by $(\partial_i^h)^* = -\partial_i^{-h}$.

Theorem 48.13. *Suppose $k \in \mathbb{N}_0$, Ω is an open subset of $\mathbb{R}^d$ and V is an open precompact subset of Ω.*

1. *If $1 \le p < \infty$, $u \in W^{k,p}(\Omega)$ and $\partial_i u \in W^{k,p}(\Omega)$, then*
$$\|\partial_i^h u\|_{W^{k,p}(V)} \le \|\partial_i u\|_{W^{k,p}(\Omega)} \tag{48.8}$$
for all $0 < |h| < \frac{1}{2}\mathrm{dist}(V, \Omega^c)$.
2. *Suppose that $1 < p \le \infty$, $u \in W^{k,p}(\Omega)$ and assume there exists a constant $C(V) < \infty$ such that*
$$\|\partial_i^h u\|_{W^{k,p}(V)} \le C(V) \text{ for all } 0 < |h| < \frac{1}{2}\mathrm{dist}(V, \Omega^c).$$

Then $\partial_i u \in W^{k,p}(V)$ and $\|\partial_i u\|_{W^{k,p}(V)} \le C(V)$. Moreover if $C := \sup_{V \subset\subset \Omega} C(V) < \infty$ then in fact $\partial_i u \in W^{k,p}(\Omega)$ and there is a constant $c < \infty$ such that

$$\|\partial_i u\|_{W^{k,p}(\Omega)} \le c\left(C + \|u\|_{L^p(\Omega)}\right).$$

Proof. 1. Let $|\alpha| \le k$, then

$$\|\partial^\alpha \partial_i^h u\|_{L^p(V)} = \|\partial_i^h \partial^\alpha u\|_{L^p(V)} \le \|\partial_i \partial^\alpha u\|_{L^p(\Omega)}$$

wherein we have used Theorem 29.22 for the last inequality. Eq. (48.8) now easily follows.

2. If $\|\partial_i^h u\|_{W^{k,p}(V)} \le C(V)$ then for all $|\alpha| \le k$,

$$\|\partial_i^h \partial^\alpha u\|_{L^p(V)} = \|\partial^\alpha \partial_i^h u\|_{L^p(V)} \le C(V).$$

So by Theorem 29.22, $\partial_i \partial^\alpha u \in L^p(V)$ and $\|\partial_i \partial^\alpha u\|_{L^p(V)} \le C(V)$. From this we conclude that $\|\partial^\beta u\|_{L^p(V)} \le C(V)$ for all $0 < |\beta| \le k+1$ and hence $\|u\|_{W^{k+1,p}(V)} \le c\left[C(V) + \|u\|_{L^p(V)}\right]$ for some constant c. ■

Notation 48.14 *Given a multi-index α and $h \neq 0$, let*

$$\partial_h^\alpha := \prod_{i=1}^d \left(\partial_i^h\right)^{\alpha_i}.$$

The following theorem is a generalization of Theorem 48.13.

Theorem 48.15. *Suppose $k \in \mathbb{N}_0$, Ω is an open subset of $\mathbb{R}^d$, V is an open precompact subset of Ω and $u \in W^{k,p}(\Omega)$.*

1. *If $1 \le p < \infty$ and $|\alpha| \le k$, then $\|\partial_h^\alpha u\|_{W^{k-|\alpha|}(V)} \le \|u\|_{W^{k,p}(\Omega)}$ for h small.*
2. *If $1 < p \le \infty$ and $\|\partial_h^\alpha u\|_{W^{k,p}(V)} \le C$ for all $|\alpha| \le j$ and h near 0, then $u \in W^{k+j,p}(V)$ and $\|\partial^\alpha u\|_{W^{k,p}(V)} \le C$ for all $|\alpha| \le j$.*

Proof. Since $\partial_h^\alpha = \prod_i \partial_h^{\alpha_i}$, item 1. follows from Item 1. of Theorem 48.13 and induction on $|\alpha|$.

For Item 2., suppose first that $k = 0$ so that $u \in L^p(\Omega)$ and $\|\partial_h^\alpha u\|_{L^p(V)} \le C$ for $|\alpha| \le j$. Then by Proposition 29.16, there exists $\{h_l\}_{l=1}^\infty \subset \mathbb{R} \setminus \{0\}$ and $v \in L^p(V)$ such that $h_l \to 0$ and $\lim_{l\to\infty} \langle \partial_{h_\ell}^\alpha u, \phi\rangle = \langle v, \phi\rangle$ for all $\phi \in C_c^\infty(V)$. Using Remark 48.12,

$$\langle v, \phi\rangle = \lim_{l\to\infty} \langle \partial_{h_\ell}^\alpha u, \phi\rangle = (-1)^{|\alpha|} \lim_{l\to\infty} \langle u, \partial_{-h_\ell}^\alpha \phi\rangle = (-1)^{|\alpha|} \langle u, \partial^\alpha \phi\rangle$$

which shows $\partial^\alpha u = v \in L^p(V)$. Moreover, since weak convergence decreases norms,

$$\|\partial^\alpha u\|_{L^p(V)} = \|v\|_{L^p(V)} \le C.$$

For the general case if $k \in \mathbb{N}$, $u \in W^{k,p}(\Omega)$ such that $\|\partial_h^\alpha u\|_{W^{k,p}(V)} \le C$, then (for $p \in (1,\infty)$, the case $p = \infty$ is similar and left to the reader)

$$\sum_{|\beta|\le k} \|\partial_h^\alpha \partial^\beta u\|_{L^p(V)}^p = \sum_{|\beta|\le k} \|\partial^\beta \partial_h^\alpha u\|_{L^p(V)}^p = \|\partial_h^\alpha u\|_{W^{k,p}(V)}^p \le C^p.$$

As above this implies $\partial^\alpha\partial^\beta u \in L^p(V)$ for all $|\alpha| \le j$ and $|\beta| \le k$ and that

$$\|\partial^\alpha u\|_{W^{k,p}(V)}^p = \sum_{|\beta|\le k} \|\partial^\alpha\partial^\beta u\|_{L^p(V)}^p \le C^p.$$

■

48.3 Sobolev Spaces on Compact Manifolds

Theorem 48.16 (Change of Variables). *Suppose that U and V are open subsets of $\mathbb{R}^d$, $T \in C^k(U,V)$ be a C^k – diffeomorphism such that $\|\partial^\alpha T\|_{BC(U)} < \infty$ for all $1 \le |\alpha| \le k$ and $\epsilon := \inf_U |\det T'| > 0$. Then the map $T^* : W^{k,p}(V) \to W^{k,p}(U)$ defined by $u \in W^{k,p}(V) \to T^*u := u \circ T \in W^{k,p}(U)$ is well defined and is bounded.*

Proof. For $u \in W^{k,p}(V) \cap C^\infty(V)$, repeated use of the chain and product rule implies,

$$\begin{aligned}
(u\circ T)' &= (u'\circ T)\,T' \\
(u\circ T)'' &= (u'\circ T)'\,T' + (u'\circ T)\,T'' = (u''\circ T)\,T'\otimes T' + (u'\circ T)\,T'' \\
(u\circ T)^{(3)} &= \left(u^{(3)}\circ T\right) T'\otimes T'\otimes T' + (u''\circ T)\,(T'\otimes T')' \\
&\quad + (u''\circ T)\,T'\otimes T'' + (u'\circ T)\,T^{(3)} \\
&\;\;\vdots \\
(u\circ T)^{(l)} &= \left(u^{(l)}\circ T\right) \overbrace{T\otimes\cdots\otimes T}^{l \text{ times}} + \sum_{j=1}^{l-1} \left(u^{(j)}\circ T\right) p_j\left(T', T'', \dots, T^{(l+1-j)}\right).
\end{aligned} \tag{48.9}$$

This equation and the boundedness assumptions on $T^{(j)}$ for $1 \le j \le k$ implies there is a finite constant K such that

$$\left|(u\circ T)^{(l)}\right| \le K \sum_{j=1}^{l} \left|u^{(j)}\circ T\right| \text{ for all } 1 \le l \le k.$$

By Hölder's inequality for sums we conclude there is a constant K_p such that

$$\sum_{|\alpha|\le k} |\partial^\alpha (u\circ T)|^p \le K_p \sum_{|\alpha|\le k} |\partial^\alpha u|^p \circ T$$

and therefore

$$\|u\circ T\|_{W^{k,p}(U)}^p \le K_p \sum_{|\alpha|\le k} \int_U |\partial^\alpha u|^p (T(x))\, dx.$$

Making the change of variables, $y = T(x)$ and using

$$dy = |\det T'(x)|\, dx \ge \epsilon dx,$$

we find

$$\begin{aligned}
\|u\circ T\|_{W^{k,p}(U)}^p &\le K_p \sum_{|\alpha|\le k} \int_U |\partial^\alpha u|^p (T(x))\, dx \\
&\le \frac{K_p}{\epsilon} \sum_{|\alpha|\le k} \int_V |\partial^\alpha u|^p (y)\, dy = \frac{K_p}{\epsilon} \|u\|_{W^{k,p}(V)}^p.
\end{aligned} \tag{48.10}$$

This shows that $T^* : W^{k,p}(V) \cap C^\infty(V) \to W^{k,p}(U) \cap C^\infty(U)$ is a bounded operator. For general $u \in W^{k,p}(V)$, we may choose $u_n \in W^{k,p}(V) \cap C^\infty(V)$ such that $u_n \to u$ in $W^{k,p}(V)$. Since T^* is bounded, it follows that T^*u_n is Cauchy in $W^{k,p}(U)$ and hence convergent. Finally, using the change of variables theorem again we know,

$$\|T^*u - T^*u_n\|_{L^p(V)}^p \le \epsilon^{-1} \|u - u_n\|_{L^p(U)}^p \to 0 \text{ as } n \to \infty$$

and therefore $T^*u = \lim_{n\to\infty} T^*u_n$ and by continuity Eq. (48.10) still holds for $u \in W^{k,p}(V)$. ■

Let M be a compact C^k – manifolds without boundary, i.e. M is a compact Hausdorff space with a collection of charts x in an "atlas" $\mathcal{A}$ such that $x : D(x) \subset_o M \to R(x) \subset_o \mathbb{R}^d$ is a homeomorphism such that

$$x \circ y^{-1} \in C^k\left(y\left(D(x)\cap D(y)\right), x\left(D(x)\cap D(y)\right)\right) \text{ for all } x, y \in \mathcal{A}.$$

Definition 48.17. *Let $\{x_i\}_{i=1}^N \subset \mathcal{A}$ such that $M = \cup_{i=1}^N D(x_i)$ and let $\{\phi_i\}_{i=1}^N$ be a partition of unity subordinate do the cover $\{D(x_i)\}_{i=1}^N$. We now define $u \in W^{k,p}(M)$ if $u : M \to \mathbb{C}$ is a function such that*

$$\|u\|_{W^{k,p}(M)} := \sum_{i=1}^N \left\|(\phi_i u)\circ x_i^{-1}\right\|_{W^{k,p}(R(x_i))} < \infty. \tag{48.11}$$

Since $\|\cdot\|_{W^{k,p}(R(x_i))}$ is a norm for all i, it easily verified that $\|\cdot\|_{W^{k,p}(M)}$ is a norm on $W^{k,p}(M)$.

Proposition 48.18. *If $f \in C^k(M)$ and $u \in W^{k,p}(M)$ then $fu \in W^{k,p}(M)$ and*

$$\|fu\|_{W^{k,p}(M)} \le C\|u\|_{W^{k,p}(M)} \tag{48.12}$$

where C is a finite constant not depending on u. Recall that $f : M \to \mathbb{R}$ is said to be C^j with $j \le k$ if $f \circ x^{-1} \in C^j(R(x), \mathbb{R})$ for all $x \in \mathcal{A}$.

Proof. Since $\left[f \circ x_i^{-1}\right]$ has bounded derivatives on $\mathrm{supp}(\phi_i \circ x_i^{-1})$, it follows from Proposition 48.6 that there is a constant $C_i < \infty$ such that

$$\begin{aligned}\left\|(\phi_i f u) \circ x_i^{-1}\right\|_{W^{k,p}(R(x_i))} &= \left\|\left[f \circ x_i^{-1}\right](\phi_i u) \circ x_i^{-1}\right\|_{W^{k,p}(R(x_i))} \\ &\le C_i \left\|(\phi_i u) \circ x_i^{-1}\right\|_{W^{k,p}(R(x_i))}\end{aligned}$$

and summing this equation on i shows Eq. (48.12) holds with $C := \max_i C_i$. ■

Theorem 48.19. *If $\{y_j\}_{j=1}^K \subset \mathcal{A}$ such that $M = \cup_{j=1}^K D(y_j)$ and $\{\psi_j\}_{j=1}^K$ is a partition of unity subordinate to the cover $\{D(y_j)\}_{j=1}^K$, then the norm*

$$|u|_{W^{k,p}(M)} := \sum_{j=1}^K \left\|(\psi_j u) \circ y_j^{-1}\right\|_{W^{k,p}(R(y_j))} \tag{48.13}$$

is equivalent to the norm in Eq. (48.11). That is to say the space $W^{k,p}(M)$ along with its topology is well defined independent of the choice of charts and partitions of unity used in defining the norm on $W^{k,p}(M)$.

Proof. Since $|\cdot|_{W^{k,p}(M)}$ is a norm,

$$\begin{aligned}|u|_{W^{k,p}(M)} &= \left|\sum_{i=1}^N \phi_i u\right|_{W^{k,p}(M)} \le \sum_{i=1}^N |\phi_i u|_{W^{k,p}(M)} \\ &= \sum_{j=1}^K \left\|\sum_{i=1}^N (\psi_j \phi_i u) \circ y_j^{-1}\right\|_{W^{k,p}(R(y_j))} \\ &\le \sum_{j=1}^K \sum_{i=1}^N \left\|(\psi_j \phi_i u) \circ y_j^{-1}\right\|_{W^{k,p}(R(y_j))}\end{aligned} \tag{48.14}$$

and since $x_i \circ y_j^{-1}$ and $y_j \circ x_i^{-1}$ are C^k diffeomorphism and the sets $y_j(\mathrm{supp}(\phi_i) \cap \mathrm{supp}(\psi_j))$ and $x_i(\mathrm{supp}(\phi_i) \cap \mathrm{supp}(\psi_j))$ are compact, an application of Theorem 48.16 and Proposition 48.6 shows there are finite constants C_{ij} such that

$$\begin{aligned}\left\|(\psi_j \phi_i u) \circ y_j^{-1}\right\|_{W^{k,p}(R(y_j))} &\le C_{ij} \left\|(\psi_j \phi_i u) \circ x_i^{-1}\right\|_{W^{k,p}(R(x_i))} \\ &\le C_{ij} \left\|\phi_i u \circ x_i^{-1}\right\|_{W^{k,p}(R(x_i))}\end{aligned}$$

which combined with Eq. (48.14) implies

$$|u|_{W^{k,p}(M)} \le \sum_{j=1}^K \sum_{i=1}^N C_{ij} \left\|\phi_i u \circ x_i^{-1}\right\|_{W^{k,p}(R(x_i))} \le C\|u\|_{W^{k,p}(M)}$$

where $C := \max_i \sum_{j=1}^K C_{ij} < \infty$. Analogously, one shows there is a constant $K < \infty$ such that $\|u\|_{W^{k,p}(M)} \le K|u|_{W^{k,p}(M)}$. ■

Lemma 48.20. *Suppose $x \in \mathcal{A}(M)$ and $U \subset_o M$ such that $U \subset \bar{U} \subset D(x)$, then there is a constant $C < \infty$ such that*

$$\left\|u \circ x^{-1}\right\|_{W^{k,p}(x(U))} \le C\|u\|_{W^{k,p}(M)} \text{ for all } u \in W^{k,p}(M). \tag{48.15}$$

Conversely a function $u : M \to \mathbb{C}$ with $\mathrm{supp}(u) \subset U$ is in $W^{k,p}(M)$ iff $\left\|u \circ x^{-1}\right\|_{W^{k,p}(x(U))} < \infty$ and in any case there is a finite constant such that

$$\|u\|_{W^{k,p}(M)} \le C\left\|u \circ x^{-1}\right\|_{W^{k,p}(x(U))}. \tag{48.16}$$

Proof. Choose charts $y_1 := x, y_2, \ldots, y_K \in \mathcal{A}$ such that $\{D(y_i)\}_{j=1}^K$ is an open cover of M and choose a partition of unity $\{\psi_j\}_{j=1}^K$ subordinate to the cover $\{D(y_j)\}_{j=1}^K$ such that $\psi_1 = 1$ on a neighborhood of $\bar{U}$. To construct such a partition of unity choose $U_j \subset_o M$ such that $U_j \subset \bar{U}_j \subset D(y_j)$, $\bar{U} \subset U_1$ and $\cup_{j=1}^K U_j = M$ and for each j let $\eta_j \in C_c^k(D(y_j), [0,1])$ such that $\eta_j = 1$ on a neighborhood of $\bar{U}_j$. Then define $\psi_j := \eta_j(1-\eta_0)\cdots(1-\eta_{j-1})$ where by convention $\eta_0 \equiv 0$. Then $\{\psi_j\}_{j=1}^K$ is the desired partition, indeed by induction one shows

$$1 - \sum_{j=1}^l \psi_j = (1-\eta_1)\cdots(1-\eta_l)$$

and in particular

$$1 - \sum_{j=1}^K \psi_j = (1-\eta_1)\cdots(1-\eta_K) = 0.$$

Using Theorem 48.19, it follows that

$$\begin{aligned}\left\|u \circ x^{-1}\right\|_{W^{k,p}(x(U))} &= \left\|(\psi_1 u) \circ x^{-1}\right\|_{W^{k,p}(x(U))} \\ &\le \left\|(\psi_1 u) \circ x^{-1}\right\|_{W^{k,p}(R(y_1))} \\ &\le \sum_{j=1}^K \left\|(\psi_j u) \circ y_j^{-1}\right\|_{W^{k,p}(R(y_j))} \\ &= |u|_{W^{k,p}(M)} \le C\|u\|_{W^{k,p}(M)}\end{aligned}$$

which proves Eq. (48.15).

Using Theorems 48.19 and 48.16 there are constants C_j for $j = 0, 1, 2 \ldots, N$ such that

$$\begin{aligned}\|u\|_{W^{k,p}(M)} &\le C_0 \sum_{j=1}^{K} \left\|(\psi_j u)\circ y_j^{-1}\right\|_{W^{k,p}(R(y_j))} \\ &= C_0 \sum_{j=1}^{K} \left\|(\psi_j u)\circ y_1^{-1}\circ y_1 \circ y_j^{-1}\right\|_{W^{k,p}(R(y_j))} \\ &\le C_0 \sum_{j=1}^{K} C_j \left\|(\psi_j u)\circ x^{-1}\right\|_{W^{k,p}(R(y_1))} \\ &= C_0 \sum_{j=1}^{K} C_j \left\|\psi_j \circ x^{-1}\cdot u\circ x^{-1}\right\|_{W^{k,p}(R(y_1))}.\end{aligned}$$

This inequality along with K – applications of Proposition 48.6 proves Eq. (48.16). ■

Theorem 48.21. *The space $(W^{k,p}(M), \|\cdot\|_{W^{k,p}(M)})$ is a Banach space.*

Proof. Let $\{x_i\}_{i=1}^N \subset \mathcal{A}$ and $\{\phi_i\}_{i=1}^N$ be as in Definition 48.17 and choose $U_i \subset_o M$ such that $\operatorname{supp}(\phi_i) \subset U_i \subset \bar{U}_i \subset D(x_i)$. If $\{u_n\}_{n=1}^\infty \subset W^{k,p}(M)$ is a Cauchy sequence, then by Lemma 48.20, $\left\{u_n \circ x_i^{-1}\right\}_{n=1}^\infty \subset W^{k,p}(x_i(U_i))$ is a Cauchy sequence for all i. Since $W^{k,p}(x_i(U_i))$ is complete, there exists $v_i \in W^{k,p}(x_i(U_i))$ such that $u_n \circ x_i^{-1} \to \tilde{v}_i$ in $W^{k,p}(x_i(U_i))$. For each i let $v_i := \phi_i(\tilde{v}_i \circ x_i)$ and notice by Lemma 48.20 that

$$\|v_i\|_{W^{k,p}(M)} \le C\left\|v_i \circ x_i^{-1}\right\|_{W^{k,p}(x_i(U_i))} = C\|\tilde{v}_i\|_{W^{k,p}(x_i(U_i))} < \infty$$

so that $u := \sum_{i=1}^N v_i \in W^{k,p}(M)$. Since $\operatorname{supp}(v_i - \phi_i u_n) \subset U_i$, it follows that

$$\begin{aligned}\|u - u_n\|_{W^{k,p}(M)} &= \left\|\sum_{i=1}^N v_i - \sum_{i=1}^N \phi_i u_n\right\|_{W^{k,p}(M)} \\ &\le \sum_{i=1}^N \|v_i - \phi_i u_n\|_{W^{k,p}(M)} \\ &\le C\sum_{i=1}^N \left\|[\phi_i(\tilde{v}_i \circ x_i - u_n)]\circ x_i^{-1}\right\|_{W^{k,p}(x_i(U_i))} \\ &= C\sum_{i=1}^N \left\|\left[\phi_i\circ x_i^{-1}\left(\tilde{v}_i - u_n\circ x_i^{-1}\right)\right]\right\|_{W^{k,p}(x_i(U_i))} \\ &\le C\sum_{i=1}^N C_i\left\|\tilde{v}_i - u_n \circ x_i^{-1}\right\|_{W^{k,p}(x_i(U_i))} \to 0 \text{ as } n\to\infty\end{aligned}$$

wherein the last inequality we have used Proposition 48.6 again. ■

48.4 Trace Theorems

For many more general results on this subject matter, see E. Stein [17, Chapter VI].

Notation 48.22 *Let $\mathbb{H}^d := \left\{x \in \mathbb{R}^d : x_d > 0\right\}$ be the open upper half space inside of $\mathbb{R}^d$ and if $D > 0$ let*

$$\mathbb{H}^d_D := \left\{x \in \mathbb{H}^d : 0 < x_d < D\right\}.$$

Lemma 48.23. *Suppose $k \ge 1$ and $D > 0$.*

1. *If $p \in [1,\infty)$ and $C_c^k\left(\overline{\mathbb{H}^d}\right)$,then for all $\alpha \in \mathbb{N}_0^{d-1}\times\{0\} \subset \mathbb{N}_0^d$ with $|\alpha| \le k-1$,*

$$\|\partial^\alpha u\|_{L^p(\partial\mathbb{H}^d)} \le D^{-1/p}\|\partial^\alpha u\|_{L^p(\mathbb{H}^d_D)} + \frac{D^{\frac{p-1}{p}}}{p^{1/p}}\left\|\partial^{\alpha+e_d}u\right\|_{L^p(\mathbb{H}^d_D)}. \quad (48.17)$$

In particular there there is a constant $C = C(p,k,D,d)$ such that

$$\|u\|_{W^{k-1,p}(\partial\mathbb{H}^d)} \le C(p,D,k,d)\|u\|_{W^{k,p}(\mathbb{H}^d)}. \quad (48.18)$$

2. *For $p = \infty$ and $u \in W^{k,\infty}\left(\mathbb{H}^d\right)$, there is a continuous version $\tilde{u}$ of u. The function $\tilde{u}$ is in $BC^{k-1}\left(\mathbb{H}^d\right)$ and has the property that $\partial^\alpha\tilde{u}$ extends to a continuous function $v_\alpha \in BC\left(\bar{\mathbb{H}}\right)$ for all $|\alpha| < k$ and the function $\tilde{u}|_{\partial\mathbb{H}^d}$ is in $BC^{k-1}\left(\partial\bar{\mathbb{H}}\right)$ and*

$$\|\tilde{u}\|_{BC^{k-1}(\partial\mathbb{H}^d)} \le \|u\|_{W^{k,\infty}(\mathbb{H}^d_D)} \text{ for any } D > 0.$$

Proof. 1. Write $x \in \overline{\mathbb{H}^d}$ as $x = (y,z) \in \mathbb{R}^{d-1}\times[0,\infty)$ and suppose $\alpha \in \mathbb{N}_0^{d-1}\times\{0\}\subset\mathbb{N}_0^d$ with $|\alpha| \le k-1$. Then by the fundamental theorem of calculus,

$$\partial^\alpha u(y,0) = \partial^\alpha u(y,z) - \int_0^z \partial^\alpha u_t(y,t)dt \quad (48.19)$$

which implies

$$|\partial^\alpha u(y,0)|\, 1_{[0,D]}(z) \le |\partial^\alpha u(y,z)|\, 1_{[0,D]}(z) + 1_{[0,D]}(z)\int_0^z |\partial^\alpha u_t(y,t)|\, dt.$$

Taking the $L^p(\mathbb{H}^d_D)$ – norm of this last equation implies

$$\|\partial^\alpha u\|_{L^p(\partial\mathbb{H}^d_D)}\cdot D^{1/p} \le \|\partial^\alpha u\|_{L^p(\mathbb{H}^d_D)} + B$$

where

$$\begin{aligned} B^p &= \int_{\mathbb{H}^d_D} \left(1_{[0,D]}(z) \int_0^z |\partial^\alpha u_t(y,t)|\, dt \right)^p dydz \\ &\leq \int_{\mathbb{H}^d_D} \left(1_{[0,D]}(z) \cdot z^{p/q} \int_0^z |\partial^\alpha u_t(y,t)|^p\, dt \right) dydz \\ &\leq \int_{\mathbb{H}^d_D} \left(1_{[0,D]}(z) \cdot z^{p/q} \int_0^D |\partial^\alpha u_t(y,t)|^p\, dt \right) dydz \\ &= \left\| \partial^{\alpha+e_d} u \right\|^p_{L^p(\mathbb{H}^d_D)} \frac{D^{p/q+1}}{p/q+1} = \frac{D^p}{p} \left\| \partial^{\alpha+e_d} u \right\|^p_{L^p(\mathbb{H}^d_D)}. \end{aligned}$$

Putting these two equations together shows

$$\| \partial^\alpha u \|_{L^p(\partial \mathbb{H}^d_D)} \leq D^{-1/p} \left[\| \partial^\alpha u \|_{L^p(\mathbb{H}^d_D)} + \frac{D}{p^{1/p}} \left\| \partial^{\alpha+e_d} u \right\|_{L^p(\mathbb{H}^d_D)} \right]$$

which is the same as Eq. (48.17).

Suppose that $p = \infty$ and $u \in W^{k,\infty}\left(\mathbb{H}^d\right)$. By Proposition 29.29, we know that $\partial^\alpha u$ has a Lipschitz continuous version v_α on $\mathbb{H}^d$ for each $|\alpha| < k$. Being Lipschitz, each v_α has a unique extension to a continuous function $\overline{\mathbb{H}^d}$. Let

$$\eta \in C_c^\infty\left(B(0,1) \cap \left(-\mathbb{H}^d\right), [0,\infty)\right)$$

be chosen so that $\int_{\mathbb{R}^n} \eta(x) dx = 1$ and $\eta_m(x) = m^n \eta(mx)$ and let $u_m := u * \eta_m = v_o * \eta_m$. Since $\operatorname{supp}(\eta) \subset \left(-\mathbb{H}^d\right)$ as in Figure 48.5, $u_m \in C^\infty\left(\overline{\mathbb{H}^d}\right)$,

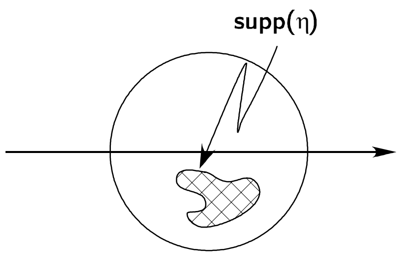


Fig. 48.5. The support of η.

$\partial^\alpha u_m = \partial^\alpha u * \eta_m = v_\alpha * \eta_m$ for all $|\alpha| < k$ and

$$\|u_m - u_n\|_{BC^{k-1}(\overline{\mathbb{H}^d} \cap B(0,R))} \to 0 \text{ as } m,n \to \infty.$$

Therefore $\partial^\alpha u_m \to v_\alpha$ uniformly on compact subsets of $\overline{\mathbb{H}^d}$ for all $|\alpha| < k$. Hence $v_0 \in C^{k-1}\left(\mathbb{H}^d\right)$ and $\partial^\alpha v_0 = v_\alpha$ extends to $\overline{\mathbb{H}^d}$ for all $|\alpha| < k$ and $v_0|_{\partial \mathbb{H}^d} \in BC^{k-1}\left(\partial \bar{\mathbb{H}}\right)$. ■

Theorem 48.24 (Trace Theorem). *Suppose $k \geq 1$ and $\Omega \subset_o \mathbb{R}^d$ such that $\bar{\Omega}$ is a compact manifold with C^k – boundary. Then there exists a unique linear map $T : W^{k,p}(\Omega) \to W^{k-1,p}(\partial\Omega)$ such that $Tu = u|_{\partial\Omega}$ for all $u \in C^k\left(\bar{\Omega}\right)$.*

Proof. Choose a covering $\{V_i\}_{i=0}^N$ of $\bar{\Omega}$ such that $\bar{V}_0 \subset \Omega$ and for each $i \geq 1$, there is C^k – diffeomorphism $x_i : V_i \to R(x_i) \subset_o \mathbb{R}^d$ such that

$$\begin{aligned} x_i\left(\partial\Omega \cap V_i\right) &= R(x_i) \cap \operatorname{bd}(\mathbb{H}^d) \text{ and} \\ x_i\left(\Omega \cap V_i\right) &= R(x_i) \cap \mathbb{H}^d \end{aligned}$$

as in Figure 48.6. Further choose $\phi_i \in C_c^\infty\left(V_i, [0,1]\right)$ such that $\sum_{i=0}^N \phi_i = 1$

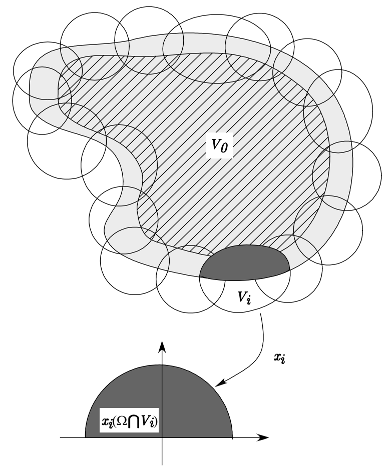


Fig. 48.6. Covering Ω (the shaded region) as described in the text.

on a neighborhood of $\bar{\Omega}$ and set $y_i := x_i|_{\partial\Omega \cap V_i}$ for $i \geq 1$. Given $u \in C^k\left(\bar{\Omega}\right)$ if $p < \infty$ and $u \in W^{k,\infty}(\Omega)$ if $p = \infty$, we compute

$$\begin{aligned}\|u|_{\partial\bar{\Omega}}\|_{W^{k-1,p}(\partial\bar{\Omega})} &= \sum_{i=1}^{N}\left\|(\phi_i u)\,|_{\partial\bar{\Omega}}\circ y_i^{-1}\right\|_{W^{k-1,p}(R(x_i)\cap\mathrm{bd}(\mathbb{H}^d))}\\ &= \sum_{i=1}^{N}\left\|\left[(\phi_i u)\circ x_i^{-1}\right]|_{\mathrm{bd}(\mathbb{H}^d)}\right\|_{W^{k-1,p}(R(x_i)\cap\mathrm{bd}(\mathbb{H}^d))}\\ &\le \sum_{i=1}^{N} C_i\left\|\left[(\phi_i u)\circ x_i^{-1}\right]\right\|_{W^{k,p}(R(x_i))}\\ &\le \max C_i\cdot\sum_{i=1}^{N}\left\|\left[(\phi_i u)\circ x_i^{-1}\right]\right\|_{W^{k,p}(R(x_i)\cap\mathbb{H}^d)}\\ &\quad+\left\|\left[(\phi_0 u)\circ x_0^{-1}\right]\right\|_{W^{k,p}(R(x_0))}\\ &\le C\,\|u\|_{W^{k,p}(\Omega)}\end{aligned}$$

where $C=\max\{1,C_1,\dots,C_N\}$. The proof is complete if $p=\infty$ and follows by the B.L.T. Theorem 2.68 and the fact that $C^k(\bar{\Omega})$ is dense inside $W^{k,p}(\Omega)$ if $p<\infty$. ■

Notation 48.25 *In the sequel will often abuse notation and simply write $u|_{\partial\bar{\Omega}}$ for the "function" $Tu\in W^{k-1,p}(\partial\bar{\Omega})$.*

Proposition 48.26 (Integration by parts). *Suppose $\Omega\subset_o\mathbb{R}^d$ such that $\bar{\Omega}$ is a compact manifold with C^1 – boundary, $p\in[1,\infty]$ and $q=\frac{p}{p-1}$ is the conjugate exponent. Then for $u\in W^{1,p}(\Omega)$ and $v\in W^{1,q}(\Omega)$,*

$$\int_\Omega \partial_i u\cdot v dm=-\int_\Omega u\cdot\partial_i v dm+\int_{\partial\bar{\Omega}} u|_{\partial\bar{\Omega}}\cdot v|_{\partial\bar{\Omega}} n_i d\sigma \tag{48.20}$$

where $n:\partial\bar{\Omega}\to\mathbb{R}^d$ is unit outward pointing norm to $\partial\bar{\Omega}$.

Proof. Equation 48.20 holds for $u,v\in C^2(\bar{\Omega})$ and therefore for $(u,v)\in W^{k,p}(\Omega)\times W^{k,q}(\Omega)$ since both sides of the equality are continuous in $(u,v)\in W^{k,p}(\Omega)\times W^{k,q}(\Omega)$ as the reader should verify.

Warning BRUCE: We might need $p\in(1,\infty)$ here. To fix this, I think if $p=1$ one should replace u by $u_M:=\psi_M(u)$ where $\psi_M(x)=\int_0^x\alpha(y/m)dy$ and $\alpha\in C_c(\mathbb{R},[0,1)$ such that $\alpha=1$ on $[-1,1]$. Then $u_M\in W^{1,\infty}(\Omega)$ and $u_M\to u$ in $W^{1,1}(\Omega)$ and hence the argument given above goes through. We need only approximate $v\in W^{1,q}(\Omega)$ with $q<\infty$ now. We should then pass to the limit as $M\to\infty$. ■

Definition 48.27. *Let $W_0^{k,p}(\Omega):=\overline{C_c^\infty(\Omega)}^{W^{k,p}(\Omega)}$ be the closure of $C_c^\infty(\Omega)$ inside $W^{k,p}(\Omega)$.*

Remark 48.28. Notice that if $T:W^{k,p}(\Omega)\to W^{k-1,p}(\partial\bar{\Omega})$ is the trace operator in Theorem 48.24, then $T\left(W_0^{k,p}(\Omega)\right)=\{0\}\subset W^{k-1,p}(\partial\bar{\Omega})$ since $Tu=u|_{\partial\bar{\Omega}}=0$ for all $u\in C_c^\infty(\Omega)$.

Corollary 48.29. *Suppose $\Omega\subset_o\mathbb{R}^d$ such that $\bar{\Omega}$ is a compact manifold with C^1 – boundary, $p\in[1,\infty)$ and $T:W^{1,p}(\Omega)\to L^p(\partial\Omega)$ is the trace operator of Theorem 48.24. Then $W_0^{1,p}(\Omega)=\mathrm{Nul}(T)$.*

Proof. It has already been observed in Remark 48.28 that $W_0^{1,p}(\Omega)\subset\mathrm{Nul}(T)$. Suppose $u\in\mathrm{Nul}(T)$ and $\mathrm{supp}(u)$ is compactly contained in Ω. The mollification $u_\epsilon(x)$ defined in Proposition 48.4 will be in $C_c^\infty(\Omega)$ for $\epsilon>0$ sufficiently small and by Proposition 48.4, $u_\epsilon\to u$ in $W^{1,p}(\Omega)$. Thus $u\in W_0^{1,p}(\Omega)$. So to finish the proof that $\mathrm{Nul}(T)\subset W_0^{1,p}(\Omega)$, it suffices to show every $u\in W_0^{1,p}(\Omega)$ may be approximated by $v\in W_0^{1,p}(\Omega)$ such that $\mathrm{supp}(v)$ is compactly contained in Ω. Two proofs of this last assertion will now be given.

Proof 1. For $u\in\mathrm{Nul}(T)\subset W^{1,p}(\Omega)$ define

$$\tilde{u}(x)=\begin{cases}u(x) & \text{for } x\in\bar{\Omega}\\ 0 & \text{for } x\notin\bar{\Omega}.\end{cases}$$

Then clearly $\tilde{u}\in L^p(\mathbb{R}^d)$ and moreover by Proposition 48.26, for $v\in C_c^\infty(\mathbb{R}^d)$,

$$\int_{\mathbb{R}^d}\tilde{u}\cdot\partial_i v dm=\int_\Omega u\cdot\partial_i v dm=-\int_\Omega\partial_i u\cdot v dm$$

from which it follows that $\partial_i\tilde{u}$ exists weakly in $L^p(\mathbb{R}^d)$ and $\partial_i\tilde{u}=1_\Omega\partial_i u$ a.e.. Thus $\tilde{u}\in W^{1,p}(\mathbb{R}^d)$ with $\|\tilde{u}\|_{W^{1,p}(\mathbb{R}^d)}=\|u\|_{W^{1,p}(\Omega)}$ and $\mathrm{supp}(\tilde{u})\subset\bar{\Omega}$. (The reader should compare this result with Proposition 48.30 below.)

Choose $V\in C_c^1(\mathbb{R}^d,\mathbb{R}^d)$ such that $V(x)\cdot n(x)>0$ for all $x\in\partial\bar{\Omega}$ and define

$$\tilde{u}_\epsilon(x)=T_\epsilon\tilde{u}(x):=\tilde{u}\circ e^{\epsilon V}(x).$$

Notice that $\mathrm{supp}(\tilde{u}_\epsilon)\subset e^{-\epsilon V}(\bar{\Omega})\sqsubset\sqsubset\Omega$ for all ϵ sufficiently small. By the change of variables Theorem 48.16, we know that $\tilde{u}_\epsilon\in W^{1,p}(\Omega)$ and since $\mathrm{supp}(\tilde{u}_\epsilon)$ is a compact subset of Ω, it follows from the first paragraph that $\tilde{u}_\epsilon\in W_0^{1,p}(\Omega)$.

To so finish this proof, it only remains to show $\tilde{u}_\epsilon\to u$ in $W^{1,p}(\Omega)$ as $\epsilon\downarrow 0$. Looking at the proof of Theorem 48.16, the reader may show there are constants $\delta>0$ and $C<\infty$ such that

$$\|T_\epsilon v\|_{W^{1,p}(\mathbb{R}^d)}\le C\,\|v\|_{W^{1,p}(\mathbb{R}^d)} \text{ for all } v\in W^{1,p}(\mathbb{R}^d). \tag{48.21}$$

By direct computation along with the dominated convergence it may be shown that

$$T_\epsilon v\to v \text{ in } W^{1,p}(\mathbb{R}^d) \text{ for all } v\in C_c^\infty(\mathbb{R}^d). \tag{48.22}$$

As is now standard, Eqs. (48.21) and (48.22) along with the density of $C_c^\infty(\mathbb{R}^d)$ in $W^{1,p}(\mathbb{R}^d)$ allows us to conclude $T_\epsilon v\to v$ in $W^{1,p}(\mathbb{R}^d)$ for all $v\in W^{1,p}(\mathbb{R}^d)$ which completes the proof that $\tilde{u}_\epsilon\to u$ in $W^{1,p}(\Omega)$ as $\epsilon\to 0$.

Proof 2. As in the first proof it suffices to show that any $u\in W_0^{1,p}(\Omega)$ may be approximated by $v\in W^{1,p}(\Omega)$ with $\mathrm{supp}(v)\sqsubset\Omega$. As above extend u

to Ω^c by 0 so that $\tilde{u} \in W^{1,p}\left(\mathbb{R}^d\right)$. Using the notation in the proof of 48.24, it suffices to show $u_i := \phi_i \tilde{u} \in W^{1,p}\left(\mathbb{R}^d\right)$ may be approximated by $u_i \in W^{1,p}(\Omega)$ with $\operatorname{supp}(u_i) \sqsubset \Omega$. Using the change of variables Theorem 48.16, the problem may be reduced to working with $w_i = u_i \circ x_i^{-1}$ on $B = R(x_i)$. But in this case we need only define $w_i^\epsilon(y) := w_i^\epsilon(y - \epsilon e_d)$ for $\epsilon > 0$ sufficiently small. Then $\operatorname{supp}(w_i^\epsilon) \subset \mathbb{H}^d \cap B$ and as we have already seen $w_i^\epsilon \to w_i$ in $W^{1,p}\left(\mathbb{H}^d\right)$. Thus $u_i^\epsilon := w_i^\epsilon \circ x_i \in W^{1,p}(\Omega)$, $u_i^\epsilon \to u_i$ as $\epsilon \downarrow 0$ with $\operatorname{supp}(u_i) \sqsubset \Omega$. ■

48.5 Extension Theorems

Proposition 48.30. *Let $k \in \mathbb{N}_0$, $p \in [1,\infty]$ and suppose Ω is **any** open subset of $\mathbb{R}^d$. Then the **extension by zero map,***

$$u \in W_0^{k,p}(\Omega) \to 1_\Omega u \in W^{k,p}(\mathbb{R}^d),$$

is a contraction. Recall $W_0^{k,p}(\Omega)$ was defined in Definition 48.27) above.

Proof. The result holds for $u \in C_c^\infty(\Omega)$ and hence for all $u \in W_0^{k,p}(\Omega)$. ■

Lemma 48.31. *Let $R > 0$, $B := B(0,R) \subset \mathbb{R}^d$, $B^\pm := \{x \in B : \pm x_d > 0\}$ and $\Gamma := \{x \in B : x_d = 0\}$. Suppose that $u \in C^k(B \setminus \Gamma) \cap C(B)$ and for each $|\alpha| \le k$, $\partial^\alpha u$ extends to a continuous function v_α on B. Then $u \in C^k(B)$ and $\partial^\alpha u = v_\alpha$ for all $|\alpha| \le k$.*

Proof. For $x \in \Gamma$ and $i < d$, then by continuity, the fundamental theorem of calculus and the dominated convergence theorem,

$$\begin{aligned} u(x + \Delta e_i) - u(x) &= \lim_{\substack{y \to x \\ y \in B \setminus \Gamma}} [u(y + \Delta e_i) - u(y)] = \lim_{\substack{y \to x \\ y \in B \setminus \Gamma}} \int_0^\Delta \partial_i u(y + s e_i) ds \\ &= \lim_{\substack{y \to x \\ y \in B \setminus \Gamma}} \int_0^\Delta v_{e_i}(y + s e_i) ds = \int_0^\Delta v_{e_i}(x + s e_i) ds \end{aligned}$$

and similarly, for $i = d$,

$$\begin{aligned} u(x + \Delta e_d) - u(x) &= \lim_{\substack{y \to x \\ y \in B^{\operatorname{sgn}(\Delta)} \setminus \Gamma}} [u(y + \Delta e_d) - u(y)] \\ &= \lim_{\substack{y \to x \\ y \in B^{\operatorname{sgn}(\Delta)} \setminus \Gamma}} \int_0^\Delta \partial_d u(y + s e_d) ds \\ &= \lim_{\substack{y \to x \\ y \in B^{\operatorname{sgn}(\Delta)} \setminus \Gamma}} \int_0^\Delta v_{e_d}(y + s e_d) ds = \int_0^\Delta v_{e_d}(x + s e_d) ds. \end{aligned}$$

These two equations show, for each i, $\partial_i u(x)$ exits and $\partial_i u(x) = v_{e_i}(x)$. Hence we have shown $u \in C^1(B)$.

Suppose it has been proven for some $l \ge 1$ that $\partial^\alpha u(x)$ exists and is given by $v_\alpha(x)$ for all $|\alpha| \le l < k$. Then applying the results of the previous paragraph to $\partial^\alpha u(x)$ with $|\alpha| = l$ shows that $\partial_i \partial^\alpha u(x)$ exits and is given by $v_{\alpha + e_i}(x)$ for all i and $x \in B$ and from this we conclude that $\partial^\alpha u(x)$ exists and is given by $v_\alpha(x)$ for all $|\alpha| \le l + 1$. So by induction we conclude $\partial^\alpha u(x)$ exists and is given by $v_\alpha(x)$ for all $|\alpha| \le k$, i.e. $u \in C^k(B)$. ■

Lemma 48.32. *Given any $k+1$ distinct points, $\{c_i\}_{i=0}^k$, in $\mathbb{R} \setminus \{0\}$, the $(k+1) \times (k+1)$ matrix C with entries $C_{ij} := (c_i)^j$ is invertible.*

Proof. Let $a \in \mathbb{R}^{k+1}$ and define $p(x) := \sum_{j=0}^k a_j x^j$. If $a \in \operatorname{Nul}(C)$, then

$$0 = \sum_{j=0}^k (c_i)^j a_j = p(c_i) \text{ for } i = 0, 1, \ldots, k.$$

Since $\deg(p) \le k$ and the above equation says that p has $k+1$ distinct roots, we conclude that $a \in \operatorname{Nul}(C)$ implies $p \equiv 0$ which implies $a = 0$. Therefore $\operatorname{Nul}(C) = \{0\}$ and C is invertible. ■

Lemma 48.33. *Let B, $B^\pm$ and Γ be as in Lemma 48.31 and $\{c_i\}_{i=0}^k$, be $k+1$ distinct points in $(\infty, -1]$ for example $c_i = -(i+1)$ will work. Also let $a \in \mathbb{R}^{k+1}$ be the unique solution (see Lemma 48.32 to $C^{tr} a = \mathbf{1}$ where $\mathbf{1}$ denotes the vector of all ones in $\mathbb{R}^{k+1}$, i.e. a satisfies*

$$1 = \sum_{i=0}^k (c_i)^j a_i \text{ for } j = 0, 1, 2 \ldots, k. \tag{48.23}$$

For $u \in C_c^k(\overline{\mathbb{H}^d})$[1] with $\operatorname{supp}(u) \subset B \cap \overline{\mathbb{H}^d}$ and $x = (y,z) \in \mathbb{R}^d$ define

$$\tilde{u}(x) = \tilde{u}(y,z) = \begin{cases} u(y,z) & \text{if } z \ge 0 \\ \sum_{i=0}^k a_i u(y, c_i z) & \text{if } z \le 0. \end{cases} \tag{48.24}$$

Then $\tilde{u} \in C_c^k(\mathbb{R}^d)$ with $\operatorname{supp}(\tilde{u}) \subset B$ and moreover there exists a constant M independent of u such that

$$\|\tilde{u}\|_{W^{k,p}(B)} \le M \|u\|_{W^{k,p}(B^+)}. \tag{48.25}$$

Proof. By Eq. (48.23) with $j = 0$,

$$\sum_{i=0}^k a_i u(y, c_i 0) = u(y,0) \sum_{i=0}^k a_i = u(y,0).$$

[1] Or more generally, one may assume $u \in C^k(\mathbb{H}^d) \cap C_c\left(\overline{\mathbb{H}^d}\right)$ such that each $\partial^\alpha u$ for $|\alpha| \le k$ extends to a continuous function on $\overline{\mathbb{H}^d}$.

This shows that $\tilde{u}$ in Eq. (48.24) is well defined and that $\tilde{u} \in C\left(\mathbb{H}^d\right)$. Let $K^- := \{(y,z) : (y,-z) \in \operatorname{supp}(u)\}$. Since $c_i \in (\infty, -1]$, if $x = (y,z) \notin K^-$ and $z < 0$ then $(y, c_i z) \notin \operatorname{supp}(u)$ and therefore $\tilde{u}(x) = 0$ and therefore $\operatorname{supp}(\tilde{u})$ is compactly contained inside of B. Similarly if $\alpha \in \mathbb{N}_0^d$ with $|\alpha| \leq k$, Eq. (48.23) with $j = \alpha_d$ implies

$$v_\alpha(x) := \begin{cases} \left(\partial^\alpha u\right)(y,z) & \text{if } z \geq 0 \\ \sum_{i=0}^k a_i c_i^{\alpha_d} \left(\partial^\alpha u\right)(y, c_i z) & \text{if } z \leq 0. \end{cases}$$

is well defined and $v_\alpha \in C\left(\mathbb{R}^d\right)$. Differentiating Eq. (48.24) shows $\partial^\alpha \tilde{u}(x) = v_\alpha(x)$ for $x \in B \setminus \Gamma$ and therefore we may conclude from Lemma 48.31 that $\tilde{u} \in C_c^k(B) \subset C^k\left(\mathbb{R}^d\right)$ and $\partial^\alpha \tilde{u} = v_\alpha$ for all $|\alpha| \leq k$.

We now verify Eq. (48.25) as follows. For $|\alpha| \leq k$,

$$\begin{aligned} \|\partial^\alpha \tilde{u}\|_{L^p(B^-)}^p &= \int_{\mathbb{R}^d} 1_{z<0} \left| \sum_{i=0}^k a_i c_i^{\alpha_d} \left(\partial^\alpha u\right)(y, c_i z) \right|^p dydz \\ &\leq C \int_{\mathbb{R}^d} 1_{z<0} \sum_{i=0}^k \left|\left(\partial^\alpha u\right)(y, c_i z)\right|^p dydz \\ &= C \int_{\mathbb{R}^d} 1_{z>0} \sum_{i=0}^k \frac{1}{|c_i|} \left|\left(\partial^\alpha u\right)(y, z)\right|^p dydz \\ &= C \left(\sum_{i=0}^k \frac{1}{|c_i|} \right) \|\partial^\alpha u\|_{L^p(B^+)}^p \end{aligned}$$

where $C := \left(\sum_{i=0}^k |a_i c_i^{\alpha_d}|^q \right)^{p/q}$. Summing this equation on $|\alpha| \leq k$ shows there exists a constant M' such that $\|\tilde{u}\|_{W^{k,p}(B^-)} \leq M' \|u\|_{W^{k,p}(B^+)}$ and hence Eq. (48.25) holds with $M = M' + 1$. ■

Corollary 48.34. *Let $k \geq 1$, B, $B^\pm$, $\{c_i\}_{i=0}^k$ be as in Lemma 48.33 and suppose that $u \in W^{k,\infty}(\mathbb{H}^d)$ with $\operatorname{supp}(u) \subset B \cap \overline{\mathbb{H}^d}$. By item 2. of Lemma 48.23, by modifying u on a null set we may assume that $u \in BC^{k-1}\left(\mathbb{H}^d\right)$ with $\partial^\alpha u \in C\left(\overline{\mathbb{H}^d}\right)$ for all $|\alpha| < k$. Then the function $\tilde{u}$ defined in Eq. (48.24) in $W^{k,\infty}(\mathbb{R}^d)$, with $\operatorname{supp}(\tilde{u}) \subset B$ and Eq. (48.25) is still valid.*

Proof. By Lemma 48.33, $\tilde{u} \in C^{k-1}(\mathbb{R}^d)$. Let $\phi \in C_c^\infty\left(\mathbb{R}^d\right)$ and $|\alpha| = k-1$ and $i \in \{1, 2, \ldots, d\}$, then by standard integration by parts

$$\begin{aligned} \langle \tilde{u}, \partial^\alpha \partial_i \phi \rangle &= (-1)^{|\alpha|} \langle \partial^\alpha \tilde{u}, \partial_i \phi \rangle \\ &= (-1)^{|\alpha|} \langle 1_{B^+} \partial^\alpha \tilde{u}, \partial_i \phi \rangle + (-1)^{|\alpha|} \langle 1_{B^-} \partial^\alpha \tilde{u}, \partial_i \phi \rangle. \end{aligned}$$

Making use of Proposition 48.26 and the change of variables theorem,

$$\begin{aligned} \langle \tilde{u}, \partial^\alpha \partial_i \phi \rangle &= -(-1)^{|\alpha|} \langle 1_{B^+} \partial_i \partial^\alpha \tilde{u}, \phi \rangle - (-1)^{|\alpha|} \langle 1_{B^-} \partial_i \partial^\alpha \tilde{u}, \phi \rangle \\ &= -(-1)^{|\alpha|} \langle \partial_i \partial^\alpha \tilde{u}, \phi \rangle \end{aligned}$$

wherein we have used the fact that $\partial^\alpha \tilde{u}|_{\partial B^+} = \partial^\alpha \tilde{u}|_{\partial B^-}$ for all $|\alpha| < k$ to see that the boundary terms from the integrals cancel. Hence it follows that $\partial^\alpha \partial_i \tilde{u}$ exists weakly and is given by the expected formula, namely by differentiating Eq. (48.24) away from Γ and piecing the results together. The verification of Eq. (48.25) is as before. ■

Theorem 48.35 (Extension Theorem). *Suppose $k \geq 1$ and $\Omega \subset_o \mathbb{R}^d$ such that $\bar{\Omega}$ is a compact manifold with C^k – boundary. Given $U \subset_o \mathbb{R}^d$ such that $\bar{\Omega} \subset U$, there exists a bounded linear (extension) operator $E : W^{k,p}(\Omega) \to W^{k,p}\left(\mathbb{R}^d\right)$ such that*

1. *$Eu = u$ a.e. in Ω and*
2. *$\operatorname{supp}(Eu) \subset U$.*

Proof. As in the proof of Theorem 48.24, choose a covering $\{V_i\}_{i=0}^N$ of $\bar{\Omega}$ such that $\bar{V}_0 \subset \Omega$, $\cup_{i=0}^N \bar{V}_i \subset U$ and for each $i \geq 1$, there is C^k – diffeomorphism $x_i : V_i \to R(x_i) \subset_o \mathbb{R}^d$ such that

$$x_i\left(\partial\Omega \cap V_i\right) = R(x_i) \cap \operatorname{bd}(\mathbb{H}^d) \text{ and } x_i\left(\Omega \cap V_i\right) = R(x_i) \cap \mathbb{H}^d = B^+$$

where B^+ is as in Lemma 48.33 and Corollary 48.34, refer to Figure 48.6. Further choose $\phi_i \in C_c^\infty\left(V_i, [0,1]\right)$ such that $\sum_{i=0}^N \phi_i = 1$ on a neighborhood of $\bar{\Omega}$ and set $y_i := x_i|_{\partial\Omega \cap V_i}$ for $i \geq 1$. Given $u \in C^k\left(\bar{\Omega}\right)$ if $p < \infty$ ($u \in W^{k,\infty}(\Omega)$ if $p = \infty$) and $i \geq 1$, the function $v_i := (\phi_i u) \circ x_i^{-1}$ may be viewed as a function in $C^k(\mathbb{H}^d) \cap C_c(\overline{\mathbb{H}^d})$ ($W^{k,\infty}(\mathbb{H}^d)$) with $\operatorname{supp}(u) \subset B$. Let $\tilde{v}_i \in C_c^k(B)$ $\left(W^{k,\infty}(B)\right)$ be defined as in Eq. (48.24) above and define $\tilde{u} := \phi_0 u + \sum_{i=1}^N \tilde{v}_i \circ x_i \in C_c^k\left(\mathbb{R}^d\right)$ $\left(W^{k,\infty}\left(\mathbb{R}^d\right)\right)$. Notice that $\tilde{u} = u$ on $\bar{\Omega}$, $\operatorname{supp}(u) \subset U$ and by Lemma 48.20,

$$\begin{aligned} \|\tilde{u}\|_{W^{k,p}(\mathbb{R}^d)} &\leq \|\phi_0 u\|_{W^{k,p}(\mathbb{R}^d)} + \sum_{i=1}^N \|\tilde{v}_i \circ x_i\|_{W^{k,p}(\mathbb{R}^d)} \\ &\leq \|\phi_0 u\|_{W^{k,p}(\Omega)} + \sum_{i=1}^N \|\tilde{v}_i\|_{W^{k,p}(R(x_i))} \\ &\leq C\left(\phi_0\right) \|u\|_{W^{k,p}(\Omega)} + \sum_{i=1}^N \|v_i\|_{W^{k,p}(B^+)} \\ &= C\left(\phi_0\right) \|u\|_{W^{k,p}(\Omega)} + \sum_{i=1}^N \left\|(\phi_i u) \circ x_i^{-1}\right\|_{W^{k,p}(B^+)} \\ &\leq C\left(\phi_0\right) \|u\|_{W^{k,p}(\Omega)} + \sum_{i=1}^N C_i \|u\|_{W^{k,p}(\Omega)}. \end{aligned}$$

This completes the proof for $p=\infty$ and shows for $p<\infty$ that the map $u \in C^k(\bar{\Omega}) \to Eu := \tilde{u} \in C_c^k(U)$ is bounded as map from $W^{k,p}(\Omega)$ to $W^{k,p}(U)$. As usual, we now extend E using the B.L.T. Theorem 2.68 to a bounded linear map from $W^{k,p}(\Omega)$ to $W^{k,p}(U)$. So for general $u \in W^{k,p}(\Omega)$, $Eu = W^{k,p}(U)$ – $\lim_{n\to\infty} \tilde{u}_n$ where $u_n \in C^k(\bar{\Omega})$ and $u = W^{k,p}(\Omega)$ – $\lim_{n\to\infty} u_n$. By passing to a subsequence if necessary, we may assume that $\tilde{u}_n$ converges a.e. to Eu from which it follows that $Eu = u$ a.e. on $\bar{\Omega}$ and $\operatorname{supp}(Eu) \subset U$. ■

48.6 Exercises

Exercise 48.36. Show the norm in Eq. (48.1) is equivalent to the norm

$$|f|_{W^{k,p}(\Omega)} := \sum_{|\alpha|\le k} \|\partial^\alpha f\|_{L^p(\Omega)}.$$

Solution 48.37. 48.36This is a consequence of the fact that all norms on $l^p(\{\alpha : |\alpha| \le k\})$ are equivalent. To be more explicit, let $a_\alpha = \|\partial^\alpha f\|_{L^p(\Omega)}$, then

$$\sum_{|\alpha|\le k} |a_\alpha| \le \left(\sum_{|\alpha|\le k} |a_\alpha|^p\right)^{1/p} \left(\sum_{|\alpha|\le k} 1^q\right)^{1/q}$$

while

$$\left(\sum_{|\alpha|\le k} |a_\alpha|^p\right)^{1/p} \le \left(\sum_{|\alpha|\le k}^{p} \left[\sum_{|\beta|\le k} |a_\beta|\right]^p\right)^{1/p} \le \left[\#\{\alpha : |\alpha| \le k\}\right]^{1/p} \sum_{|\beta|\le k} |a_\beta|.$$

49

Sobolev Inequalities

49.1 Morrey's Inequality

Notation 49.1 *Let S^{d-1} be the sphere of radius one centered at zero inside $\mathbb{R}^d$. For a set $\Gamma \subset S^{d-1}$, $x \in \mathbb{R}^d$, and $r \in (0,\infty)$, let*

$$\Gamma_{x,r} \equiv \{x + s\omega : \omega \in \Gamma \text{ such that } 0 \le s \le r\}.$$

So $\Gamma_{x,r} = x + \Gamma_{0,r}$ where $\Gamma_{0,r}$ is a cone based on Γ, see Figure 49.1 below.

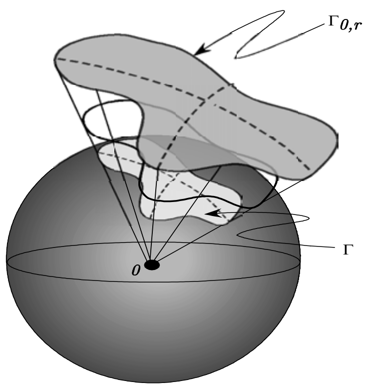


Fig. 49.1. The cone $\Gamma_{0,r}$.

Notation 49.2 *If $\Gamma \subset S^{d-1}$ is a measurable set let $|\Gamma| = \sigma(\Gamma)$ be the surface "area" of Γ.*

Notation 49.3 *If $\Omega \subset \mathbb{R}^d$ is a measurable set and $f : \mathbb{R}^d \to \mathbb{C}$ is a measurable function let*

$$f_\Omega := \fint_\Omega f(x)dx := \frac{1}{m(\Omega)} \int_\Omega f(x)dx.$$

By Theorem 9.35,

$$\int_{\Gamma_{x,r}} f(y)dy = \int_{\Gamma_{0,r}} f(x+y)dy = \int_0^r dt\, t^{d-1} \int_\Gamma f(x+t\omega)\, d\sigma(\omega) \tag{49.1}$$

and letting $f = 1$ in this equation implies

$$m(\Gamma_{x,r}) = |\Gamma|\, r^d/d. \tag{49.2}$$

Lemma 49.4. *Let $\Gamma \subset S^{d-1}$ be a measurable set such that $|\Gamma| > 0$. For $u \in C^1(\overline{\Gamma}_{x,r})$,*

$$\fint_{\Gamma_{x,r}} |u(y) - u(x)|dy \le \frac{1}{|\Gamma|} \int_{\Gamma_{x,r}} \frac{|\nabla u(y)|}{|x-y|^{d-1}}\, dy. \tag{49.3}$$

Proof. Write $y = x + s\omega$ with $\omega \in S^{d-1}$, then by the fundamental theorem of calculus,

$$u(x+s\omega) - u(x) = \int_0^s \nabla u(x+t\omega)\cdot\omega\, dt$$

and therefore,

$$\begin{aligned}
\int_\Gamma |u(x+s\omega) - u(x)|d\sigma(\omega) &\le \int_0^s \int_\Gamma |\nabla u(x+t\omega)|d\sigma(\omega)dt \\
&= \int_0^s t^{d-1}dt \int_\Gamma \frac{|\nabla u(x+t\omega)|}{|x+t\omega-x|^{d-1}} d\sigma(\omega) \\
&= \int_{\Gamma_{x,s}} \frac{|\nabla u(y)|}{|y-x|^{d-1}}\, dy \le \int_{\Gamma_{x,r}} \frac{|\nabla u(y)|}{|x-y|^{d-1}}\, dy,
\end{aligned}$$

wherein the second equality we have used Eq. (49.1). Multiplying this inequality by s^{d-1} and integrating on $s \in [0,r]$ gives

$$\int_{\Gamma_{x,r}} |u(y) - u(x)|dy \le \frac{r^d}{d} \int_{\Gamma_{x,r}} \frac{|\nabla u(y)|}{|x-y|^{d-1}}\, dy = \frac{m(\Gamma_{x,r})}{|\Gamma|} \int_{\Gamma_{x,r}} \frac{|\nabla u(y)|}{|x-y|^{d-1}}\, dy$$

which proves Eq. (49.3). ■

Corollary 49.5. *Suppose $d < p \le \infty$, $\Gamma \in \mathcal{B}_{S^{d-1}}$ such that $|\Gamma| > 0$, $r \in (0,\infty)$ and $u \in C^1(\overline{\Gamma}_{x,r})$. Then*

$$|u(x)| \le C(|\Gamma|, r, d, p)\ \|u\|_{W^{1,p}(\Gamma_{x,r})} \tag{49.4}$$

where

$$C(|\Gamma|, r, d, p) := \frac{1}{|\Gamma|^{1/p}} \max\left(\frac{d^{-1/p}}{r}, \left(\frac{p-1}{p-d}\right)^{1-1/p}\right) \cdot r^{1-d/p}.$$

Proof. For $y \in \Gamma_{x,r}$,

$$|u(x)| \le |u(y)| + |u(y) - u(x)|$$

and hence using Eq. (49.3) and Hölder's inequality,

$$\begin{aligned}
|u(x)| &\le \fint_{\Gamma_{x,r}} |u(y)| dy + \frac{1}{|\Gamma|} \int_{\Gamma_{x,r}} \frac{|\nabla u(y)|}{|x-y|^{d-1}}\, dy \\
&\le \frac{1}{m(\Gamma_{x,r})} \|u\|_{L^p(\Gamma_{x,r})} \|1\|_{L^p(\Gamma_{x,r})} \\
&\quad + \frac{1}{|\Gamma|} \|\nabla u\|_{L^p(\Gamma_{x,r})} \|\frac{1}{|x-\cdot|^{d-1}}\|_{L^q(\Gamma_{x,r})}
\end{aligned} \tag{49.5}$$

where $q = \frac{p}{p-1}$ as before. Now

$$\begin{aligned}
\|\frac{1}{|\cdot|^{d-1}}\|^q_{L^q(\Gamma_{0,r})} &= \int_0^r dt\, t^{d-1} \int_\Gamma \left(t^{d-1}\right)^{-q}\ d\sigma(\omega) \\
&= |\Gamma| \int_0^r dt\ \left(t^{d-1}\right)^{1-\frac{p}{p-1}} = |\Gamma| \int_0^r dt\, t^{-\frac{d-1}{p-1}}
\end{aligned}$$

and since

$$1 - \frac{d-1}{p-1} = \frac{p-d}{p-1}$$

we find

$$\|\frac{1}{|\cdot|^{d-1}}\|_{L^q(\Gamma_{0,r})} = \left(\frac{p-1}{p-d}|\Gamma|\, r^{\frac{p-d}{p-1}}\right)^{1/q} = \left(\frac{p-1}{p-d}|\Gamma|\right)^{\frac{p-1}{p}} r^{1-\frac{d}{p}}. \tag{49.6}$$

Combining Eqs. (49.5), Eq. (49.6) along with the identity,

$$\frac{1}{m(\Gamma_{x,r})}\|1\|_{L^q(\Gamma_{x,r})} = \frac{1}{m(\Gamma_{x,r})} m(\Gamma_{x,r})^{1/q} = \left(|\Gamma|\, r^d/d\right)^{-1/p}, \tag{49.7}$$

shows

$$\begin{aligned}
|u(x)| &\le \|u\|_{L^p(\Gamma_{x,r})} \left(|\Gamma|\, r^d/d\right)^{-1/p} + \frac{1}{|\Gamma|}\ \|\nabla u\|_{L^p(\Gamma_{x,r})} \left(\frac{p-1}{p-d}|\Gamma|\right)^{1-1/p} r^{1-d/p} \\
&= \frac{1}{|\Gamma|^{1/p}} \left[\|u\|_{L^p(\Gamma_{x,r})} \frac{d^{-1/p}}{r} + \|\nabla u\|_{L^p(\Gamma_{x,r})} \left(\frac{p-1}{p-d}\right)^{1-1/p}\right] r^{1-d/p}. \\
&\le \frac{1}{|\Gamma|^{1/p}} \max\left(\frac{d^{-1/p}}{r}, \left(\frac{p-1}{p-d}\right)^{1-1/p}\right) \|u\|_{W^{1,p}(\Gamma_{x,r})} \cdot r^{1-d/p}.
\end{aligned}$$

■

Corollary 49.6. *For $d \in \mathbb{N}$ and $p \in (d, \infty]$ there are constants $\alpha = \alpha_d$ and $\beta = \beta_d$ such that if $u \in C^1(\mathbb{R}^d)$ then for all $x, y \in \mathbb{R}^d$,*

$$|u(y) - u(x)| \le 2\beta\alpha^{1/p} \left(\frac{p-1}{p-d}\right)^{\frac{p-1}{p}} \|\nabla u\|_{L^p(B(x,r)\cap B(y,r))} \cdot |x-y|^{\left(1-\frac{d}{p}\right)} \tag{49.8}$$

where $r := |x-y|$.

Proof. Let $r := |x-y|$, $V := B_x(r) \cap B_y(r)$ and $\Gamma, \Lambda \subset S^{d-1}$ be chosen so that $x + r\Gamma = \partial B_x(r) \cap B_y(r)$ and $y + r\Lambda = \partial B_y(r) \cap B_x(r)$, i.e.

$$\Gamma = \frac{1}{r}\left(\partial B_x(r) \cap B_y(r) - x\right) \text{ and } \Lambda = \frac{1}{r}\left(\partial B_y(r) \cap B_x(r) - y\right) = -\Gamma.$$

Also let $W = \Gamma_{x,r} \cap \Lambda_{y,r}$, see Figure 49.2 below. By a scaling,

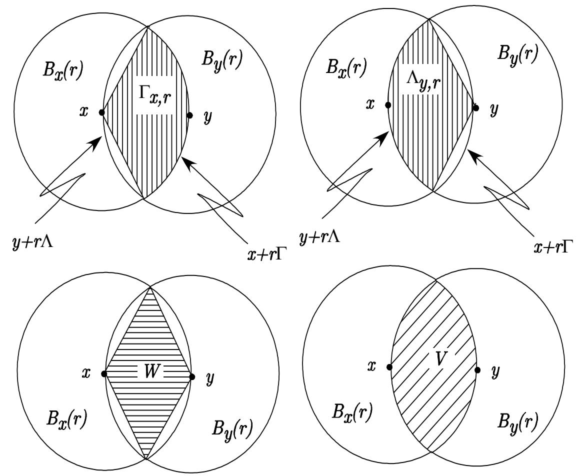


Fig. 49.2. The geometry of two intersecting balls of radius $r := |x-y|$. Here $W = \Gamma_{x,r} \cap \Lambda_{y,r}$ and $V = B(x,r) \cap B(y,r)$.

$$\beta_d := \frac{|\Gamma_{x,r} \cap \Lambda_{y,r}|}{|\Gamma_{x,r}|} = \frac{|\Gamma_{x,1} \cap \Lambda_{y,1}|}{|\Gamma_{x,1}|} \in (0,1)$$

is a constant only depending on d, i.e. we have $|\Gamma_{x,r}| = |\Lambda_{y,r}| = \beta|W|$. Integrating the inequality

$$|u(x) - u(y)| \le |u(x) - u(z)| + |u(z) - u(y)|$$

over $z \in W$ gives

$$|u(x) - u(y)| \le \fint_W |u(x) - u(z)|dz + \fint_W |u(z) - u(y)|dz$$
$$= \frac{\beta}{|\Gamma_{x,r}|}\left(\int_W |u(x) - u(z)|dz + \int_W |u(z) - u(y)|dz\right)$$
$$\le \frac{\beta}{|\Gamma_{x,r}|}\left(\int_{\Gamma_{x,r}} |u(x) - u(z)|dz + \int_{\Lambda_{y,r}} |u(z) - u(y)|dz\right).$$

Hence by Lemma 49.4, Hölder's inequality and translation and rotation invariance of Lebesgue measure,

$$|u(x) - u(y)| \le \frac{\beta}{|\Gamma|}\left(\int_{\Gamma_{x,r}} \frac{|\nabla u(z)|}{|x-z|^{d-1}}\,dz + \int_{\Lambda_{y,r}} \frac{|\nabla u(z)|}{|z-y|^{d-1}}\,dz\right)$$
$$\le \frac{\beta}{|\Gamma|}\left(\begin{array}{c}\|\nabla u\|_{L^p(\Gamma_{x,r})}\|\frac{1}{|x-\cdot|^{d-1}}\|_{L^q(\Gamma_{x,r})}\\ +\|\nabla u\|_{L^p(\Lambda_{y,r})}\|\frac{1}{|y-\cdot|^{d-1}}\|_{L^q(\Lambda_{y,r})}\end{array}\right)$$
$$\le \frac{2\beta}{|\Gamma|}\,\|\nabla u\|_{L^p(V)}\|\frac{1}{|\cdot|^{d-1}}\|_{L^q(\Gamma_{0,r})} \tag{49.9}$$

where $q = \frac{p}{p-1}$ is the conjugate exponent to p. Combining Eqs. (49.9) and (49.6) gives Eq. (49.8) with $\alpha := |\Gamma|^{-1}$. ■

Theorem 49.7 (Morrey's Inequality). *If $d < p \le \infty$, $u \in W^{1,p}(\mathbb{R}^d)$, then there exists a unique version u^* of u (i.e. $u^* = u$ a.e.) such that u^* is continuous. Moreover $u^* \in C^{0,1-\frac{d}{p}}(\mathbb{R}^d)$ and*

$$\|u^*\|_{C^{0,1-\frac{d}{p}}(\mathbb{R}^d)} \le C\|u\|_{W^{1,p}(\mathbb{R}^d)} \tag{49.10}$$

where $C = C(p,d)$ is a universal constant. Moreover, the estimates in Eqs. (49.3), (49.4) and (49.8) still hold when u is replaced by u^*.

Proof. For $p < \infty$ and $u \in C_c^1(\mathbb{R}^d)$, Corollaries 49.5 and 49.6 imply

$$\|u\|_{BC(\mathbb{R}^d)} \le C\|u\|_{W^{1,p}(\mathbb{R}^d)} \text{ and } \frac{|u(y) - u(x)|}{|x-y|^{1-\frac{d}{p}}} \le C\|\nabla u\|_{L^p(\mathbb{R}^d)}$$

which implies $[u]_{1-\frac{d}{p}} \le C\|\nabla u\|_{L^p(\mathbb{R}^d)} \le C\|u\|_{W^{1,p}(\mathbb{R}^d)}$ and hence

$$\|u\|_{C^{0,1-\frac{d}{p}}(\mathbb{R}^d)} \le C\|u\|_{W^{1,p}(\mathbb{R}^d)}. \tag{49.11}$$

Now suppose $u \in W^{1,p}(\mathbb{R}^d)$, choose (using Exercise 29.32) $u_n \in C_c^1(\mathbb{R}^d)$ such that $u_n \to u$ in $W^{1,p}(\mathbb{R}^d)$. Then by Eq. (49.11), $\|u_n - u_m\|_{C^{0,1-\frac{d}{p}}(\mathbb{R}^d)} \to 0$ as $m,n \to \infty$ and therefore there exists $u^* \in C^{0,1-\frac{d}{p}}(\mathbb{R}^d)$ such that $u_n \to u^*$ in $C^{0,1-\frac{d}{p}}(\mathbb{R}^d)$. Clearly $u^* = u$ a.e. and Eq. (49.10) holds.

If $p = \infty$ and $u \in W^{1,\infty}\left(\mathbb{R}^d\right)$, then by Proposition 29.29 there is a version u^* of u which is Lipschitz continuous. Now in both cases, $p < \infty$ and $p = \infty$, the sequence $u_m := u * \eta_m = u^* * \eta_m \in C^\infty\left(\mathbb{R}^d\right)$ and $u_m \to u^*$ uniformly on compact subsets of $\mathbb{R}^d$. Using Eq. (49.3) with u replaced by u_m along with a (by now) standard limiting argument shows that Eq. (49.3) still holds with u replaced by u^*. The proofs of Eqs. (49.4) and (49.8) only relied on Eq. (49.3) and hence go through without change. Similarly the argument in the first paragraph only relied on Eqs. (49.4) and (49.8) and hence Eq. (49.10) is also valid for $p = \infty$. ■

Corollary 49.8 (Morrey's Inequality). *Suppose $\Omega \subset_o \mathbb{R}^d$ such that $\overline{\Omega}$ is compact C^1-manifold with boundary and $d < p \le \infty$. Then for $u \in W^{1,p}(\Omega)$, there exists a unique version u^* of u such that $u^* \in C^{0,1-\frac{d}{p}}(\mathbb{R}^d)$ and we further have*

$$\|u^*\|_{C^{0,1-\frac{d}{p}}(\Omega)} \le C\|u\|_{W^{1,p}(\Omega)}, \tag{49.12}$$

where $C = C(p,d,\Omega)$.

Proof. Let U be a precompact open subset of $\mathbb{R}^d$ and $E : W^{1,p}(\Omega) \to W^{1,p}(\mathbb{R}^d)$ be an extension operator as in Theorem 48.35. For $u \in W^{1,p}(\Omega)$ with $d < p \le \infty$, Theorem 49.7 implies there is a version $U^* \in C^{0,1-\frac{d}{p}}(\mathbb{R}^d)$ of Eu. Letting $u^* := U^*|_\Omega$, we have and moreover,

$$\|u^*\|_{C^{0,1-\frac{d}{p}}(\Omega)} \le \|U^*\|_{C^{0,1-\frac{d}{p}}(\mathbb{R}^d)} \le C\|Eu\|_{W^{1,p}(\mathbb{R}^d)} \le C\,\|u\|_{W^{1,p}(\Omega)}.$$

■

The following example shows that $L^\infty(\mathbb{R}^d) \not\subseteq W^{1,d}(\mathbb{R}^d)$, i.e. $W^{1,d}(\mathbb{R}^d)$ contains unbounded elements. Therefore Theorem 49.7 and Corollary 49.8 are not valid for $p = d$. It turns out that for $p = d$, $W^{1,d}\left(\mathbb{R}^d\right)$ embeds into $BMO(\mathbb{R}^d)$ – the space of functions with "bounded mean oscillation."

Example 49.9. Let $u(x) = \psi(x)\log\log\left(1 + \frac{1}{|x|}\right)$ where $\psi \in C_c^\infty(\mathbb{R}^d)$ is chosen so that $\psi(x) = 1$ for $|x| \le 1$. Then $u \notin L^\infty(\mathbb{R}^d)$ while $u \in W^{1,d}(\mathbb{R}^d)$. Let us check this claim. Using Theorem 9.35, one easily shows $u \in L^p(\mathbb{R}^d)$. A short computation shows, for $|x| < 1$, that

$$\nabla u(x) = \frac{1}{\log\left(1 + \frac{1}{|x|}\right)}\,\frac{1}{1 + \frac{1}{|x|}}\nabla\frac{1}{|x|}$$
$$= \frac{1}{1 + \frac{1}{|x|}}\,\frac{1}{\log\left(1 + \frac{1}{|x|}\right)}\,\left(-\frac{1}{|x|}\,\hat{x}\right)$$

where $\hat{x} = x/|x|$ and so again by Theorem 9.35,

$$\int_{\mathbb{R}^d} |\nabla u(x)|^d dx \geq \int_{|x|<1} \left(\frac{1}{|x|^2 + |x|} \frac{1}{\log\left(1 + \frac{1}{|x|}\right)} \right)^d dx$$
$$\geq \sigma(S^{d-1}) \int_0^1 \left(\frac{2}{r \log\left(1 + \frac{1}{r}\right)} \right)^d r^{d-1} dr = \infty.$$

49.2 Rademacher's Theorem

Theorem 49.10. *Suppose that $u \in W^{1,p}_{loc}(\Omega)$ for some $d < p \leq \infty$. Then u is differentiable almost everywhere and w-$\partial_i u = \partial_i u$ a.e. on Ω.*

Proof. We clearly may assume that $p < \infty$. For $v \in W^{1,p}_{loc}(\Omega)$ and $x, y \in \Omega$ such that $\overline{B(x,r) \cap B(y,r)} \subset \Omega$ where $r := |x - y|$, the estimate in Corollary 49.6, gives

$$|v(y) - v(x)| \leq C\|\nabla u\|_{L^p(B(x,r)\cap B(y,r))} \cdot |x-y|^{\left(1-\frac{d}{p}\right)}$$
$$= C\|\nabla v\|_{L^p(B(x,r)\cap B(y,r))} \cdot r^{\left(1-\frac{d}{p}\right)}. \tag{49.13}$$

Let u now denote the unique continuous version of $u \in W^{1,p}_{loc}(\Omega)$. The by the Lebesgue differentiation Theorem 20.12, there exists an exceptional set $E \subset \Omega$ such that $m(E) = 0$ and

$$\lim_{r\downarrow 0} \fint_{B(x,r)} |\nabla u(y) - \nabla u(x)|^p dy = 0 \text{ for } x \in \Omega \setminus E.$$

Fix a point $x \in \Omega \setminus E$ and let $v(y) := u(y) - u(x) - \nabla u(x) \cdot (y - x)$ and notice that $\nabla v(y) = \nabla u(y) - \nabla u(x)$. Applying Eq. (49.13) to v then implies

$$|u(y) - u(x) - \nabla u(x) \cdot (y - x)|$$
$$\leq C\|\nabla u(\cdot) - \nabla u(x)\|_{L^p(B(x,r)\cap B(y,r))} \cdot r^{\left(1-\frac{d}{p}\right)}$$
$$\leq C\left(\int_{B(x,r)} |\nabla u(y) - \nabla u(x)|^p dy\right)^{1/p} \cdot r^{\left(1-\frac{d}{p}\right)}$$
$$= C\sigma\left(S^{d-1}\right)^{1/p} r^{d/p} \left(\fint_{B(x,r)} |\nabla u(y) - \nabla u(x)|^p dy\right)^{1/p} \cdot r^{\left(1-\frac{d}{p}\right)}$$
$$= C\sigma\left(S^{d-1}\right)^{1/p} \left(\fint_{B(x,r)} |\nabla u(y) - \nabla u(x)|^p dy\right)^{1/p} \cdot |x-y|$$

which shows u is differentiable at x and $\nabla u(x) = w$-$\nabla u(x)$. ■

Theorem 49.11 (Rademacher's Theorem). *Let u be locally Lipschitz continuous on $\Omega \subset_o \mathbb{R}^d$. Then u is differentiable almost everywhere and w-$\partial_i u = \partial_i u$ a.e. on Ω.*

Proof. By Proposition 29.29 $\partial_i^{(w)} u$ exists weakly and is in $\partial_i u \in L^\infty(\mathbb{R}^d)$ for $i = 1, 2, \ldots, d$. The result now follows from Theorem 49.10. ■

49.3 Gagliardo-Nirenberg-Sobolev Inequality

In this section our goal is to prove an inequality of the form:

$$\|u\|_{L^q} \leq C\|\nabla u\|_{L^p(\mathbb{R}^d)} \text{ for } u \in C_c^1(\mathbb{R}^d). \tag{49.14}$$

For $\lambda > 0$, let $u_\lambda(x) = u(\lambda x)$. Then

$$\|u_\lambda\|_{L^q}^q = \int_{\mathbb{R}^d} |u(\lambda x)|^q dx = \int_{\mathbb{R}^d} |u(y)|^q \frac{dy}{\lambda^d}$$

and hence $\|u_\lambda\|_{L^q} = \lambda^{-d/q}\|u\|_{L^q}$. Moreover, $\nabla u_\lambda(x) = \lambda(\nabla u)(\lambda x)$ and thus

$$\|\nabla u_\lambda\|_{L^p} = \lambda\|(\nabla u)_\lambda\|_{L^p} = \lambda\,\lambda^{-d/p}\|\nabla u\|_{L^p}.$$

If (49.14) is to hold for all $u \in C_c^1(\mathbb{R}^d)$ then we must have

$$\lambda^{-d/q}\|u\|_{L^q} = \|u_\lambda\|_{L^q} \leq C\|\nabla u_\lambda\|_{L^p(\mathbb{R}^d)} = C\lambda^{1-d/p}\|\nabla u\|_{L^p} \text{ for all } \lambda > 0$$

which is only possible if

$$1 - d/p + d/q = 0, \text{ i.e. } 1/p = 1/d + 1/q. \tag{49.15}$$

Notation 49.12 *For $p \in [1, d]$, let $p^* := \frac{dp}{d-p}$ with the convention that $p^* = \infty$ if $p = d$. That is $p^* = q$ where q solves Eq. (49.15).*

Theorem 49.13. *Let $p = 1$ so $1^* = \frac{d}{d-1}$, then*

$$\|u\|_{1^*} = \|u\|_{\frac{d}{d-1}} \leq \prod_{i=1}^d \left(\int_{\mathbb{R}^d} |\partial_i u(x)| dx\right)^{\frac{1}{d}} \leq d^{-\frac{1}{2}} \|\nabla u\|_1 \tag{49.16}$$

for all $u \in W^{1,1}(\mathbb{R}^d)$.

Proof. Since there exists $u_n \in C_c^1(\mathbb{R}^d)$ such that $u_n \to u$ in $W^{1,1}(\mathbb{R}^d)$, a simple limiting argument shows that it suffices to prove Eq. (49.16) for $u \in C_c^1(\mathbb{R}^d)$. To help the reader understand the proof, let us give the proof for $d \leq 3$ first and with the constant $d^{-1/2}$ being replaced by 1. After that the general induction argument will be given. (The adventurous reader may skip directly to the paragraph containing Eq. (49.17).

($d=1$, $p^*=\infty$) By the fundamental theorem of calculus,

$$|u(x)| = \left|\int_{-\infty}^{x} u'(y)dy\right| \le \int_{-\infty}^{x} |u'(y)|\, dy \le \int_{\mathbb{R}} |u'(x)|\, dx.$$

Therefore $\|u\|_{L^\infty} \le \|u'\|_{L^1}$[1], proving the $d=1$ case.

($d=2$, $p^*=2$) Applying the same argument as above to $y_1 \to u(y_1, x_2)$ and $y_2 \to u(x_1, y_2)$,,

$$|u(x_1,x_2)| \le \int_{-\infty}^{\infty} |\partial_1 u(y_1,x_2)|\, dy_1 \le \int_{-\infty}^{\infty} |\nabla u(y_1,x_2)|\, dy_1 \text{ and}$$
$$|u(x_1,x_2)| \le \int_{-\infty}^{\infty} |\partial_2 u(x_1,y_2)|\, dy_2 \le \int_{-\infty}^{\infty} |\nabla u(x_1,y_2)|\, dy_2$$

and therefore

$$|u(x_1,x_2)|^2 \le \int_{-\infty}^{\infty} |\partial_1 u(y_1,x_2)| dy_1 \cdot \int_{-\infty}^{\infty} |\partial_2 u(x_1,y_2)|\, dy_2.$$

Integrating this equation relative to x_1 and x_2 gives

$$\|u\|_{L^2}^2 = \int_{\mathbb{R}^2} |u(x)|^2 dx \le \left(\int_{-\infty}^{\infty} |\partial_1 u(x)|\ dx\right)\left(\int_{-\infty}^{\infty} |\partial_2 u(x)|\ dx\right)$$
$$\le \left(\int_{-\infty}^{\infty} |\nabla u(x)|\ dx\right)^2$$

which proves the $d=2$ case.

($d=3$, $p^*=3/2$) Let $x^1=(y_1,x_2,x_3)$, $x^2=(x_1,y_2,x_3)$, and $x^3=(x_1,x_2,y_3)$. Then as above,

$$|u(x)| \le \int_{-\infty}^{\infty} |\partial_i u(x^i)| dy_i \text{ for } i=1,2,3$$

[1] Actually we may do better here by observing

$$|u(x)| = \frac{1}{2}\left|\int_{-\infty}^{x} u'(y)dy - \int_{x}^{\infty} u'(y)dy\right|$$
$$\le \frac{1}{2}\int_{\mathbb{R}} |u'(x)|\, dx$$

and this leads to an improvement in Eq. (49.17) to

$$\|u\|_{1^*} \le \frac{1}{2}\prod_{i=1}^{d}\left(\int_{\mathbb{R}^d} |\partial_i u(x)| dx\right)^{\frac{1}{d}}$$
$$\le \frac{1}{2} d^{-\frac{1}{2}} \|\nabla u\|_1 .$$

and hence

$$|u(x)|^{\frac{3}{2}} \le \prod_{i=1}^{3}\left(\int_{-\infty}^{\infty} |\partial_i u(x^i)| dy_i\right)^{\frac{1}{2}}.$$

Integrating this equation on x_1 gives,

$$\int_{\mathbb{R}} |u(x)|^{\frac{3}{2}} dx_1 \le \left(\int_{-\infty}^{\infty} |\partial_1 u(x^1)| dy_1\right)^{\frac{1}{2}} \int \prod_{i=2}^{3}\left(\int_{-\infty}^{\infty} |\partial_i u(x^i)| dy_i\right)^{\frac{1}{2}} dx_1$$
$$\le \left(\int_{-\infty}^{\infty} |\partial_1 u(x)| dx_1\right)^{\frac{1}{2}} \prod_{i=2}^{3}\left(\int_{-\infty}^{\infty} |\partial_i u(x^i)| dx_1 dy_i\right)^{\frac{1}{2}}$$

wherein the second equality we have used the Hölder's inequality with $p=q=2$. Integrating this result on x_2 and using Hölder's inequality gives

$$\int_{\mathbb{R}^2} |u(x)|^{\frac{3}{2}} dx_1 dx_2$$
$$\le \left(\int_{\mathbb{R}^2} |\partial_2 u(x)| dx_1 dx_2\right)^{\frac{1}{2}} \int_{\mathbb{R}} dx_2 \left(\int_{-\infty}^{\infty} |\partial_1 u(x)| dx_1\right)^{\frac{1}{2}} \times$$
$$\left(\int_{\mathbb{R}^2} |\partial_3 u(x^3)| dx_1 dy_3\right)^{\frac{1}{2}}$$
$$\le \left(\int_{\mathbb{R}^2} |\partial_2 u(x)| dx_1 dx_2\right)^{\frac{1}{2}} \left(\int_{\mathbb{R}^2} |\partial_1 u(x)| dx_1 dx_2\right)^{\frac{1}{2}} \left(\int_{\mathbb{R}^3} |\partial_3 u(x)| dx\right)^{\frac{1}{2}}.$$

One more integration of x_3 and application of Hölder's inequality, implies

$$\int_{\mathbb{R}^3} |u(x)|^{\frac{3}{2}} dx \le \prod_{i=1}^{3}\left(\int_{\mathbb{R}^3} |\partial_i u(x)| dx\right)^{\frac{1}{2}} \le \left(\int_{\mathbb{R}^3} |\nabla u(x)| dx\right)^{\frac{3}{2}}$$

proving the $d=3$ case.

For general d ($p^* = \frac{d}{d-1}$), as above let $x^i = (x_1, \dots, x_{i-1}, y_i, x_{i+1} \dots, x_d)$. Then

$$|u(x)| \le \int_{-\infty}^{\infty} |\partial_i u(x^i)| dy_i$$

and

$$|u(x)|^{\frac{d}{d-1}} \le \prod_{i=1}^{d}\left(\int_{-\infty}^{\infty} |\partial_i u(x^i)| dy_i\right)^{\frac{1}{d-1}}. \tag{49.17}$$

Integrating this equation relative to x_1 and making use of Hölder's inequality in the form

$$\left\|\prod_{i=2}^{d} f_i\right\|_1 \le \prod_{i=2}^{d} \|f_i\|_{d-1} \tag{49.18}$$

(see Corollary 10.3) we find

$$\int_{\mathbb{R}} |u(x)|^{\frac{d}{d-1}} dx_1 \le \left(\int_{\mathbb{R}} \partial_1 u(x) dx_1\right)^{\frac{1}{d-1}} \int_{\mathbb{R}} dx_1 \prod_{i=2}^{d} \left(\int_{\mathbb{R}} |\partial_i u(x^i)| dy_i\right)^{\frac{1}{d-1}}$$
$$\le \left(\int_{\mathbb{R}} \partial_1 u(x) dx_1\right)^{\frac{1}{d-1}} \prod_{i=2}^{d} \left(\int_{\mathbb{R}^2} |\partial_i u(x^i)| dx_1 dy_i\right)^{\frac{1}{d-1}}$$
$$= \left(\int_{\mathbb{R}} \partial_1 u(x) dx_1\right)^{\frac{1}{d-1}} \left(\int_{\mathbb{R}^2} |\partial_2 u(x)| dx_1 dx_2\right)^{\frac{1}{d-1}} \times \prod_{i=3}^{d} \left(\int_{\mathbb{R}^2} |\partial_i u(x^i)| dx_1 dy_i\right)^{\frac{1}{d-1}}.$$

Integrating this equation on x_2 and using Eq. (49.18) once again implies,

$$\int_{\mathbb{R}^2} |u(x)|^{\frac{d}{d-1}} dx_1 dx_2$$
$$\le \left(\int_{\mathbb{R}^2} |\partial_2 u(x)| dx_1 dx_2\right)^{\frac{1}{d-1}} \int_{\mathbb{R}} dx_2 \left(\int_{\mathbb{R}} \partial_1 u(x) dx_1\right)^{\frac{1}{d-1}} \times \prod_{i=3}^{d} \left(\int_{\mathbb{R}^2} |\partial_i u(x^i)| dx_1 dy_i\right)^{\frac{1}{d-1}}$$
$$\le \left(\int_{\mathbb{R}^2} |\partial_2 u(x)| dx_1 dx_2\right)^{\frac{1}{d-1}} \left(\int_{\mathbb{R}^2} |\partial_1 u(x)| dx_1 dx_2\right)^{\frac{1}{d-1}} \times \prod_{i=3}^{d} \left(\int_{\mathbb{R}^3} |\partial_i u(x^i)| dx_1 dx_2 dy_i\right)^{\frac{1}{d-1}}.$$

Continuing this way inductively, one shows

$$\int_{\mathbb{R}^k} |u(x)|^{\frac{d}{d-1}} dx_1 dx_2 \dots dx_k$$
$$\le \prod_{i=1}^{k} \left(\int_{\mathbb{R}^k} |\partial_i u(x)| dx_1 dx_2 \dots dx_k\right)^{\frac{1}{d-1}} \times \prod_{i=k+1}^{d} \left(\int_{\mathbb{R}^3} |\partial_i u(x^i)| dx_1 dx_2 \dots dx_k dy_{k+1}\right)^{\frac{1}{d-1}}$$

and in particular when $k = d$,

$$\int_{\mathbb{R}^d} |u(x)|^{\frac{d}{d-1}} dx \le \left(\frac{1}{2}\right)^{\frac{d}{d-1}} \prod_{i=1}^{d} \left(\int_{\mathbb{R}^d} |\partial_i u(x)| dx_1 dx_2 \dots dx_d\right)^{\frac{1}{d-1}} \tag{49.19}$$
$$\le \prod_{i=1}^{d} \left(\int_{\mathbb{R}^d} |\nabla u(x)| dx\right)^{\frac{1}{d-1}} = \left(\int_{\mathbb{R}^d} |\nabla u(x)| dx\right)^{\frac{d}{d-1}}.$$

This estimate may now be improved on by using Young's inequality (see Exercise 49.33) in the form $\prod_{i=1}^{d} a_i \le \frac{1}{d} \sum_{i=1}^{d} a_i^d$. Indeed by Eq. (49.19) and Young's inequality,

$$\|u\|_{\frac{d}{d-1}} \le \prod_{i=1}^{d} \left(\int_{\mathbb{R}^d} |\partial_i u(x)| dx\right)^{\frac{1}{d}} \le \frac{1}{d} \sum_{i=1}^{d} \left(\int_{\mathbb{R}^d} |\partial_i u(x)| dx\right)$$
$$= \frac{1}{d} \int_{\mathbb{R}^d} \sum_{i=1}^{d} |\partial_i u(x)| dx \le \frac{1}{d} \int_{\mathbb{R}^d} \sqrt{d}\, |\nabla u(x)|\, dx$$

wherein the last inequality we have used Hölder's inequality for sums,

$$\sum_{i=1}^{d} |a_i| \le \left(\sum_{i=1}^{d} 1\right)^{1/2} \left(\sum_{i=1}^{d} |a_i|^2\right)^{1/2} = \sqrt{d}\, |a|.$$

■

The next theorem generalizes Theorem 49.13 to an inequality of the form in Eq. (49.14).

Theorem 49.14. *If $p \in [1, d)$ then,*

$$\|u\|_{L^{p^*}} \le d^{-1/2} \frac{p(d-1)}{d-p} \|\nabla u\|_{L^p} \text{ for all } u \in W^{1,p}(\mathbb{R}^d). \tag{49.20}$$

Proof. As usual since $C_c^1(\mathbb{R}^d)$ is dense in $W^{1,p}(\mathbb{R}^d)$ it suffices to prove Eq. (49.20) for $u \in C_c^1(\mathbb{R}^d)$. For $u \in C_c^1(\mathbb{R}^d)$ and $s > 1$, $|u|^s \in C_c^1(\mathbb{R}^d)$ and $\nabla |u|^s = s|u|^{s-1}\mathrm{sgn}(u)\nabla u$. Applying Eq. (49.16) with u replaced by $|u|^s$ and then using Holder's inequality gives

$$\||u|^s\|_{1^*} \le d^{-\frac{1}{2}} \|\nabla |u|^s\|_1 = s d^{-\frac{1}{2}} \||u|^{s-1} \nabla u\|_{L^1}$$
$$\le \frac{s}{\sqrt{d}} \|\nabla u\|_{L^p} \cdot \||u|^{s-1}\|_{L^q} \tag{49.21}$$

where $q = \frac{p}{p-1}$. We will now choose s so that $s1^* = (s-1)q$, i.e.

$$s = \frac{q}{q - 1^*} = \frac{1}{1 - 1^* \frac{1}{q}} = \frac{1}{1 - \frac{d}{d-1}\left(1 - \frac{1}{p}\right)}$$
$$= \frac{p(d-1)}{p(d-1) - d(p-1)} = \frac{p(d-1)}{d-p} = p^* \frac{d-1}{d}.$$

For this choice of s, $s1^* = p^* = (s-1)q$ and Eq. (49.21) becomes

$$\left[\int_{\mathbb{R}^d} |u|^{p^*} dm\right]^{1/1^*} \le \frac{s}{\sqrt{d}} \|\nabla u\|_{L^p} \cdot \left[\int_{\mathbb{R}^d} |u|^{p^*} dm\right]^{1/q}. \tag{49.22}$$

Since

$$\frac{1}{1^*}-\frac{1}{q}=\frac{d-1}{d}-\frac{p-1}{p}=\frac{p(d-1)-d(p-1)}{dp}$$
$$=\frac{d-p}{pd}=\frac{1}{p^*},$$

Eq. (49.22) implies Eq. (49.20). ∎

Corollary 49.15. *Suppose $\Omega\subset\mathbb{R}^d$ is bounded open set with C^1-boundary, then for all $p\in[1,d)$ and $1\le q\le p^*$ there exists $C=C(\Omega,p,q)$ such that*

$$\|u\|_{L^q(\Omega)}\le C\|u\|_{W^{1,p}(\Omega)}.$$

Proof. Let U be a precompact open subset of $\mathbb{R}^d$ such that $\bar{\Omega}\subset U$ and $E:W^{1,p}(\Omega)\to W^{1,p}(\mathbb{R}^d)$ be an extension operator as in Theorem 48.35. Then for $u\in C^1(\overline{\Omega})\cap W^{1,p}(\Omega)$,

$$\|u\|_{L^{p^*}(\Omega)}\le C\|Eu\|_{L^{p^*}(\mathbb{R}^d)}\le C\|\nabla(Eu)\|_{L^p(\mathbb{R}^d)}\le C\|u\|_{W^{1,p}(\Omega)},$$

i.e.

$$\|u\|_{L^{p^*}(\Omega)}\le C\|u\|_{W^{1,p}(\Omega)} \tag{49.23}$$

Since $C^1(\overline{\Omega})$ is dense in $W^{1,p}(\Omega)$, Eq. (49.23) holds for all $u\in W^{1,p}(\Omega)$. Finally for all $1\le q<p^*$,

$$\|u\|_{L^q(\Omega)}\le\|u\|_{L^{P^*}(\Omega)}\cdot\|1\|_{L^r(\Omega)}=\|u\|_{L^{p^*}}(\lambda(\Omega))^{\frac{1}{r}}$$
$$\le C(\lambda(\Omega))^{\frac{1}{r}}\|u\|_{W^{1,p}(\Omega)}$$

where $\frac{1}{r}+\frac{1}{p^*}=\frac{1}{q}$. ∎

49.4 Sobolev Embedding Theorems Summary

Let us summarize what we have proved up to this point in the following theorem.

Theorem 49.16. *Let $p\in[1,\infty]$ and $u\in W^{1,p}(\mathbb{R}^d)$. Then*

1. ***Morrey's Inequality.*** *If $p>d$, then $W^{1,p}\hookrightarrow C^{0,1-\frac{d}{p}}$ and*

$$\|u^*\|_{C^{0,1-\frac{d}{p}}(\mathbb{R}^d)}\le C\|u\|_{W^{1,p}(\mathbb{R}^d)}.$$

2. *When $p=d$ there is an L^∞ – like space called BMO (which is **not** defined in these notes) such that $W^{1,p}\hookrightarrow BMO$.*

3. ***GNS Inequality.*** *If $1\le p<d$, then $W^{1,p}\hookrightarrow L^{p^*}$*

$$\|u\|_{L^{p^*}}\le d^{-1/2}\frac{p(d-1)}{d-p}\|\nabla u\|_{L^p}$$

where $p^*=\frac{dp}{d-p}$ or equivalently $\frac{1}{p^*}=\frac{1}{p}-\frac{1}{d}$.

Our next goal is write out the embedding theorems for $W^{k,p}(\Omega)$ for general k and p.

Notation 49.17 *Given a number $s\ge0$, let*

$$s_+=\begin{cases} s & \text{if } n\notin\mathbb{N}_0\\ s+\delta & \text{if } n\in\mathbb{N}_0\end{cases}$$

*where $\delta>0$ is some arbitrarily small number. When $s=k+\alpha$ with $k\in\mathbb{N}_0$ and $0\le\alpha<1$ we will write $C^{k,\alpha}(\Omega)$ simply as $C^s(\Omega)$. **Warning,** although $C^{k,1}(\Omega)\subset C^{k+1}(\Omega)$ it is **not true** that $C^{k,1}(\Omega)=C^{k+1}(\Omega)$.*

Theorem 49.18 (Sobolev Embedding Theorems). *Suppose $\Omega=\mathbb{R}^d$ or $\Omega\subset\mathbb{R}^d$ is bounded open set with C^1-boundary, $p\in[1,\infty)$, $k,l\in\mathbb{N}$ with $l\le k$.*

1. *If $p<d/l$ then $W^{k,p}(\Omega)\hookrightarrow W^{k-l,q}(\Omega)$ provided $q:=\frac{dp}{d-pl}$, i.e. q solves*

$$\frac{1}{q}=\frac{1}{p}-\frac{l}{d}>0$$

and there is a constant $C<\infty$ such that

$$\|u\|_{W^{k-l,q}(\Omega)}\le C\|u\|_{W^{k,p}(\Omega)} \text{ for all } u\in W^{k,p}(\Omega).$$

2. *If $p>d/k$, then $W^{k,p}(\Omega)\hookrightarrow C^{k-(d/p)_+}(\Omega)$ and there is a constant $C<\infty$ such that*

$$\|u\|_{C^{k-(d/p)_+}(\Omega)}\le C\|u\|_{W^{k,p}(\Omega)} \text{ for all } u\in W^{k,p}(\Omega).$$

Proof. 1. $(p<d/l)$ If $u\in W^{k,p}(\Omega)$, then $\partial^\alpha u\in W^{1,p}(\Omega)$ for all $|\alpha|\le k-1$. Hence by Corollary 49.15, $\partial^\alpha u\in L^{p^*}(\Omega)$ for all $|\alpha|\le k-1$ and therefore $W^{k,p}(\Omega)\hookrightarrow W^{k-1,p^*}(\Omega)$ and there exists a constant C_1 such that

$$\|u\|_{W^{k-1,p_1}(\Omega)}\le C\|u\|_{W^{k,p}(\Omega)} \text{ for all } u\in W^{k,p}(\Omega). \tag{49.24}$$

Define p_j inductively by, $p_1:=p^*$ and $p_j:=p^*_{j-1}$. Since $\frac{1}{p_j}=\frac{1}{p_{j-1}}-\frac{1}{d}$ it is easily checked that $\frac{1}{p_l}=\frac{1}{p}-\frac{l}{d}>0$ since $p<d/l$. Hence using Eq. (49.24) repeatedly we learn that the following inclusion maps are all bounded:

$$W^{k,p}(\Omega)\hookrightarrow W^{k-1,p_1}(\Omega)\hookrightarrow W^{k-2,p_2}(\Omega)\ldots\hookrightarrow W^{k-l,p_l}(\Omega).$$

This proves the first item of the theorem. The following lemmas will be used in the proof of item 2. ∎

Lemma 49.19. *Suppose $j \in \mathbb{N}$ and $p \geq d$ and $j > d/p$ (i.e. $j \geq 1$ if $p > d$ and $j \geq 2$ if $p = d$) then*

$$W^{j,p}(\Omega) \hookrightarrow C^{j-(d/p)_+}(\Omega)$$

and there is a constant $C < \infty$ such that

$$\|u\|_{C^{j-(d/p)_+}(\Omega)} \leq C \|u\|_{W^{j,p}(\Omega)}. \tag{49.25}$$

Proof. By the usual methods, it suffices to show that the estimate in Eq. (49.25) holds for all $u \in C^\infty(\bar{\Omega})$.

For $p > d$ and $|\alpha| \leq j-1$,

$$\|\partial^\alpha u\|_{C^{0,1-d/p}(\Omega)} \leq C \|\partial^\alpha u\|_{W^{1,p}(\Omega)} \leq C \|u\|_{W^{j,p}(\Omega)}$$

and hence

$$\|u\|_{C^{j-d/p}(\Omega)} := \|u\|_{C^{j-1,1-d/p}(\Omega)} \leq C \|u\|_{W^{j,p}(\Omega)}$$

which is Eq. (49.25).

When $p = d$ (so now $j \geq 2$), choose $q \in (1,d)$ be close to d so that $j > d/q$ and $q^* = \frac{qd}{d-q} > d$. Then

$$W^{j,d}(\Omega) \hookrightarrow W^{j,q}(\Omega) \hookrightarrow W^{j-1,q^*}(\Omega) \hookrightarrow C^{j-2,1-d/q^*}(\Omega).$$

Since $d/q^* \downarrow 0$ as $q \uparrow d$, we conclude that $W^{j,d}(\Omega) \hookrightarrow C^{j-2,\alpha}(\Omega)$ for any $\alpha \in (0,1)$ which we summarizes by writing

$$W^{j,d}(\Omega) \hookrightarrow C^{j-(d/d)}(\Omega).$$

■

Proof. Continuation of the proof of Theorem 49.18. Item 2., $(p > d/k)$. If $p \geq d$, the result follows from Lemma 49.19. So nos suppose that $d > p > d/k$ and choose the largest l such that $1 \leq l < k$ and $d/l > p$ and let $q = \frac{dp}{d-pl}$, i.e. q solves $q \geq d$ and

$$\frac{1}{q} = \frac{1}{p} - \frac{l}{d} \text{ or } \frac{d}{q} = \frac{d}{p} - l$$

Then

$$W^{k,p}(\Omega) \hookrightarrow W^{k-l,q}(\Omega) \hookrightarrow C^{k-l-(d/q)_+}(\Omega) = C^{k-l-\left(\frac{d}{p}-l\right)_+}(\Omega) = C^{k-\left(\frac{d}{p}\right)_+}(\Omega)$$

as desired. ■

Remark 49.20 (Rule of thumb.). Assign the "degrees of regularity" $k-(d/p)_+$ to the space $W^{k,p}$ and $k+\alpha$ to the space $C^{k,\alpha}$. If

$$X, Y \in \left\{W^{k,p} : k \in \mathbb{N}_0, p \in [1,\infty]\right\} \cup \left\{C^{k,\alpha} : k \in \mathbb{N}_0,\ \alpha \in [0,1]\right\}$$

with $\deg_{\text{reg}}(X) \geq \deg_{\text{reg}}(Y)$, then $X \hookrightarrow Y$.

Example 49.21. 1. $W^{k,p} \hookrightarrow W^{k-\ell,q}$ iff $k - \frac{d}{p} \geq k - \ell - \frac{d}{q}$ iff $\ell \geq \frac{d}{p} - \frac{d}{q}$ iff $\frac{1}{q} \geq \frac{1}{p} - \frac{\ell}{d}$.
2. $W^{k,p} \subset C^{0,\alpha}$ iff $k - \left(\frac{d}{p}\right)_+ \geq \alpha$.

49.5 Compactness Theorems

Lemma 49.22. *Suppose $K_m : X \to Y$ are compact operators and $\|K - K_m\|_{L(X,Y)} \to 0$ as $n \to \infty$ then K is compact.*

Proof. Let $\{x_n\}_{n=1}^\infty \subset X$ be given such that $\|x_n\| \leq 1$. By Cantor's diagonalization scheme we may choose $\{x_n'\} \subset \{x_n\}$ such that $y_m := \lim_{n\to\infty} K_m x_n' \in Y$ exists for all m. Hence

$$\begin{aligned}\|Kx_n' - Kx_\ell'\| &= \|K(x_n' - x_\ell')\| \leq \|K(x_n' - x_\ell')\| \\ &\leq \|K - K_m\| \|x_n' - x_\ell'\| + \|K_m(x_n' - x_\ell')\| \\ &\leq \|K - K_m\| + \|K_m(x_n' - x_\ell')\|\end{aligned}$$

and therefore,

$$\limsup_{l,n\to\infty} \|Kx_n' - Kx_\ell'\| \leq \|K - K_m\| \to 0 \text{ as } m \to \infty.$$

■

Lemma 49.23. *Let $\eta \in C_c^\infty(\mathbb{R}^d)$, $C_\eta f = \eta * f$, $\Omega \subset \mathbb{R}^d$ be a bounded open set with C^1-boundary, U be an open precompact subset of $\mathbb{R}^d$ such that $\bar{\Omega} \subset U$ and $E : W^{1,1}(\Omega) \to W^{1,1}(\mathbb{R}^d)$ be an extension operator as in Theorem 48.35. Then to every bounded sequence $\{\tilde{u}_n\}_{n=1}^\infty \subset W^{1,1}(\Omega)$ there has a subsequence $\{u_n'\}_{n=1}^\infty$ such that $C_\eta E u_n'$ is uniformly convergent to a function in $C_c(\mathbb{R}^d)$.*

Proof. Let $u_n := E\tilde{u}_n$ and $C := \sup \|u_n\|_{W^{1,1}(\mathbb{R}^d)}$ which is finite by assumption. So $\{u_n\}_{n=1}^\infty \subset W^{1,1}(\mathbb{R}^d)$ is a bounded sequence such that $\text{supp}(u_n) \subset U \subset \bar{U} \sqsubset\sqsubset \mathbb{R}^d$ for all n. Since η is compactly supported there exists a precompact open set V such that $\bar{U} \subset V$ and $v_n := \eta * u_n \in C_c^\infty(V) \subset C_c^\infty(\mathbb{R}^d)$ for all n. Since,

$$\begin{aligned}\|v_n\|_{L^\infty} &\leq \|\eta\|_{L^\infty} \|u_n\|_{L^1} \leq \|\eta\|_{L^\infty} \|u_n\|_{L^1} \leq C\|\eta\|_{L^\infty} \text{and} \\ \|Dv_n\|_{L^\infty} &= \|\eta * Du_n\|_{L^\infty} \leq \|\eta\|_{L^\infty} \|Du_n\|_{L^1} \leq C\|\eta\|_{L^\infty},\end{aligned}$$

it follows by the Arzela-Ascoli theorem that $\{v_n\}_{n=1}^\infty$ has a uniformly convergent subsequence. ■

Lemma 49.24. *Let $\eta \in C_c^\infty(B(0,1),[0,\infty))$ such that $\int_{\mathbb{R}^d} \eta dm = 1$, $\eta_m(x) = m^n \eta(mx)$ and $K_m u = (C_{\eta_m} Eu)|_\Omega$. Then for all $p \in [1,d)$ and $q \in [1,p^*)$,*

$$\lim_{m\to\infty} \|K_m - i\|_{B(W^{1,p}(\Omega),L^q(\Omega))} = 0$$

where $i : W^{1,p}(\Omega) \to L^q(\Omega)$ is the inclusion map.

Proof. For $u \in C_c^1(U)$ let $v_m := \eta_m * u - u$, then

$$\begin{aligned}|v_m(x)| \le |\eta_m * u(x) - u(x)| &= \left|\int_{\mathbb{R}^d} \eta_m(y)(u(x-y) - u(x))dy\right| \\ &= \left|\int_{\mathbb{R}^d} \eta(y)\left[u(x - \frac{y}{m}) - u(x)\right] dy\right| \\ &\le \frac{1}{m}\int_{\mathbb{R}^d} dy\ |y|\,\eta(y)\int_0^1 dt\ \left|\nabla u(x - t\frac{y}{m})\right|\end{aligned}$$

and so by Minikowski's inequality for integrals,

$$\begin{aligned}\|v_m\|_{L^r} &\le \frac{1}{m}\int_{\mathbb{R}^d} dy\ |y|\,\eta(y)\int_0^1 dt\ \left\|\nabla u(\cdot - t\frac{y}{m})\right\|_{L^r} \\ &\le \frac{1}{m}\left(\int_{\mathbb{R}^d} |y|\,\eta(y)dy\right)\|\nabla u\|_{L^r} \le \frac{1}{m}\|u\|_{W^{1,r}(\mathbb{R}^d)}. \end{aligned} \tag{49.26}$$

By the interpolation inequality in Corollary 10.25, Theorem 49.14 and Eq. (49.26) with $r = 1$,

$$\begin{aligned}\|v_m\|_{L^q} &\le \|v_m\|_{L^1}^{\lambda}\|v_m\|_{L^{p^*}}^{1-\lambda} \\ &\le \frac{1}{m^\lambda}\|u\|_{W^{1,1}(\mathbb{R}^d)}^{\lambda}\left[d^{-1/2}\frac{p(d-1)}{d-p}\|\nabla v_m\|_{L^p}\right]^{1-\lambda} \\ &\le Cm^{-\lambda}\|u\|_{W^{1,1}(\mathbb{R}^d)}^{\lambda}\|v_m\|_{W^{1,p}(\mathbb{R}^d)}^{1-\lambda} \\ &\le Cm^{-\lambda}\|u\|_{W^{1,1}(\mathbb{R}^d)}^{\lambda}\|v_m\|_{W^{1,p}(\mathbb{R}^d)}^{1-\lambda} \\ &\le C(p,|U|)m^{-\lambda}\|u\|_{W^{1,p}(\mathbb{R}^d)}^{\lambda}\|v_m\|_{W^{1,p}(\mathbb{R}^d)}^{1-\lambda}\end{aligned}$$

where $\lambda \in (0,1)$ is determined by

$$\frac{1}{q} = \frac{\lambda}{1} + \frac{1-\lambda}{p^*} = \lambda\left(1 - \frac{1}{p^*}\right) + \frac{1}{p^*}.$$

Now using Proposition 11.12,

$$\begin{aligned}\|v_m\|_{W^{1,p}(\mathbb{R}^d)} &= \|\eta_m * u - u\|_{W^{1,p}(\mathbb{R}^d)} \\ &\le \|\eta_m * u\|_{W^{1,p}(\mathbb{R}^d)} + \|u\|_{W^{1,p}(\mathbb{R}^d)} \le 2\|u\|_{W^{1,p}(\mathbb{R}^d)}.\end{aligned}$$

Putting this all together shows

$$\begin{aligned}\|K_m u - u\|_{L^q(\Omega)} \le \|K_m u - Eu\|_{L^q} &= \|v_m\|_{L^q} \\ &\le C(p,|U|)m^{-\lambda}\|u\|_{W^{1,p}(\mathbb{R}^d)}^{\lambda}\left(2\|u\|_{W^{1,p}(\mathbb{R}^d)}\right)^{1-\lambda} \\ &\le C(p,|U|)m^{-\lambda}\|Eu\|_{W^{1,p}(\mathbb{R}^d)} \\ &\le C(p,|U|)m^{-\lambda}\|u\|_{W^{1,p}(\Omega)}\end{aligned}$$

from which it follows that

$$\|K_m - i\|_{B(W^{1,p}(\Omega),L^q(\Omega))} \le Cm^{-\lambda} \to 0 \text{ as } m \to \infty.$$

■

Theorem 49.25 (Rellich - Kondrachov Compactness Theorem). *Suppose $\Omega \subset \mathbb{R}^d$ is a precompact open subset with C^1-boundary, $p \in [1, d)$ and $1 \le q < p^*$ then $W^{1,p}(\Omega)$ is compactly embedded in $L^q(\Omega)$.*

Proof. If $\{u_n\}_{n=1}^\infty$ is contained in the unit ball in $W^{1,p}(\Omega)$, then by Lemma 49.23 $\{K_m u_n\}_{n=1}^\infty$ has a uniformly convergent subsequence and hence is convergent in $L^q(\Omega)$. This shows $K_m : W^{1,p}(\Omega) \to L^q(\Omega)$ is compact for every m. By Lemma 49.24, $K_m \to i$ in the $L\left(W^{1,p}(\Omega), L^q(\Omega)\right)$ – norm and so by Lemma 49.22 $i : W^{1,p}(\Omega) \to L^q(\Omega)$ is compact. ■

Corollary 49.26. *The inclusion of $W^{k,p}(\Omega)$ into $W^{k-\ell,q}(\Omega)$ is compact provided $l \ge 1$ and $\frac{1}{q} > \frac{1}{p} - \frac{l}{d} = \frac{d-pl}{dp} > 0$, i.e. $q < \frac{dp}{d-pl}$.*

Proof. Case (i) Suppose $\ell = 1$, $q \in [1, p^*)$ and $\{u_n\}_{n=1}^\infty \subset W^{k,p}(\Omega)$ is bounded. Then $\{\partial^\alpha u_n\}_{n=1}^\infty \subset W^{1,p}(\Omega)$ is bounded for all $|\alpha| \le k-1$ and therefore there exist a subsequence $\{\tilde{u}_n\}_{n=1}^\infty \subset \{u_n\}_{n=1}^\infty$ such that $\partial^\alpha \tilde{u}_n$ is convergent in $L^q(\Omega)$ for all $|\alpha| \le k-1$. This shows that $\{\tilde{u}_n\}$ is $W^{k-1,q}(\Omega)$ – convergent and so proves this case.

Case (ii) $\ell > 1$. Let $\tilde{p}$ be defined so that $\frac{1}{\tilde{p}} = \frac{1}{p} - \frac{\ell-1}{d}$. Then

$$W^{k,p}(\Omega) \subset W^{k-\ell+1,\tilde{p}}(\Omega) \subset\subset W^{k-\ell,q}(\Omega).$$

and therefore $W^{k,p}(\Omega) \subset\subset W^{k-\ell,q}(\Omega)$. ■

Example 49.27. It is necessary to assume that The inclusion of $L^2([0,1]) \hookrightarrow L^1([0,1])$ is continuous (in fact a contraction) but not compact. To see this, take $\{u_n\}_{n=1}^\infty$ to be the Haar basis for L^2. Then $u_n \to 0$ weakly in both L^2 and L^1 so if $\{u_n\}_{n=1}^\infty$ were to have a convergent subsequence the limit would have to be $0 \in L^1$. On the other hand, since $|u_n| = 1$, $\|u_n\|_2 = \|u_n\|_1 = 1$ and any subsequential limit would have to have norm one and in particular not be 0.

Lemma 49.28. *Let Ω be a precompact open set such that $\bar{\Omega}$ is a manifold with C^1 – boundary. Then for all $p \in [1,\infty)$, $W^{1,p}(\Omega)$ is compactly embedded in $L^p(\Omega)$. Moreover if $p > d$ and $0 \le \beta < 1 - \frac{d}{p}$, then $W^{1,p}(\Omega)$ is compactly embedded in $C^{0,\beta}(\Omega)$. In particular, $W^{1,p}(\Omega) \subset\subset L^\infty(\Omega)$ for all $d < p \le \infty$.*

Proof. Case 1, $p \in [1,d)$. By Theorem 49.25, $W^{1,p}(\Omega) \subset\subset L^q(\Omega)$ for all $1 \le q < p^*$. Since $p^* > p$ we may choose $q = p$ to learn $W^{1,p}(\Omega) \subset\subset L^p(\Omega)$.

Case 2, $p \in [d,\infty)$. For any $p_0 \in [1,d)$, we have

$$W^{1,p}(\Omega) \hookrightarrow W^{1,p_0}(\Omega) \subset\subset L^{p_0^*}(\Omega).$$

Since $p_0^* = \frac{p_0 d}{d-p_o} \uparrow \infty$ as $p_0 \uparrow d$, we see that $W^{1,p}(\Omega) \subset\subset L^q(\Omega)$ for all $q < \infty$. Moreover by Morrey's inequality (Corollary 49.8) and Proposition5.13 we have $W^{1,p}(\Omega) \hookrightarrow C^{0,1-\frac{d}{p}}(\Omega) \subset\subset C^{0,\beta}(\Omega)$ which completes the proof. ■

Remark 49.29. Similar proofs may be given to show $W^{k,p} \subset\subset C^{k-\frac{d}{p}-\delta}$ for all $\delta > 0$ provided $k - \frac{d}{p} > 0$ and $k - \frac{d}{p} - \delta > 0$.

Lemma 49.30 (Poincaré Lemma). *Assume $1 \le p \le \infty$, Ω is a precompact open connected subset of $\mathbb{R}^d$ such that $\overline{\Omega}$ is a manifold with C^1-boundary. Then exist $C = C(\Omega, \rho)$ such that*

$$\|u - u_\Omega\|_{L^p(\Omega)} \le C\|\nabla u\|_{L^p(\Omega)} \text{ for all } u \in W^{1,p}(\Omega), \tag{49.27}$$

where $u_\Omega := \fint_\Omega u dm$ is the average of u on Ω as in Notation 49.3.

Proof. For sake of contradiction suppose there is no $C < \infty$ such that Eq. (49.27) holds. Then there exists a sequence $\{u_n\}_{n=1}^\infty \subset W^{1,p}(\Omega)$ such that

$$\|u_n - (u_n)_\Omega\|_{L^p(\Omega)} > n\|\nabla u_n\|_{L^p(\Omega)} \text{ for all } n.$$

Let

$$u_n := \frac{u_n - (u_n)_\Omega}{\|u_n - (u_n)_\Omega\|_{L^p(\Omega)}}.$$

Then $u_n \in W^{1,p}(\Omega)$, $(u_n)_\Omega = 0$, $\|u_n\|_{L^p(\Omega)} = 1$ and $1 = \|u_n\|_{L^p(\Omega)} > n\|\nabla u_n\|_{L^p(\Omega)}$ for all n. Therefore $\|\nabla u_n\|_{L^p(\Omega)} < \frac{1}{n}$ and in particular $\sup_n \|u_n\|_{W^{1,p}(\Omega)} < \infty$ and hence by passing to a subsequence if necessary there exists $u \in L^p(\Omega)$ such that $u_n \to u$ in $L^p(\Omega)$. Since $\nabla u_n \to 0$ in $L^p(\Omega)$, it follows that u_n is convergent in $W^{1,p}(\Omega)$ and hence $u \in W^{1,p}(\Omega)$ and $\nabla u = \lim_{n\to\infty} \nabla u_n = 0$ in $L^p(\Omega)$. Since $\nabla u = 0$, $u \in W^{k,p}(\Omega)$ for all $k \in \mathbb{N}$ and hence $u \in C^\infty(\Omega)$ and $\nabla u = 0$ and Ω is connected implies u is constant. Since $0 = \lim_{n\to\infty}(u_n)_\Omega = u_\Omega$ we must have $u \equiv 0$ which is clearly impossible since $\|u\|_{L^p(\Omega)} = \lim_{n\to\infty} \|u_n\|_{L^p(\Omega)} = 1$. ■

Theorem 49.31 (Poincaré Lemma). *Let Ω be a precompact open subset of $\mathbb{R}^d$ and $p \in [1, \infty)$. Then*

$$\|u\|_{L^p} \le \text{diam}(\Omega)\|\nabla u\|_{L^p} \text{ for all } u \in W_0^{1,p}(\Omega). \tag{49.28}$$

Proof. Let $\text{diam}(\Omega) = M$. By translating Ω if necessary we may assume $\Omega \subset [-M, M]^d$. For $1 \le p < \infty$ we may assume $u \in C_c^\infty(\Omega)$ since $C_c^\infty(\Omega)$ is dense in $W_0^{1,p}(\Omega)$. Then by the fundamental theorem of calculus,

$$\begin{aligned} |u(x)| &= \frac{1}{2}\left|\int_{-M}^{x_1} \partial_1 u(y, x_2, \ldots, x_d)dy - \int_{x_1}^M \partial_1 u(y, x_2, \ldots, x_d)dy\right| \\ &\le \frac{1}{2}\int_{-M}^M |\partial_1 u(y, x_2, \ldots, x_d)|dy = M\int_{-M}^M |\partial_1 u(y, x_2, \ldots, x_d)|\frac{dy}{2M} \end{aligned}$$

and hence by Jensen's inequality,

$$|u(x)|^p \le M^p \int_{-M}^M |\partial_1 u(y, x_2, \ldots, x_d)|^p \frac{dy}{2M}$$

Integrating this equation over x implies,

$$\|u\|_{L^p}^p \le M^p \int_\Omega |\partial_1 u(x)|^p dx \le M^p \int_\Omega |\nabla u(x)|^p dx$$

which gives Eq. (49.28). ■

49.6 Fourier Transform Method

See L^2 – Sobolev spaces for another proof of the following theorem.

Theorem 49.32. *Suppose $s > t \ge 0$, $\{u_n\}_{n=1}^\infty$ is a bounded sequence (say by 1) in $H^s(\mathbb{R}^d)$ such that $K = \overline{\cup_n \text{supp}(u_n)} \sqsubset\sqsubset \mathbb{R}^d$. Then there exist a subsequence $\{v_n\}_{n=1}^\infty \subset \{u_n\}_{n=1}^\infty$ which is convergent in $H^t(\mathbb{R}^d)$.*

Proof. Since

$$\begin{aligned} \left|\partial_\xi^\alpha \hat{u}_n(\xi)\right| &= \left|\partial_\xi^\alpha \int_{\mathbb{R}^d} e^{-i\xi\cdot x} u_n(x)dx\right| = \left|\int_{\mathbb{R}^d} (-ix)^\alpha e^{-i\xi\cdot x} u_n(x)dx\right| \\ &\le \|x^\alpha\|_{L^2(K)} \|u_n\|_{L^2} \le C_\alpha \|u_n\|_{H^s(\mathbb{R}^d)} \le C_\alpha \end{aligned}$$

$\hat{u}_n$ and all of it's derivatives are uniformly bounded. By the Arzela-Ascoli theorem and Cantor's Diagonalization argument, there exists a subsequence $\{v_n\}_{n=1}^\infty \subset \{u_n\}_{n=1}^\infty$ such that $\hat{v}_n$ and all of its derivatives converge uniformly on compact subsets in ξ –space. If $\hat{v}(\xi) := \lim_{n\to\infty} \hat{v}_n(\xi)$, then by the dominated convergence theorem,

$$\begin{aligned} \int_{|\xi|\le R} (1 + |\xi|^2)^s |\hat{v}(\xi)|^2 d\xi &= \lim_{n\to\infty} \int_{|\xi|\le R} (1 + |\xi|^2)^s |\hat{v}_n(\xi)|^2 d\xi \\ &\le \limsup_{n\to\infty} \|v_n\|_{H^s(\mathbb{R}^d)}^2 \le 1. \end{aligned}$$

Since R is arbitrary this implies $\hat{v} \in L^2((1 + |\xi|^2)^s d\xi)$ and $\|v\|_{H^s(\mathbb{R}^d)} \le 1$. Set $g_n := v - v_n$ while $v = \mathcal{F}^{-1}\hat{v}$. Then $\{g_n\}_{n=1}^\infty \subset H^s(\mathbb{R}^d)$ and we wish to show $g_n \to 0$ in $H^t(\mathbb{R}^d)$. Let $d\mu_t(\xi) = (1 + |\xi|^2)^t d\xi$, then for any $R < \infty$,

$$\begin{aligned} \|g_n\|_{H^t}^2 &= \int |\hat{g}(\xi) - \hat{g}_n(\xi)|^2 d\mu_t(\xi) \\ &= \int_{|\xi|\le R} |\hat{g}(\xi) - \hat{g}_n(\xi)|^2 d\mu_t(\xi) + \int_{|\xi|\ge R} |\hat{g}(\xi) - \hat{g}_n(\xi)|^2 d\mu_t(\xi). \end{aligned}$$

The first term goes to zero by the dominated convergence theorem, hence

$$\begin{aligned}\limsup_{n\to\infty}\|g_n\|_{H^t}^2 &\le \limsup_{n\to\infty}\int_{|\xi|\ge R}|\hat g(\xi)-\hat g_n(\xi)|^2 d\mu_t(\xi)\\ &=\limsup_{n\to\infty}\int_{|\xi|\ge R}|\hat g-\hat g_n(\xi)|^2\frac{(1+|\xi|^2)^s}{(1+|\xi|^2)^{s-t}}d\xi\\ &\le \limsup_{n\to\infty}\frac{1}{(1+R^2)^{s-t}}\int_{|\xi|\ge R}|\hat g-\hat g_n(\xi)|^2 d\mu_s(\xi)\\ &\le \limsup_{n\to\infty}\frac{1}{(1+R^2)^{s-t}}\|g_n-g\|_{H^t}^2\\ &\le 4\left(\frac{1}{1+R^2}\right)^{s-t}\to 0 \text{ as } R\to\infty.\end{aligned}$$

■

49.7 Other theorems along these lines

Another theorem of this form is derived as follows. Let $\rho>0$ be fixed and $g\in C_c((0,1),[0,1])$ such that $g(t)=1$ for $|t|\le 1/2$ and set $\tau(t):=g(t/\rho)$. Then for $x\in\mathbb{R}^d$ and $\omega\in\Gamma$ we have

$$\int_0^\rho \frac{d}{dt}[\tau(t)u(x+t\omega)]\,dt=-u(x)$$

and then by integration by parts repeatedly we learn that

$$\begin{aligned}u(x)&=\int_0^\rho \partial_t^2[\tau(t)u(x+t\omega)]\,tdt=\int_0^\rho \partial_t^2[\tau(t)u(x+t\omega)]\,d\frac{t^2}{2}\\ &=-\int_0^\rho \partial_t^3[\tau(t)u(x+t\omega)]\,d\frac{t^3}{3!}=\dots\\ &=(-1)^m\int_0^\rho \partial_t^m[\tau(t)u(x+t\omega)]\,d\frac{t^m}{m!}\\ &=(-1)^m\int_0^\rho \partial_t^m[\tau(t)u(x+t\omega)]\,\frac{t^{m-1}}{(m-1)!}dt.\end{aligned}$$

Integrating this equation on $\omega\in\Gamma$ then implies

$$\begin{aligned}|\Gamma|\,u(x)&=(-1)^m\int_\gamma d\omega\int_0^\rho \partial_t^m[\tau(t)u(x+t\omega)]\,\frac{t^{m-1}}{(m-1)!}dt\\ &=\frac{(-1)^m}{(m-1)!}\int_\gamma d\omega\int_0^\rho t^{m-d}\partial_t^m[\tau(t)u(x+t\omega)]\,t^{d-1}dt\\ &=\frac{(-1)^m}{(m-1)!}\int_\gamma d\omega\int_0^\rho t^{m-d}\sum_{k=0}^m\binom{m}{k}\left[\tau^{(m-k)}(t)\left(\partial_\omega^k u\right)(x+t\omega)\right]t^{d-1}dt\\ &=\frac{(-1)^m}{(m-1)!}\int_\gamma d\omega\int_0^\rho t^{m-d}\sum_{k=0}^m\binom{m}{k}\rho^{k-m}\left[g^{(m-k)}(t)\left(\partial_\omega^k u\right)(x+t\omega)\right]t^{d-1}\\ &=\frac{(-1)^m}{(m-1)!}\sum_{k=0}^m\binom{m}{k}\rho^{k-m}\int_{\Gamma_{x,\rho}}|y-x|^{m-d}\left[g^{(m-k)}(|y-x|)\left(\partial_{\widehat{y-x}}^k u\right)(y)\right]\end{aligned}$$

and hence

$$u(x)=\frac{(-1)^m}{|\Gamma|\,(m-1)!}\sum_{k=0}^m\binom{m}{k}\rho^{k-m}\int_{\Gamma_{x,\rho}}|y-x|^{m-d}\left[g^{(m-k)}(|y-x|)\left(\partial_{\widehat{y-x}}^k u\right)(y)\right]$$

and hence by the Hölder's inequality,

$$|u(x)|\le C(g)\frac{(-1)^m}{|\Gamma|\,(m-1)!}\sum_{k=0}^m\binom{m}{k}\rho^{k-m}\left[\int_{\Gamma_{x,\rho}}|y-x|^{q(m-d)}dy\right]^{1/q}\left[\int_{\Gamma_{x,\rho}}\left|\left(\partial_{\widehat{y-}}^k\right.\right.\right.$$

From the same computation as in Eq. (48.4) we find

$$\begin{aligned}\int_{\Gamma_{x,\rho}}|y-x|^{q(m-d)}dy&=\sigma(\Gamma)\int_0^\rho r^{q(m-d)}r^{d-1}dr=\sigma(\Gamma)\frac{\rho^{q(m-d)+d}}{q(m-d)+d}\\ &=\sigma(\Gamma)\frac{\rho^{\frac{pm-d}{p-1}}}{pm-d}(p-1).\end{aligned}$$

provided that $pm-d>0$ (i.e. $m>d/p$) wherein we have used

$$q(m-d)+d=\frac{p}{p-1}(m-d)+d=\frac{p(m-d)+d(p-1)}{p-1}=\frac{pm-d}{p-1}.$$

This gives the estimate

$$\left[\int_{\Gamma_{x,\rho}}|y-x|^{q(m-d)}dy\right]^{1/q}\le\left[\frac{\sigma(\Gamma)(p-1)}{pm-d}\right]^{\frac{p-1}{p}}\rho^{\frac{pm-d}{p}}=\left[\frac{\sigma(\Gamma)(p-1)}{pm-d}\right]^{\frac{p-1}{p}}\rho^{m}$$

Thus we have obtained the estimate that

$$|u(x)|\le\frac{C(g)}{|\Gamma|\,(m-1)!}\left[\frac{\sigma(\Gamma)(p-1)}{pm-d}\right]^{\frac{p-1}{p}}\times\\ \rho^{m-d/p}\sum_{k=0}^m\binom{m}{k}\rho^{k-m}\left\|\partial_{\widehat{y-x}}^k u\right\|_{L^p(\Gamma_{x,p})}.$$

49.8 Exercises

Exercise 49.33. Let $a_i \geq 0$ and $p_i \in [1, \infty)$ for $i = 1, 2, \ldots, d$ satisfy $\sum_{i=1}^d p_i^{-1} = 1$, then

$$\prod_{i=1}^d a_i \leq \sum_{i=1}^d \frac{1}{p_i} a_i^{p_i}.$$

Hint: This may be proved by induction on d making use of Lemma 1.27. Alternatively see Example 10.11, where this is already proved using Jensen's inequality.

Solution 49.34 (49.33). We may assume that $a_i > 0$, in which case

$$\prod_{i=1}^d a_i = e^{\sum_{i=1}^d \ln a_i} = e^{\sum_{i=1}^d \frac{1}{p_i} \ln a_i^{p_i}} \leq \sum_{i=1}^d \frac{1}{p_i} e^{\ln a_i^{p_i}} = \sum_{i=1}^d \frac{1}{p_i} a_i^{p_i}.$$

This was already done in Example 10.11.

Part XV

Variable Coefficient Equations

50

2^{nd} order differential operators

Notations 50.1 *Let Ω be a precompact open subset of $\mathbb{R}^d$, $A_{ij} = A_{ji}, A_i, A_0 \in BC^\infty(\overline{\Omega})$ for $i, j = 1, \dots, d$,*

$$p(x, \xi) := -\sum_{i,j=1}^{d} A_{ij}\xi_i\xi_j + \sum_{i=1}^{d} A_i\xi_i + A_0$$

and

$$L = p(x, \partial) = -\sum_{i,j=1}^{d} A_{ij}\partial_i\partial_j + \sum_{i=1}^{d} A_i\partial_i + A_0.$$

We also let

$$L^\dagger = -\sum_{i,j=1}^{d} \partial_i\partial_j M_{A_{ij}} - \sum_{i=1}^{d} \partial_i M_{A_i} + A_0.$$

Remark 50.2. The operators L and $L^\dagger$ have the following properties.

1. The operator $L^\dagger$ is the formal adjoint of L, i.e.
$$(Lu, v) = (u, L^\dagger v) \text{ for all } u, v \in \mathcal{D}(\Omega) = C_c^\infty(\Omega).$$
2. We may view L as an operator on $\mathcal{D}'(\Omega)$ via the formula $u \in \mathcal{D}'(\Omega) \to Lu \in \mathcal{D}'(\Omega)$ where
$$\langle Lu, \phi\rangle := \langle u, L^\dagger\phi\rangle \text{ for all } \phi \in C_c^\infty(\Omega).$$
3. The restriction of L to $H^{k+2}(\Omega)$ gives a bounded linear transformation
$$L : H^{k+2}(\Omega) \to H^k(\Omega) \text{ for } k \in \mathbb{N}_0.$$
Indeed, L may be written as
$$L = -\sum_{i,j=1}^{d} M_{A_{ij}}\partial_i\partial_j + \sum_{i=1}^{d} M_{A_i}\partial_i + M_{A_0}.$$

Now $\partial_i : H^k(\Omega) \to H^{k+1}(\Omega)$ is bounded and $M_\psi : H^k(\Omega) \to H^k(\Omega)$ is bounded where $\psi \in BC^\infty(\overline{\Omega})$. Therefore, for $k \in \mathbb{N}_0$, $L : H^{k+2}(\Omega) \to H^k(\Omega)$ is bounded.

Definition 50.3. *For $u \in \mathcal{D}'(\Omega)$, let*

$$\|u\|_{H^{-1}(\Omega)} := \sup_{0\neq\phi\in\mathcal{D}(\Omega)} \frac{|\langle u, \phi\rangle|}{\|\phi\|_{H_0^1(\Omega)}}$$

and

$$H^{-1}(\Omega) := \left\{u \in \mathcal{D}'(\Omega) : \|u\|_{H^{-1}(\Omega)} < \infty\right\}.$$

Example 50.4. Let $\Omega = \mathbb{R}^d$ and $S \subset \Omega$ be the unit sphere in $\mathbb{R}^d$. Then define $\sigma \in \mathcal{D}'(\Omega)$ by

$$\langle\sigma, \phi\rangle := \int_S \phi d\sigma.$$

Let us shows that $\sigma \in H^{-1}(\Omega)$. For this let $T : H^1(\Omega) \to L^2(S, d\sigma)$ denote the trace operator, i.e. the unique bounded linear operator such that $T\phi = \phi|_S$ for all $\phi \in C_c^\infty(\mathbb{R}^d)$. Since T is bounded,

$$|\langle\sigma, \phi\rangle| \le \sigma(S)^{1/2} \|T\phi\|_{L^2(S)} \le \sigma(S)^{1/2} \|T\|_{L(H^1(\Omega), L^2(S))} \|\phi\|_{H^{-1}(\Omega)}.$$

This shows $\sigma \in H^{-1}(\Omega)$ and $\|\sigma\|_{H^{-1}(\Omega)} \le \sigma(S)^{1/2} \|T\|_{L(H^1(\Omega), L^2(S))}$.

Lemma 50.5. *Suppose Ω is an open subset of $\mathbb{R}^d$ such that $\bar{\Omega}$ is a manifold with C^0 – boundary and $\Omega = \bar{\Omega}^o$, then the map $u \in \left[H_0^1(\Omega)\right]^* \to u|_{\mathcal{D}(\Omega)} \in H^{-1}(\Omega)$ is a unitary map of Hilbert spaces.*

Proof. By definition $C_c^\infty(\Omega)$ is dense in $H_0^1(\Omega)$, and hence it follows that the map $u \in \left[H_0^1(\Omega)\right]^* \to u|_{\mathcal{D}(\Omega)} \in H^{-1}(\Omega)$ is isometric. If $u \in H^{-1}(\Omega)$, it has a unique extension to $H_0^1(\Omega) = \overline{C_c^\infty(\Omega)}^{H^1(\Omega)}$ and this provides the inverse map. ■

If we identify $L^2(\Omega) = H^0(\Omega)$ with elements of $\mathcal{D}'(\Omega)$ via $u \to (u, \cdot)_{L^2(\Omega)}$, then

$$\mathcal{D}'(\Omega) \supset H^{-1}(\Omega) \supset H^0(\Omega) = L^2(\Omega) \supset H^1(\Omega) \supset H^2(\Omega) \supset \dots$$

Proposition 50.6. *The following mapping properties hold:*

1. *If $\chi \in BC^1(\overline{\Omega})$. Then $M_\chi : H^{-1}(\Omega) \to H^{-1}(\Omega)$ is a bounded operator.*
2. *If $V = \sum_{i=1}^d M_{A_i}\partial_i + M_{A_0}$ with $A_i, A_0 \in BC^1(\bar{\Omega})$, then $V : L^2(\Omega) \to H^{-1}(\Omega)$ is a bounded operator.*
3. *The map $L : \mathcal{D}'(\Omega) \to \mathcal{D}'(\Omega)$ restricts to a bounded linear map from $H^1(\Omega)$ to $H^{-1}(\Omega)$. Also*

Proof. Let us begin by showing $M_\chi : H_0^1(\Omega) \to H_0^1(\Omega)$ is a bounded linear map. In order to do this choose $\chi_n \in C_c^\infty(\mathbb{R}^d)$ such that $\chi_n \to \chi$ in $BC^1(\overline{\Omega})$. Then for $\phi \in C_c^\infty(\Omega)$, $\chi_n\phi \in C_c^\infty(\Omega) \subset H_0^1(\Omega)$ and there is a constant $K < \infty$ such that

$$\|\chi_n\phi\|^2_{H_0^1(\Omega)} \le K \|\chi_n\|_{BC^1(\overline{\Omega})} \|\phi\|^2_{H_0^1(\Omega)}.$$

By density this estimate holds for all $\phi \in H_0^1(\Omega)$ and by replacing χ_n by $\chi_n - \chi_m$ we also learn that

$$\|(\chi_n - \chi_m)\phi\|^2_{H_0^1(\Omega)} \le K \|\chi_n - \chi_m\|_{BC^1(\overline{\Omega})} \|\phi\|^2_{H_0^1(\Omega)} \to 0 \text{ as } m,n \to \infty.$$

By completeness of $H_0^1(\Omega)$ it follows that $\chi\phi \in H_0^1(\Omega)$ for all $\phi \in H_0^1(\Omega)$ and

$$\|\chi_n\phi\|^2_{H_0^1(\Omega)} \le K \|\chi\|_{BC^1(\overline{\Omega})} \|\phi\|^2_{H_0^1(\Omega)}.$$

1. If $u \in H^{-1}(\Omega)$ and $\phi \in H_0^1(\Omega)$, then by definition, $\langle M_\chi u, \phi\rangle = \langle u, \chi\phi\rangle$ and therefore,
$$\begin{aligned}|\langle M_\chi u, \phi\rangle| &= |\langle u, \chi\phi\rangle| \le \|u\|_{H^{-1}(\Omega)} \|\chi\phi\|_{H_0^1(\Omega)} \\ &\le K\|\chi\|_{BC^1(\overline{\Omega})} \|u\|_{H^{-1}(\Omega)} \|\phi\|_{H_0^1(\Omega)}\end{aligned}$$
which implies $M_\chi u \in H^{-1}(\Omega)$ and
$$\|M_\chi u\|_{H^{-1}(\Omega)} \le K\|\chi\|_{BC^1(\overline{\Omega})} \|u\|_{H^{-1}(\Omega)}.$$
2. For $u \in L^2(\Omega)$ and $\phi \in C_c^\infty(\Omega)$
$$\begin{aligned}|\langle \partial_i u, \phi\rangle| &= |\langle u, \partial_i\phi\rangle| \le \|u\|_{L^2(\Omega)} \cdot \|\partial_i\phi\|_{L^2(\Omega)} \\ &\le \|u\|_{L^2(\Omega)} \|\phi\|_{H_0^1(\Omega)}\end{aligned}$$
and therefore $\|\partial_i u\|_{H^{-1}(\Omega)} \le \|u\|_{L^2(\Omega)}$. For general $V = \sum_{i=1}^d M_{A_i}\partial_i + M_{A_0}$, we have
$$\begin{aligned}\|Au\|_{H^{-1}(\Omega)} &\le \sum_{i=1}^d K\|A_i\|_{BC^1(\overline{\Omega})} \|\partial_i u\|_{H^{-1}(\Omega)} + \|A_0\|_\infty \|u\|_{L^2(\Omega)} \\ &\le \left[\sum_{i=1}^d K\|A_i\|_{BC^1(\overline{\Omega})} + \|A_0\|_\infty\right] \|u\|_{L^2(\Omega)}.\end{aligned}$$
3. Since $V : H^1(\Omega) \to L^2(\Omega)$ and $i : L^2(\Omega) \hookrightarrow H^{-1}(\Omega)$ are both bounded maps, to prove $L = -\sum_{i,j=1}^d M_{A_{ij}}\partial_i\partial_j + V$ is bounded from $H^1(\Omega) \to H^{-1}(\Omega)$ it suffices to show $M_{A_{ij}}\partial_i\partial_j : H^1(\Omega) \to H^{-1}(\Omega)$ is a bounded. But $M_{A_{ij}}\partial_i\partial_j : H^1(\Omega) \to H^{-1}(\Omega)$ is bounded since it is the composition of the following bounded maps:
$$H^1(\Omega) \overset{\partial_j}{\to} L^2(\Omega) \overset{\partial_i}{\to} H^{-1}(\Omega) \overset{M_{A_{ij}}}{\to} H^{-1}(\Omega).$$

■

Lemma 50.7. *Suppose $\chi \in BC^\infty(\overline{\Omega})$ then*

1. *$[L, M_\chi] = V$ is a first order differential operator acting on $\mathcal{D}'(\Omega)$ which necessarily satisfies $V : H^k(\Omega) \to H^{k-1}(\Omega)$ for $k = 0, 1, 2, \ldots$ etc.*
2. *If $u \in H^k(\Omega)$, then*
$$[L, M_\chi]u \in H^{k-1}(\Omega) \text{ for } k = 0, 1, 2, \ldots$$
and
$$\|[L, M_\chi]u\|_{H^{k-1}(\Omega)} \le C_k(\chi)\|u\|_{H^k(\Omega)}.$$

Proof. On smooth functions $u \in C^\infty(\Omega)$,

$$L(\chi u) = \chi L u - 2\sum_{i,j=1}^d A_{ij}\partial_i\chi \cdot \partial_j u + \left(\sum_{i=1}^d A_i\partial_i\chi - \sum_{i,j=1}^d A_{ij}\partial_i\partial_j\chi\right) \cdot u$$

and therefore

$$\begin{aligned}[L, M_\chi]u &= -2\sum_{i,j=1}^d A_{ij}\partial_i\chi \cdot \partial_j u + \left(\sum_{i=1}^d A_i\partial_i\chi - \sum_{i,j=1}^d A_{ij}\partial_i\partial_j\chi\right) \cdot u \\ &=: Vu.\end{aligned}$$

Similarly,

$$\begin{aligned}L^\dagger(\chi u) &= -\sum_{i,j=1}^d \partial_i\partial_j[\chi A_{ij}u] - \sum_{i=1}^d \partial_i(\chi A_i u) + A_0\chi u \\ &= \chi L^\dagger u - 2\sum_{i,j=1}^d \partial_i\chi \cdot \partial_j[A_{ij}u] \\ &\quad - \sum_{i=1}^d A_i\partial_i\chi \cdot u - \left(\sum_{i,j=1}^d A_{ij}\partial_i\partial_j\chi\right) u. \qquad (50.1)\end{aligned}$$

Noting that

$$\begin{aligned}V^\dagger u &= 2\sum_{i,j=1}^d \partial_j[\partial_i\chi \cdot A_{ij}u] + \left(\sum_{i=1}^d A_i\partial_i\chi - \sum_{i,j=1}^d A_{ij}\partial_i\partial_j\chi\right) \cdot u \\ &= 2\sum_{i,j=1}^d \partial_i\chi \cdot \partial_j[A_{ij}u] + \left(\sum_{i=1}^d A_i\partial_i\chi + \sum_{i,j=1}^d A_{ij}\partial_i\partial_j\chi\right) \cdot u,\end{aligned}$$

Eq. (50.1) may be written as

$$[L^\dagger, M_\chi] = -V^\dagger.$$

Now suppose $k = 0$, then in this case for $\phi \in \mathcal{D}(\Omega)$,

$$\begin{aligned}|\langle [L, M_\chi]u, \phi\rangle| &= \left|\langle u, [M_\chi, L^\dagger]\phi\rangle\right| = \left|\langle u, V^\dagger\phi\rangle\right| \\ &\le \|u\|_{L^2(\Omega)}\|V^\dagger\phi\|_{L^2(\Omega)} \le C\|u\|_{L^2(\Omega)}\|\phi\|_{H_0^1(\Omega)}.\end{aligned}$$

This implies $\|[L, M_\chi]u\|_{H^{-1}(\Omega)} \le C\|u\|_{L^2}$ and in particular $[L, M_\chi]u \in H^{-1}(\Omega)$. For $k > 0$, $[L, M_\chi]u = Vu$ with V as above and therefore by Proposition48.6, there exists $C < \infty$ such that $\|Vu\|_{H^{k-1}(\Omega)} \le C\|u\|_{H^k(\Omega)}$. ■

Definition 50.8. *The operator L is* ***uniformly elliptic*** *on Ω if there exists $\epsilon > 0$ such that $(A_{ij}(x)) \ge \epsilon I$ for all $x \in \Omega$, i.e. $A_{ij}(x)\xi_i\xi_j \ge \epsilon|\xi|^2$ for all $x \in \Omega$ and $\xi \in \mathbb{R}^d$.*

Suppose now that L is uniformly elliptic. Let us outline the results to be proved below.

50.1 Outline of future results

1. We consider L with Dirichlet boundary conditions meaning we will view L as a mapping from $H_0^1(\Omega) \to H^{-1}(\Omega) = \left[H_0^1(\Omega)\right]^*$. Proposition 51.13 below states there exists $C = C(L) < \infty$ such that $(L + C) : H_0^1(\Omega) \to H^{-1}(\Omega)$ is an isomorphism of Hilbert spaces. The proof uses the Dirichlet form
$$\mathcal{E}(u, v) := \langle Lu, v\rangle \text{ for } u, v \in H_0^1(\Omega).$$
Notice for $v \in \mathcal{D}(\Omega)$ and $u \in H_0^1(\Omega)$,
$$\begin{aligned}\mathcal{E}(u, v) &= \langle Lu, v\rangle = \langle u, L^\dagger v\rangle \\ &= \int_\Omega u\left(-\partial_i\partial_j(A_{ij}v) - \partial_i(A_iv) + A_0v\right)dm \\ &= \int_\Omega \left[\partial_iu\cdot\partial_j(A_{ij}v) - u\partial_i(A_iv) + uA_0v\right]dm \\ &= \int_\Omega \left[A_{ij}\partial_iu\cdot\partial_jv + (A_i + \partial_jA_{ij})\partial_iu\cdot v + A_0uv\right]dm.\end{aligned}$$
Since the last expression is continuous for $(u, v) \in H_0^1(\Omega)\times H_0^1(\Omega)$, we have shown
$$\mathcal{E}(u, v) = \int_\Omega \left[A_{ij}\partial_iu\cdot\partial_jv + (A_i + \partial_jA_{ij})\partial_iu\cdot v + A_0uv\right]dm$$
for all $u, v \in H_0^1(\Omega)$.

2. To implement other boundary conditions, we will need to consider L acting on subspaces of $H^2(\Omega)$ which are determined by the boundary conditions. Rather than describe the general case here, let us consider an example where $L = -\Delta$ and the boundary condition is $\frac{\partial u}{\partial n} = \rho u$ on $\partial\Omega$ where $\partial_nu = \nabla u\cdot n$, n is the outward normal on $\partial\Omega$ and $\rho : \partial\Omega \to \mathbb{R}$ is a smooth function. In this case, let
$$D := \left\{u \in H^2(\Omega) : \frac{\partial u}{\partial n} = \rho u \text{ on } \partial\Omega\right\}.$$
We will eventually see that D is a dense subspace of $H^1(\Omega)$. For $u \in D$ and $v \in H^1(\Omega)$,
$$\begin{aligned}(-\Delta u, v) &= \int_\Omega \nabla u\cdot\nabla v\, dm - \int_{\partial\Omega} v\partial_nu\, d\sigma \\ &= \int_\Omega \nabla u\cdot\nabla v\, dm - \int_{\partial\Omega} \rho uv\, d\sigma =: \mathcal{E}(u, v).\end{aligned} \tag{50.2}$$
The latter expression extends by continuity to all $u \in H^1(\Omega)$. Given $\mathcal{E}$ as in Eq. (50.2) let $-\Delta_{\mathcal{E}} : H^1(\Omega) \to \left[H^1(\Omega)\right]^*$ be defined by $-\Delta_{\mathcal{E}}u := \mathcal{E}(u, \cdot)$ so that $-\Delta_{\mathcal{E}}u$ is an extension of $-\Delta u$ as a linear functional on $H_0^1(\Omega)$ to one on $H^1(\Omega) \supset H_0^1(\Omega)$. It will be shown below that there exists $C < \infty$ such that $(-\Delta_{\mathcal{E}} + C) : H^1(\Omega) \to \left[H^1(\Omega)\right]^*$ is an isomorphism of Hilbert spaces.
3. The Dirichlet form $\mathcal{E}$ in Eq. (50.2) may be rewritten in a way as to avoid the surface integral term. To do this, extend the normal vector field n along $\partial\Omega$ to a smooth vector field on $\bar{\Omega}$. Then by integration by parts,
$$\begin{aligned}\int_{\partial\Omega} \rho uv\, d\sigma &= \int_{\partial\Omega} n_i^2\rho uv\, d\sigma = \int_\Omega \partial_i\left[n_i\rho uv\right]dm \\ &= \left[\int_\Omega \nabla\cdot(\rho n)\; uv + \rho n_i\partial_iu\cdot v + \rho n_iu\cdot\partial_iv\right]dm.\end{aligned}$$
In this way we see that the Dirichlet form $\mathcal{E}$ in Eq. (50.2) may be written as
$$\mathcal{E}(u, v) = \int_\Omega \left[\nabla u\cdot\nabla v + a_{i0}\partial_iu\cdot v + a_{0i}u\partial_iv + a_{00}uv\right]\, dm \tag{50.3}$$
with $a_{00} = \nabla\cdot(\rho n)$, $a_{i0} = \rho n_i = a_{0i}$. This should motivate the next section where we consider generalizations of the form $\mathcal{E}$ in Eq. (50.3).

51

Dirichlet Forms

In this section Ω will be an open subset of $\mathbb{R}^d$.

51.1 Basics

Notation 51.1 (Dirichlet Forms) *For $\alpha, \beta \in \mathbb{N}_0^d$ with $|\alpha|, |\beta| \le 1$, suppose $a_{\alpha,\beta} \in BC^\infty\left(\bar{\Omega}\right)$ and $\rho \in BC^\infty\left(\bar{\Omega}\right)$ with $\rho > 0$, let*

$$\mathcal{E}(u,v) = \sum_{|\alpha|,|\beta|\le 1} \int_\Omega a_{\alpha\beta}\partial^\alpha u \cdot \partial^\beta v \; d\mu \tag{51.1}$$

where $d\mu := \rho dm$. We will also write $(u,v) := \int_\Omega uv \; d\mu$ and L^2 for $L^2(\Omega,\mu)$. In the sequel we will often write $a_{i,\beta}$ for $a_{e_i,\beta}$, $a_{\alpha,j}$ for a_{α,e_j} and a_{ij} for a_{e_i,e_j}.

Proposition 51.2. *Let $\mathcal{E}$ be as in Notation 51.1 then*

$$|\mathcal{E}(u,v)| \le C\|u\|_{H^1}\|v\|_{H^1} \text{ for all } u,v \in H^1$$

where C is a constant depending on d and upper bounds for $\left\{\|a_{\alpha\beta}\|_{BC(\bar{\Omega})} : |\alpha|, |\beta| :\right.$

Proof. To simplify notation in the proof, let $\|\cdot\|$ denote the $L^2(\Omega,\mu)$ – norm. Then

$$\begin{aligned} |\mathcal{E}(u,v)| &\le C\sum_{ij}\{\|\partial_i u\| \; \|\partial_j v\| + \|\partial_i u\| \; \|v\| + \|u\| \; \|\partial_i v\| + \|u\| \; \|v\|\} \\ &\le C\|u\|_{H^1} \cdot \|v\|_{H^1}. \end{aligned}$$

■

Notation 51.3 *Given $\mathcal{E}$ as in Notation 51.1, let $\mathcal{L}_\mathcal{E}$ and $\mathcal{L}_\mathcal{E}^\dagger$ be the bounded linear operators from $H^1(\Omega)$ to $\left[H^1(\Omega)\right]^*$ defined by*

$$\mathcal{L}_\mathcal{E} u := \mathcal{E}(u,\cdot) \text{ and } \mathcal{L}_\mathcal{E}^\dagger u := \mathcal{E}(\cdot,u).$$

It follows directly from the definitions that $\langle \mathcal{L}_\mathcal{E} u, v\rangle = \langle u, \mathcal{L}_\mathcal{E}^\dagger v\rangle$ for all $u, v \in H^1(\Omega)$. The Einstein summation convention will be used below when convenient.

Proposition 51.4. *Suppose Ω is a precompact open subset of $\mathbb{R}^d$ such that $\bar{\Omega}$ is a manifold with C^2 – boundary, Then for all $u \in H^2(\Omega)$ and $v \in H^1(\Omega)$,*

$$\langle \mathcal{L}_\mathcal{E} u, v\rangle = \mathcal{E}(u,v) = (Lu,v) + \int_{\partial\Omega} Bu \cdot v \; \rho d\sigma \tag{51.2}$$

and for all $u \in H^1(\Omega)$ and $v \in H^2(\Omega)$,

$$\langle u, \mathcal{L}_\mathcal{E}^\dagger v\rangle = \mathcal{E}(u,v) = (u, L^\dagger v) + \int_{\partial\Omega} u \cdot B^\dagger v \; \rho d\sigma, \tag{51.3}$$

where

$$B = n_j a_{ij}\partial_i + n_j a_{0j} = n \cdot a\nabla + n \cdot a_{0,\cdot}, \tag{51.4}$$

$$B^\dagger = n_i\left[a_{ij}\partial_j + a_{i0}\right] = an \cdot \nabla + n \cdot a_{\cdot,0}, \tag{51.5}$$

$$Lu := \rho^{-1}\sum_{|\alpha|,|\beta|\le 1}(-1)^{|\beta|}\partial^\beta\left[\rho a_{\alpha\beta}\partial^\alpha u\right] \tag{51.6}$$

and

$$L^\dagger v := \rho^{-1}\sum_{|\alpha|,|\beta|\le 1}(-1)^{|\alpha|}\partial^\alpha\left[\rho a_{\alpha\beta}\partial^\beta v\right] \tag{51.7}$$

We may also write L, $L^\dagger$ as

$$L = -a_{ij}\partial_j\partial_i + \left(a_{i0} - a_{0j} - \rho^{-1}\partial_j\left[\rho a_{ij}\right]\right)\partial_i + \left(a_{00} - \rho^{-1}\partial_j\left[\rho a_{0j}\right]\right), \tag{51.8}$$

$$L^\dagger = -a_{ij}\partial_i\partial_j + \left(a_{0j} - a_{j0} - \rho^{-1}\partial_i\left[\rho a_{ij}\right]\right)\partial_j + \left(a_{00} - \rho^{-1}\partial_i\left[\rho a_{i0}\right]\right). \tag{51.9}$$

Proof. Suppose $u \in H^2(\Omega)$ and $v \in H^1(\Omega)$, then by integration by parts,

$$\begin{aligned} \mathcal{E}(u,v) &= \sum_{|\alpha|,|\beta|\le 1}\int_\Omega (-1)^{|\beta|}\rho^{-1}\partial^\beta\left[\rho a_{\alpha\beta}\partial^\alpha u\right] \cdot v \; d\mu + \sum_{|\alpha|\le 1}\sum_{j=1}^d \int_{\partial\Omega} n_j\left[a_{\alpha j}\partial^\alpha u\right] \cdot v \; \rho \\ &= (Lu,v) + \int_{\partial\Omega} Bu \cdot v \; \rho d\sigma, \end{aligned}$$

where

$$Lu = \rho^{-1} \sum_{|\alpha|,|\beta|\le 1} (-1)^{|\beta|} \partial^\beta \left[\rho a_{\alpha\beta}\partial^\alpha u\right] = -\rho^{-1} \sum_{|\alpha|\le 1} \sum_{j=1}^{d} \partial_j \left(\rho a_{\alpha j}\partial^\alpha u\right) + \sum_{|\alpha|\le 1} a_{\alpha 0}\hat{\partial}$$

$$= -\rho^{-1} \sum_{i,j=1}^{d} \partial_j \left(\rho a_{ij}\partial_i u\right) - \rho^{-1}\sum_{j=1}^{d} \partial_j \left(\rho a_{0j} u\right) + \sum_{i=1}^{d} a_{i0}\partial_i u + a_{00}u$$

$$= -\sum_{i,j=1}^{d} a_{ij}\partial_j\partial_i u - \rho^{-1}\sum_{i,j=1}^{d} \left(\partial_j\left[\rho a_{ij}\right]\right)\partial_i u - \sum_{j=1}^{d} a_{0j}\partial_j u$$

$$- \rho^{-1}\sum_{j=1}^{d}\left(\partial_j\left[\rho a_{0j}\right]\right) u + \sum_{i=1}^{d} a_{i0}\partial_i u + a_{00}u$$

and

$$Bu = \sum_{|\alpha|\le 1}\sum_{j=1}^{d} n_j \left(a_{\alpha j}\partial^\alpha u\right) = \sum_{i,j=1}^{d} n_j a_{ij}\partial_i u + \sum_{j=1}^{d} n_j a_{0j} u.$$

Similarly for $u \in H^1(\Omega)$ and $v \in H^2(\Omega)$,

$$\mathcal{E}(u,v) = \sum_{|\alpha|,|\beta|\le 1} \int_\Omega u\cdot (-1)^{|\alpha|}\rho^{-1}\partial^\alpha\left[\rho a_{\alpha\beta}\partial^\beta v\right]\, d\mu + \sum_{i=1}^{d}\sum_{|\beta|\le 1}\int_\Omega u\cdot n_i\left[a_{i\beta}\partial^\beta v\right]$$

$$= (u, L^\dagger v) + \int_{\partial\Omega} u\cdot B^\dagger v\ \rho d\sigma,$$

where $B^\dagger v = n_i\left[a_{ij}\partial_j + a_{i0}\right]$ and

$$\begin{aligned} L^\dagger v &= -\rho^{-1}\partial_i\left(\rho a_{ij}\partial_j v\right) + a_{0j}\partial_j v - \rho^{-1}\partial_i\left(\rho a_{i0} v\right) + a_{00}v \\ &= -a_{ij}\partial_i\partial_j v - \rho^{-1}\left(\partial_i\left[\rho a_{ij}\right]\right)\partial_j v \\ &\quad + a_{0j}\partial_j v - a_{i0}\partial_i v - \rho^{-1}\left(\partial_i\left[\rho a_{i0}\right]\right) v + a_{00}v \\ &= \left[-a_{ij}\partial_i\partial_j + \left(a_{0j} - a_{j0} - \rho^{-1}\partial_i\left[\rho a_{ij}\right]\right)\partial_j + a_{00} - \rho^{-1}\partial_i\left[\rho a_{i0}\right]\right] v. \end{aligned}$$

■

Proposition 51.4 shows that to the Dirichlet form $\mathcal{E}$ there is an associated second order elliptic operator L along with boundary conditions B as in Eqs. (51.6) and (51.4). The next proposition shows how to reverse this procedure and associate a Dirichlet form $\mathcal{E}$ to a second order elliptic operator L with boundary conditions.

Proposition 51.5 (Following Folland p. 240.). *Let $A_j, A_0\rho \in BC^\infty(\Omega)$ and $A_{ij} = A_{ji} \in BC^\infty(\Omega)$ with $(A_{ij}) > 0$ and $\rho > 0$ and let*

$$L = -A_{ij}\partial_i\partial_j + A_i\partial_i + A_0 \tag{51.10}$$

and $(u,v) := \int_\Omega uv\rho dm$. Also suppose $\alpha : \partial\Omega \to \mathbb{R}$ and $V : \partial\Omega \to \mathbb{R}^d$ are smooth functions such that $V(x)\cdot n(x) > 0$ for all $x \in \partial\Omega$ and let $B_0 u :=$ $V\cdot\nabla u + \alpha u$. Then there exists a Dirichlet form $\mathcal{E}$ as in Notation 51.1 and $\beta \in C^\infty(\partial\Omega \to (0,\infty))$ such that Eq. (51.2) holds with $Bu = \beta B_0 u$. In particular if $u \in H^2(\Omega)$, then $Bu = 0$ iff $B_0 u = 0$ on $\partial\Omega$.

Proof. Since mixed partial derivatives commute on $H^2(\Omega)$, the term $a_{ij}\partial_j\partial_i$ in Eq. (51.8) may be written as

$$\frac{1}{2}\left(a_{ij} + a_{ji}\right)\partial_j\partial_i.$$

With this in mind we must find coefficients $\{a_{\alpha,\beta} : |\alpha|, |\beta| \le 1\}$ as in Notation 51.1, such that

$$A_{ij} = \frac{1}{2}\left(a_{ij} + a_{ji}\right), \tag{51.11}$$

$$A_i = \left(a_{i0} - a_{0j} - \rho^{-1}\partial_j\left[\rho a_{ij}\right]\right), \tag{51.12}$$

$$A_0 = a_{00} - \rho^{-1}\partial_j\left[\rho a_{0j}\right], \tag{51.13}$$

$$a^{\mathrm{tr}} n = \beta V \text{ and} \tag{51.14}$$

$$n_i a_{0i} = \beta\alpha. \tag{51.15}$$

Eq. (51.11) will be satisfied if

$$a_{ij} = A_{ij} + c_{ij}$$

where $c_{ij} = -c_{ji}$ are any functions in $BC^\infty(\Omega)$. Dotting Eq. (51.14) with n shows that

$$\beta = \frac{a^{\mathrm{tr}} n\cdot n}{V\cdot n} = \frac{n\cdot an}{V\cdot n} = \frac{n\cdot An}{V\cdot n} \tag{51.16}$$

and Eq. (51.14) may now be written as

$$w := An - \frac{n\cdot An}{V\cdot n}V = cn \tag{51.17}$$

which means we have to choose $c = (c_{ij})$ so that Eq. (51.17) holds. This is easily done, since $w\cdot n = 0$ by construction we may define $c\xi := w(n\cdot\xi) - n(w\cdot\xi)$ for $\xi \in \mathbb{R}^d$. Then c is skew symmetric and $cn = w$ as desired. Since c_{ij} are smooth functions on $\partial\Omega$, a partition of unity argument shows that $c_{ij} = -c_{ji}$ may be extended to element of $C^\infty(\bar{\Omega})$. (These extensions are highly non-unique but it does not matter.) With these choices, Eq. (51.11) and Eq. (51.14) now hold with β as in Eq. (51.16). We now choose $a_{0i} \in C^\infty(\bar{\Omega})$ such that $a_{0i} = \beta\alpha n_i$ on $\partial\Omega$. Once these choices are made, it should be clear that Eqs. (51.13) and (51.14) may be solved uniquely for the functions a_{0j} and a_{00}. ■

51.2 Weak Solutions for Elliptic Operators

For the rest of this subsection we will assume $\rho = 1$. This can be done here by absorbing ρ into the coefficient $a_{\alpha\beta}$.

Definition 51.6. *The Dirichlet for $\mathcal{E}$ is* ***uniformly elliptic*** *on Ω if there exists $\epsilon > 0$ such that $(a_{ij}(x)) \geq \epsilon I$ for all $x \in \Omega$, i.e. $a_{ij}(x)\xi_i\xi_j \geq \epsilon |\xi|^2$ for all $x \in \Omega$ and $\xi \in \mathbb{R}^d$.*

Assumption 4 *For the remainder of this chapter, it will be assumed that $\mathcal{E}$ is uniformly elliptic on Ω.*

Lemma 51.7. *If $\xi^2 \leq A\xi + B$ then $\xi^2 \leq A^2 + 2B$.*

Proof. $\xi^2 \leq \frac{1}{2}A^2 + \frac{1}{2}\xi^2 + B$. Therefore $\frac{1}{2}\xi^2 \leq \frac{1}{2}A^2 + B$ or $\xi^2 \leq A^2 + 2B$. ■

Theorem 51.8. *Keeping the notation and assumptions of Proposition 51.2 along with Assumption 4, then*

$$\mathcal{E}(u,u) + C_\epsilon \|u\|_{L^2(\Omega)} \geq \frac{\epsilon}{2}\|u\|_{H^1(\Omega)}, \tag{51.18}$$

where $C_\epsilon = \frac{2C^2}{\epsilon} + C + \frac{\epsilon}{2}$.

Proof. To simplify notation in the proof, let $\|\cdot\|$ denote the $L^2(\Omega)$ – norm. Since

$$\int_\Omega a_{ij}\partial_i u \cdot \partial_j u \, dm \geq \epsilon \int_\Omega |\nabla u|^2 dx = \epsilon \|\nabla u\|_{L^2}^2,$$

$$\mathcal{E}(u,u) \geq \epsilon \|\nabla u\|_{L^2}^2 - C(\|\nabla u\| \, \|u\| + \|u\|^2)$$

and so

$$\|\nabla u\|^2 \leq \frac{C}{\epsilon}\|u\|\|\nabla u\| + \left(\frac{1}{\epsilon}\mathcal{E}(u,u) + \frac{C}{\epsilon}\|u\|^2\right).$$

Therefore by Lemma 51.7 with $A = \frac{C}{\epsilon}\|u\|$, $B = \left(\frac{1}{\epsilon}\mathcal{E}(u,u) + \frac{C}{\epsilon}\|u\|^2\right)$ and $\xi = \|\nabla u\|$,

$$\begin{aligned}\|\nabla u\|^2 &\leq \frac{C^2}{\epsilon^2}\|u\|^2 + \frac{2}{\epsilon}(\mathcal{E}(u,u) + \frac{C}{\epsilon}\|u\|^2) \\ &= \frac{2}{\epsilon}\mathcal{E}(u,u) + \left(\frac{C^2}{\epsilon^2} + \frac{2C}{\epsilon}\right)\|u\|^2.\end{aligned}$$

Hence

$$\frac{\epsilon}{2}\|\nabla u\|^2 \leq \mathcal{E}(u,u) + \left(\frac{2C^2}{\epsilon} + C\right)\|u\|^2$$

which, after adding $\frac{\epsilon}{2}\|u\|^2$ to both sides of this equation, gives Eq. (51.18). ■

The following theorem is an immediate consequence of Theorem 51.8 and the Lax-Milgram Theorem 53.9.

Corollary 51.9. *The quadratic form*

$$Q(u,v) := \mathcal{E}(u,v) + C_\epsilon(u,v)$$

satisfies the assumptions of the Lax Milgram Theorem 53.9 on $H^1(\Omega)$ or any closed subspace X of $H^1(\Omega)$.

Theorem 51.10 (Weak Solutions). *Let $\mathcal{E}$ be as in Notation 51.1 and C_ϵ be as in Theorem 51.8,*

$$Q(u,v) := \mathcal{E}(u,v) + C_\epsilon(u,v) \text{ for } u,v \in H^1(\Omega)$$

and X be a closed subspace of $H^1(\Omega)$. Then the maps $\mathcal{L} : X \to X^*$ and $\mathcal{L}^\dagger : X \to X^*$ defined by

$$\begin{aligned}\mathcal{L}v &:= Q(v,\cdot) = (\mathcal{L}_\mathcal{E} + C)\,v \text{ and} \\ \mathcal{L}^\dagger v &:= Q(\cdot,v) = \left(\mathcal{L}_\mathcal{E}^\dagger + C\right)v\end{aligned}$$

are linear isomorphisms of Hilbert spaces satisfying

$$\left\|\mathcal{L}^{-1}\right\|_{L(X^*,X)} \leq \frac{2}{\epsilon} \text{ and } \left\|(\mathcal{L}^\dagger)^{-1}\right\|_{L(X^*,X)} \leq \frac{2}{\epsilon}.$$

In particular for $f \in X^*$, there exist a unique solution $u \in X$ to $\mathcal{L}u = f$ and this solution satisfies the estimate

$$\|u\|_{H^1(\Omega)} \leq \frac{2}{\epsilon}\|f\|_{X^*}.$$

Remark 51.11. If $X \supset H_0^1(\Omega)$ and $u \in X$ then for $\phi \in C_c^\infty(\Omega) \subset X$,

$$\langle \mathcal{L}u, \phi\rangle = Q(u,\phi) = (u, \left(L^\dagger + C\right)\phi) = \langle (L+C)u, \phi\rangle.$$

That is to say $\mathcal{L}u|_{C_c^\infty(\Omega)} = (L+C)\,u$. In particular any solution $u \in X$ to $\mathcal{L}u = f \in X^*$ solves

$$(L+C)\,u = f|_{C_c^\infty(\Omega)} \in \mathcal{D}'(\Omega).$$

Remark 51.12. Suppose that $\Gamma \subset \partial\Omega$ is a measurable set such that $\sigma(\Gamma) > 0$ and $X_\Gamma := \left\{u \in H^1(\Omega) : 1_\Gamma u|_{\partial\Omega} = 0\right\}$. If $u \in H^2(\Omega)$ solves $\mathcal{L}u = f$ for some $f \in L^2(\Omega) \subset X^*$, then by Proposition 51.4,

$$(f,u) := \langle \mathcal{L}u, v\rangle = \mathcal{E}(u,v) + C(u,v) = ((L+C)\,u, v) + \int_{\partial\Omega} Bu \cdot v \, d\sigma \tag{51.19}$$

for all $v \in X_\Gamma \subset H^1(\Omega)$. Taking $v \in \mathcal{D}(\Omega) \subset X_\Gamma$ in Eq. (51.19) shows $(L+C)\,u = f$ a.e. and

$$\int_{\partial\Omega} Bu \cdot v \, d\sigma = 0 \text{ for all } v \in X_\Gamma.$$

Therefore we may conclude, u solves

$$\begin{aligned}(L+C)\,u &= f \text{ a.e. with} \\ Bu(x) &= 0 \text{ for } \sigma - \text{a.e. } x \in \partial\Omega \setminus \Gamma \text{ and} \\ u(x) &= 0 \text{ for } \sigma - \text{a.e. } x \in \Gamma.\end{aligned}$$

The following proposition records the important special case of Theorem 51.10 when $X = H_0^1(\Omega)$ and hence $X^* = H^{-1}(\Omega)$. The point to note here is that $\mathcal{L}u = (L + C)\,u$ when $X = H_0^1(\Omega)$, i.e. $\mathcal{L}u$ equals $[(L + C)\,u]$ extended by continuity to a linear functional on $X^* = \left[H_0^1(\Omega)\right]^*$.

Proposition 51.13. *Assume L is elliptic as above. Then there exist $C > 0$ sufficiently large such that $(L + C) : H_0^1(\Omega) \to H^{-1}(\Omega)$ is bijective with bounded inverse. Moreover*

$$\|(L + C)^{-1}\|_{L\left(H_0^1(\Omega),H^{-1}(\Omega)\right)} \leq 2/\epsilon$$

or equivalently

$$\|u\|_{H_0^1(\Omega)} \leq \frac{2}{\epsilon}\|(L + C)u\|_{H^{-1}(\Omega)} \text{ for all } u \in H_0^1(\Omega).$$

Our next goal, see Theorem 52.15, is to prove the elliptic regularity result, namely if $X = H_0^1(\Omega)$ or $X = H^1(\Omega)$ and $u \in X$ satisfies $\mathcal{L}u \in H^k(\Omega)$, then $u \in H^{k+2}(\Omega) \cap X$.

52

Elliptic Regularity

Assume that $\bar{\Omega}$ is a compact manifold with C^2 – boundary and satisfying $\bar{\Omega}^o = \Omega$ and let $\mathcal{E}$ be the Dirichlet form defined in Notation 51.1 and L be as in Eq. (51.6) or Eq. (51.8). We will assume $\mathcal{E}$ or equivalently that L is uniformly elliptic on Ω. This section is devoted to proving the following elliptic regularity theorem.

Theorem 52.1 (Elliptic Regularity Theorem). *Suppose $X = H_0^1(\Omega)$ or $H^1(\Omega)$ and $\mathcal{E}$ is as above. If $u \in X$ such that $\mathcal{L}_{\mathcal{E}} u \in H^k(\Omega)$ for some $k \in \mathbb{N}_0 \cup \{-1\}$, then $u \in H^{k+2}(\Omega)$ and*

$$\|u\|_{H^{k+2}}(\Omega) \le C(\|\mathcal{L}_{\mathcal{E}} u\|_{H^k(\Omega)} + \|u\|_{L^2(\Omega)}). \tag{52.1}$$

52.1 Interior Regularity

Theorem 52.2 (Elliptic Interior Regularity). *To each $\chi \in C_c^\infty(\Omega)$ there exist a constant $C = C(\chi)$ such that*

$$\|\chi u\|_{H^1(\Omega)} \le C\{\|Lu\|_{H^{-1}(\Omega)} + \|u\|_{L^2(\Omega)}\} \text{ for all } u \in H^1(\Omega). \tag{52.2}$$

In particular, if W is a precompact open subset of Ω, then

$$\|u\|_{H^1(W)} \le C\{\|Lu\|_{H^{-1}(\Omega)} + \|u\|_{L^2(\Omega)}\}. \tag{52.3}$$

Proof. For $u \in H^1(\Omega)$, $\chi u \in H_0^1(\Omega)$ and hence by Proposition 51.13, Proposition 50.6 and Lemma 50.7,

$$\begin{aligned}
\|\chi u\|_{H^1(\Omega)} &\le \frac{2}{\epsilon}\|(L + C_\epsilon)(\chi u)\|_{H^{-1}(\Omega)} \\
&= \frac{2}{\epsilon}\|\chi(L + C_\epsilon)u + [L, M_\chi]u\|_{H^{-1}(\Omega)} \\
&\le \frac{2}{\epsilon}C(\chi)\left\{\|(L + C_\epsilon)u\|_{H^{-1}(\Omega)} + \|u\|_{L^2(\Omega)}\right\} \\
&\le \frac{2}{\epsilon}C(\chi)\left\{\|Lu\|_{H^{-1}(\Omega)} + C_\epsilon\|u\|_{H^{-1}(\Omega)} + \|u\|_{L^2(\Omega)}\right\}
\end{aligned}$$

from which Eq. (52.2) follows. To prove Eq. (52.3), choose $\chi \in C_c^\infty(\Omega, [0,1])$ such that $\chi = 1$ on a neighborhood of $\bar{W}$ in which case

$$\|u\|_{H^1(W)} = \|\chi u\|_{H^1(W)} \le \|\chi u\|_{H^1(\Omega)} \le C\{\|Lu\|_{H^{-1}(\Omega)} + \|u\|_{L^2(\Omega)}\}.$$

■

Exercise 52.3. Let $v \in \mathbb{R}^d$ with $|v| = 1$, $u \in L^2(\Omega)$ and W be an open set such that $\bar{W} \sqsubset\sqsubset \Omega$. For all $h \neq 0$ sufficiently show

$$\left\|\partial_v^h u\right\|_{H^{-1}(W)} \le \|u\|_{L^2(\Omega)}.$$

Notice that $\partial_v^h u \in L^2(W) \subset H^{-1}(W)$.

Solution 52.4 (52.3). Let W_1 be a precompact open subset of Ω such that $\bar{W} \subset W_1 \subset \bar{W}_1 \subset \Omega$. Then for $\phi \in \mathcal{D}(W)$ and h close to zero,

$$\begin{aligned}
\left|\langle \partial_v^h u, \phi\rangle\right| &= \left|\langle u, \partial_v^{-h}\phi\rangle\right| \le \|u\|_{L^2(W_1)} \left\|\partial_v^{-h}\phi\right\|_{L^2(W_1)} \\
&\le \|u\|_{L^2(W_1)} \|\partial_v \phi\|_{L^2(\Omega)} \ \text{(Theorem 48.13)} \\
&\le \|u\|_{L^2(\Omega)} \|\phi\|_{H^1(\Omega)}.
\end{aligned}$$

Hence

$$\begin{aligned}
\left\|\partial_v^h u\right\|_{H^{-1}(W)} &= \sup\left\{\left|\langle \partial_v^h u, \phi\rangle\right| : \phi \in \mathcal{D}(W) \text{ with } \|\phi\|_{H^1(\Omega)} = 1\right\} \\
&\le \|u\|_{L^2(\Omega)}.
\end{aligned}$$

Theorem 52.5 (Interior Regularity). *Suppose L is 2^{nd} order uniformly elliptic operator on Ω and $u \in H^1(\Omega)$ satisfies $Lu \in H^k(\Omega)$[1] for some $k = -1, 0, 1, 2, \ldots$, then $u \in H^{k+2}_{loc}(\Omega)$. Moreover, if $W \subset\subset \Omega$ then there exists $C = C_k(W) < \infty$ such that*

$$\|u\|_{H^{k+2}(W)} \le C(\|Lu\|_{H^k(\Omega)} + \|u\|_{L^2(\Omega)}). \tag{52.4}$$

Proof. The proof is by induction on k with Theorem 52.2 being the case $k = -1$. Suppose that the interior regularity theorem holds for $-1 \le k \le k_0$. We will now complete the induction proof by showing it holds for $k = k_0 + 1$.

So suppose that $u \in H^1(\Omega)$ such that $Lu \in H^{k_0+1}(\Omega)$ and $W = W_0 \subset \Omega$ is fixed. Choose open sets W_1, W_2 and W_3 such that $\bar{W}_0 \subset W_1 \subset \bar{W}_1 \subset W_2 \subset \bar{W}_2 \subset W_3 \subset \bar{W}_3 \subset \Omega$ as in Figure 52.1. The idea now is to apply the induction hypothesis to the function $\partial_v^h u$ where $v \in \mathbb{R}^d$ and ∂_v^h is the finite difference operator in Definition 29.14. For the remainder of the proof $h \neq 0$ will be assumed to be sufficiently small so that the following computations make sense. To simplify notation let $D^h = \partial_v^h$.

[1] A priori, $Lu \in H^{-1}(\Omega) \subset \mathcal{D}'(\Omega)$).

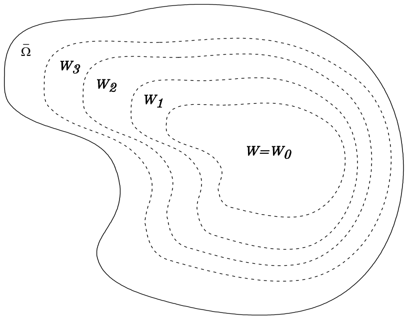


Fig. 52.1. The sets W_i for $i = 0, 1, 2$.

For h small, $D^h u \in H^1(W_3)$ and $D^h Lu \in H^{k_0+1}(W_3)$ and by Exercise 52.3 for $k_0 = -1$ and Theorem 48.13 for $k_0 \geq 0$,

$$\|D^h Lu\|_{H^{k_0}(W_1)} \leq \|Lu\|_{H^{k_0+1}(W_2)}. \tag{52.5}$$

We now compute $LD^h u$ as

$$LD^h u = D^h Lu + [L, D^h]u, \tag{52.6}$$

where

$$\begin{aligned}
[L, D^h]u &= LD^h u - D^h Lu \\
&= P(x, \partial) D^h u(x) - D^h P(x, \partial) u(x) \\
&= P(x, \partial) \left(\frac{u(x + hv) - u(x)}{h} \right) \\
&\quad - \frac{P(x + hv, \partial) u(x + hv) - P(x, \partial) u(x)}{h} \\
&= \frac{P(x, \partial) - P(x + hv, D)}{h} \, u(x + hv) = L^h \tau_v^h u(x),
\end{aligned}$$

$$\tau_v^h u(x) = u(x + hv)$$

and

$$L^h u := \sum_{|\alpha| \leq 2} \frac{A_\alpha(x) - A_\alpha(x + he)}{h} \, \partial^\alpha u.$$

The meaning of Eq. (52.6) and the above computations require a bit more explanation in the case $k_0 = -1$ in which case $Lu \in L^2(\Omega)$. What is being claimed is that

$$LD^h u = D^h Lu + L^h \tau_v^h u$$

as elements of $H^{-1}(W_3)$. By definition this means that

$$\begin{aligned}
-\langle u, D^{-h} L^\dagger \phi \rangle &= \langle LD^h u, \phi \rangle = \langle D^h Lu + L^h \tau_v^h u, \phi \rangle \\
&= -\langle u, L^\dagger D^{-h} \phi \rangle + \langle \tau_v^h u, \left(L^h\right)^\dagger \phi \rangle.
\end{aligned}$$

So the real identity which needs to be proved here is that $\left[D^{-h}, L^\dagger\right] \phi = -\tau_v^{-h} \left(L^h\right)^\dagger \phi$ for all $\phi \in \mathcal{D}(W_3)$. This can be done as above or it can be inferred (making use of the properties $L^\dagger$ is the formal adjoint of L and $-D^{-h}$ is the formal adjoint of D^h) from the computations already done in the previous paragraph with u being a smooth function.

Since L^h is a second order differential operator with coefficients which have bounded derivatives to all orders with bounds independent of h small, $[L, D^h]u = L^h \tau_v^h u \in H^{k_0}(W_1)$ and there is a constant $C < \infty$ such that

$$\begin{aligned}
\|[L, D^h]u\|_{H^{k_0}(W_1)} &= \|L^h \tau_v^h u\|_{H^{k_0}(W_1)} \\
&\leq C \|\tau_v^h u\|_{H^{k_0+2}(W_2)} \leq C \|u\|_{H^{k_0+2}(W_3)}.
\end{aligned} \tag{52.7}$$

Combining Eqs. (52.5 – 52.7) implies that $LD^h u \in H^{k_0}(W_2)$ and

$$\|LD^h u\|_{H^{k_0}(W_1)} \lesssim \|Lu\|_{H^{k_0+1}(W_2)} + \|u\|_{H^{k_0+2}(W_3)}.$$

Therefore by the induction hypothesis, $D^h u \in H^{k_0+2}(W_0)$ and

$$\begin{aligned}
\left\|D^h u\right\|_{H^{k_0+2}(W_0)} &\lesssim \|LD^h u\|_{H^{k_0}(W_1)} + \|D^h u\|_{L^2(W_1)} \\
&\lesssim \|Lu\|_{H^{k_0+1}(W_2)} + \|u\|_{H^{k_0+2}(W_3)} + \|u\|_{H^1(W_2)} \\
&\lesssim \|Lu\|_{H^{k_0+1}(W_2)} + \|u\|_{H^{k_0+2}(W_3)} \\
&\lesssim \|Lu\|_{H^{k_0+1}(\Omega)} + \|Lu\|_{H^{k_0}(\Omega)} + \|u\|_{L^2(\Omega)} \text{ (by induction)} \\
&\lesssim \|Lu\|_{H^{k_0+1}(\Omega)} + \|u\|_{L^2(\Omega)}.
\end{aligned}$$

So by Theorem 48.13, $\partial_v u \in H^{k_0+2}(W_0)$ for all $v = e_i$ with $i = 1, 2, \ldots, d$ and

$$\|\partial_i u\|_{H^{k_0+2}(W)} = \|\partial_i u\|_{H^{k_0+2}(W_0)} \lesssim \|Lu\|_{H^{k_0+1}(\Omega)} + \|u\|_{L^2(\Omega)}.$$

Thus $u \in H^{k_0+3}(W_0)$ and Eq. (52.4) holds. ■

Corollary 52.6. *Suppose L is as above and $u \in H^1(\Omega)$ such that $Lu \in BC^\infty(\Omega)$ then $u \in C^\infty(\Omega)$.*

Proof. Choose $\Omega_0 \subset\subset \Omega$ so $Lu \in BC^\infty(\overline{\Omega}_0)$. Therefore $Lu \in H^k(\Omega_0)$ for all $k = 0, 1, 2, \ldots$. Hence $u \in H^{k+2}_{loc}(\Omega_0)$ for all $k = 0, 1, 2, \ldots$. Then by Sobolev embedding Theorem 49.18, $u \in C^\infty(\Omega_0)$. Since Ω_0 is an arbitrary precompact open subset of Ω, $u \in C^\infty(\Omega)$. ■

52.2 Boundary Regularity Theorem

Example 52.7. Let $\Omega = D(0,1)$ and $u(z) = (1+z)\log(1+z)$. Since $u(z)$ is holomorphic on Ω it is also harmonic, i.e. $\Delta u = 0 \in H^k(\Omega)$ for all k. However we will now show that while $u \in H^1(\Omega)$ it is not in $H^2(\Omega)$. Because u is holomorphic,

$$u_x = \frac{\partial u}{\partial z} = 1 + \log(1+z) \text{ and } u_y = i\frac{\partial u}{\partial z} = i + i\log(1+z)$$

from which it is easily shown $u \in H^1(\Omega)$. On the other hand,

$$u_{xx} = \frac{\partial^2}{\partial z^2}u = \frac{1}{1+z}$$

and

$$\int_\Omega \left|\frac{1}{1+z}\right|^2 dx\,dy = \int_{\Omega+1} \left|\frac{1}{z}\right|^2 dx\,dy$$
$$\geq \int_C \left|\frac{1}{z}\right|^2 dx\,dy = \frac{\pi}{2}\int_0^1 \frac{1}{r^2} r dr = \infty,$$

where C is the cone in Figure 52.2. This shows $u \notin H^2(\Omega)$ and the problems come from the bad behavior of u near $-1 \in \partial\Omega$.

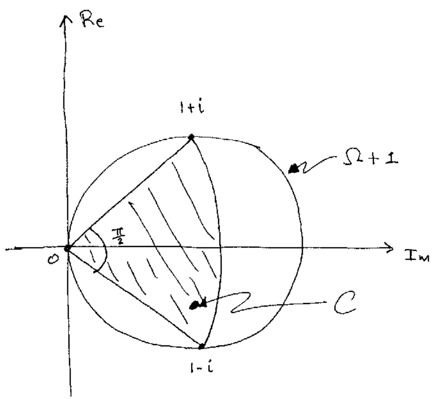


Fig. 52.2. The cone used in showing u not in $H^2(\Omega)$.

This example shows that in order to get an elliptic regularity result which is valid all the way up to the boundary, it is necessary to impose some sort of

boundary conditions on the solution which will rule out the bad behavior of the example. Since the Dirichlet form contains boundary information, we will do this by working with $\mathcal{E}$ rather than the operator L on $\mathcal{D}'(\Omega)$ associated to $\mathcal{E}$. Having to work with the quadratic form makes life a bit more difficult.

Notations 52.8 *Let*

1. $N_r := \{x \in \mathbb{H}^n : |x| < r\}$.
2. $X = H_0^1(N_r)$ *or* X *be the closed subspace* $H^1(N_r)$ *given by*
$$X = \left\{u \in H^1(N_r) : u|_{\partial\mathbb{H}^n \cap \bar{N}_r} = 0\right\}. \tag{52.8}$$
3. *For* $s \leq r$ *let* $X_s = \{u \in X : u = 0 \text{ on } \mathbb{H}^n \setminus N_\rho \text{ for some } \rho < s\}$.

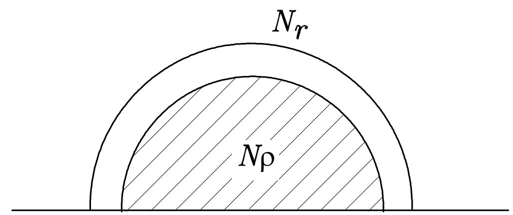


Fig. 52.3. Nested half balls.

Remark 52.9. 1. If $\phi \in C^\infty(\overline{\mathbb{H}}^n)$ and vanishes on $\mathbb{H}^n \setminus N_\rho$ for some $\rho < r$ then $\phi u \in X_r$ for all $u \in X$.
2. If $u \in X_r$ then $\partial_h^\alpha u \in X_r$ for all α such that $\alpha_d = 0$ and $|h|$ sufficiently small.

Lemma 52.10 (Commutator). *If* $\psi \in C^\infty(\overline{N}_r)$ *then for* $\gamma \in \mathbb{N}_0^{d-1} \times \{0\}$ *there exists* $C_\gamma(\psi) < \infty$ *such that*

$$\|[\psi, \partial_h^\gamma]f\|_{L^2(N_\rho)} \leq C_\gamma(\psi) \sum_{\alpha<\gamma} \|\partial^\alpha f\|_{L^2(N_r)}. \tag{52.9}$$

for all $f \in L^2(N_r)$ *with* $\partial^\alpha f \in L^2(N_r)$ *for* $\alpha < \gamma$.

Proof. The proof will be by induction on $|\gamma|$. If $\gamma = e_i$ for some $i < d$, then

$$\begin{aligned}\partial_h^i(\psi f)(x) &:= \frac{\psi(x+he_i)f(x+he_i) - \psi(x)f(x)}{h}\\ &= \frac{[\psi(x+he_i) - \psi(x)]\,f(x+he_i) + \psi(x)\,[f(x+he_i) - f(x)]}{h}\\ &= \partial_h^i\psi(x)f(x+he_i) + \psi(x)\partial_h^i f(x)\end{aligned}$$

which gives

$$[\partial_h^i,\psi]f=(\partial_h^i\psi)\tau_h^i f. \tag{52.10}$$

This then implies that

$$\|[\psi,\partial_h^i]f\|_{L^2(N_\rho)}\le C(\psi)\|f\|_{L^2(N_r)}.$$

Now suppose $|\gamma|>1$ with $\gamma=e_i+\gamma'$ so that $\partial_h^\gamma=\partial_h^{\gamma'}\partial_h^i$ with$|\gamma'|=|\gamma|-1.$ Then

$$[\psi,\partial_h^\gamma]=[\psi,\partial_h^{\gamma'}]\partial_h^i+\partial_h^{\gamma'}[\psi,\partial_h^i]$$

and therefore by the induction hypothesis and Theorems 48.13 and 48.15,

$$\begin{aligned}\|[\psi,\partial_h^\gamma]f\|_{L^2}&\le C_{\gamma'}(\psi)\sum_{\alpha<\gamma'}\|\partial^\alpha\partial_h^i f\|_{L^2}+\|\partial^{\gamma'}[\psi,\partial_h^i]f\|_{L^2}\\&\le C_{\gamma'}(\psi)\sum_{\alpha<\gamma'}\|\partial^{\alpha+e_i}f\|_{L^2}+\|\partial^{\gamma'}\left[(\partial_h^i\psi)\tau_h^i f\right]\|_{L^2}.\end{aligned} \tag{52.11}$$

But

$$\partial^{\gamma'}\left[(\partial_h^i\psi)\tau_h^i f\right]=\sum_{\beta_1+\beta_2=\gamma'}\frac{\gamma'!}{\beta_1!\beta_2!}(\partial_h^i\partial^{\beta_1}\psi)\tau_h^i\partial^{\beta_2}f$$

and hence

$$\|\partial^{\gamma'}\left[(\partial_h^i\psi)\tau_h^i f\right]\|\le C\sum_{\beta\le\gamma'}\|\partial^\beta f\|_{L^2}. \tag{52.12}$$

Combining Eqs. (52.11) and (52.12) gives the desired result,

$$\|[\psi,\partial_h^\gamma]f\|_{L^2}\le C_\gamma(\psi)\sum_{\alpha<\gamma}\|\partial^\alpha f\|_{L^2}.$$

■

Lemma 52.11 (Warmup for Proposition 52.12). *Let $a_{\alpha\beta}\in BC^\infty\left(\mathbb{H}^d\right)$ with $(a_{ij})\ge\epsilon\delta_{ij}$ for some $\epsilon>0$,*

$$\langle\mathcal{L}u,v\rangle=\mathcal{E}(u,v)=\int_{\mathbb{H}^d}\sum_{|\alpha|,|\beta|\le1}a_{\alpha\beta}\partial^\alpha u\cdot\partial^\beta v\,dx, \tag{52.13}$$

$X=H_0^1(\mathbb{H}^d)$ or $H^1(\mathbb{H}^d)$. There exists $C<\infty$ such that if $u\in X$ such that $\mathcal{L}u=:f\in L^2(\mathbb{H}^d)$, then $u\in H^2\left(\mathbb{R}^d\right)$ and

$$\|u\|_{H^2(\mathbb{H}^d)}\le C(\|f\|_{L^2(\mathbb{H}^d)}+\|u\|_{X^*}). \tag{52.14}$$

Proof. If $\mathcal{L}u=f\in X^*$ then $(\mathcal{L}+C)u=f+Cu$, so by the Lax-Milgram method,

$$\|u\|_X\lesssim\|f+Cu\|_{X^*}\le\|f\|_{X^*}+C\|u\|_{X^*}\lesssim\|\mathcal{L}u\|_{X^*}+\|u\|_{X^*}.$$

We wish to prove $\partial_i u\in H^1(\mathbb{H}^d)$ for all $i<d$ and

$$\|\partial_i u\|_{H^1(\mathbb{H}^d)}\lesssim\|\mathcal{L}u\|_{L^2(\mathbb{H}^d)}+\|u\|_{X^*}.$$

To do this consider

$$\begin{aligned}\langle\mathcal{L}\partial_i^h u,v\rangle&=\int_{\mathbb{H}^d}\sum_{|\alpha|,|\beta|\le1}a_{\alpha\beta}\partial_i^h\partial^\alpha u\cdot\partial^\beta v\,dx\\&=\int_{\mathbb{H}^d}\sum_{|\alpha|,|\beta|\le1}\left\{\partial_i^h\left(a_{\alpha\beta}\partial^\alpha u\right)+\left[a_{\alpha\beta},\partial_i^h\right]\partial^\alpha u\right\}\cdot\partial^\beta v\,dx\\&=:-\langle\mathcal{L}u,\partial_i^{-h}v\rangle-\int_{\mathbb{H}^d}\sum_{|\alpha|,|\beta|\le1}(\partial_h^i a_{\alpha\beta})\tau_h^i\partial^\alpha u\cdot\partial^\beta v\,dx\\&=-\langle\mathcal{L}u,\partial_i^{-h}v\rangle-\mathcal{E}_{\partial_h^i a}(\tau_h^i u,v=-(f,\partial_i^{-h}v)-\mathcal{E}_{\partial_h^i a}(\tau_h^i u,v)\\&=(\partial_i^h f,v)-\mathcal{E}_{\partial_h^i a}(\tau_h^i u,v)\end{aligned}$$

wherein we have made use of Eq. (52.10) in the third equality. From this it follows that

$$\mathcal{L}\partial_i^h u=\partial_i^h\mathcal{L}u-\mathcal{E}_{\partial_h^i a}(\tau_h^i u,\cdot)\in X^*$$

and

$$\begin{aligned}\left\|\mathcal{L}\partial_i^h u\right\|_{X^*}&\le\left\|\partial_i^h\mathcal{L}u\right\|_{X^*}+\left\|\mathcal{E}_{\partial_h^i a}(\tau_h^i u,\cdot)\right\|_{X^*}\lesssim\|\mathcal{L}u\|_{L^2}+\|u\|_X\\&\lesssim\|\mathcal{L}u\|_{L^2}+\|\mathcal{L}u\|_{X^*}+\|u\|_{X^*}\lesssim\|\mathcal{L}u\|_{L^2}+\|u\|_{X^*}.\end{aligned}$$

Therefore,

$$\begin{aligned}\left\|\partial_i^h u\right\|_X&\lesssim\left\|\mathcal{L}\partial_i^h u\right\|_{X^*}+\left\|\partial_i^h u\right\|_{X^*}\lesssim\|\mathcal{L}u\|_{L^2}+\|u\|_{X^*}+\|u\|_{L^2}\\&\lesssim\|\mathcal{L}u\|_{L^2}+\|u\|_{X^*}.\end{aligned}$$

Since h is small but arbitrary we conclude that $\partial_i u\in X$ and

$$\|\partial_i u\|_X\lesssim\|\mathcal{L}u\|_{L^2}+\|u\|_{X^*}\text{ for all }i<d.$$

Finally if $i=d$, we have that $f=\mathcal{L}u=\sum_{\alpha\ne2e_d}A_\alpha\partial^\alpha u+\partial_d^2u$ which implies (writing $A_{d,d}$ for A_{2e_d})

$$\partial_d^2u=A_{d,d}^{-1}\left(f-\sum_{\alpha\ne2e_d}A_\alpha\partial^\alpha u\right)\in L^2$$

because we have shown that $\partial_i\partial_j u\in L^2$ if $\{i,j\}\ne\{d\}$. Moreover we have the estimate that

$$\left\|\partial_d^2 u\right\|_{L^2} \lesssim \left\|f - \sum_{\alpha\neq 2e_d} A_\alpha \partial^\alpha u\right\|_{L^2} \lesssim \|f\|_{L^2} + \sum_{\alpha\neq 2e_d} \|\partial^\alpha u\|_{L^2}$$
$$\lesssim \|f\|_{L^2} + \sum_{j<d} \left\|\partial^j u\right\|_{X^*} \lesssim \|\mathcal{L}u\|_{L^2(\mathbb{H}^d)} + \|u\|_{X^*}.$$

Thus we have shown that $u \in X \cap H^2(\mathbb{H}^d)$ and

$$\|u\|_{H^2(\mathbb{H}^d)} \lesssim \|\mathcal{L}u\|_{L^2(\mathbb{H}^d)} + \|u\|_{X^*}.$$

■

If we try to use the above proof inductively to get higher regularity we run into a snag. To see this suppose now that $f \in H^1$. Then as above

$$\mathcal{L}\partial_j^h u = \partial_j^h \mathcal{L}u - \mathcal{E}_{\partial_h^j a}(\tau_h^j u, \cdot) = \partial_j^h f - \mathcal{E}_{\partial_h^j a}(\tau_h^j u, \cdot).$$

Let $\partial_h^j a = b$ and $\tau_h^j u = w$ and consider

$$\mathcal{E}_b(w,v) = \int_{\mathbb{H}^d} b_{\alpha,\beta}\partial^\alpha w \cdot \partial^\beta v dm.$$

Since $w \in H^2$ we may integrate by parts to find

$$\mathcal{E}_b(w,v) = \int_{\mathbb{H}^d} (-1)^{|\beta|}\, \partial^\beta \left(b_{\alpha,\beta}\partial^\alpha w\right) \cdot v dm - \int_{\partial\mathbb{H}^d} b_{\alpha,d}\partial^\alpha w \cdot v d\sigma.$$

This shows that $\mathcal{E}_b(w,\cdot)$ is representable by $(-1)^{|\beta|}\, \partial^\beta \left(b_{\alpha,\beta}\partial^\alpha w\right) \in L^2$ plus the boundary term

$$v \to \int_{\partial\mathbb{H}^d} b_{\alpha,d}\partial^\alpha w \cdot v d\sigma.$$

To continue on by this method, we would have to show that the boundary term is representable by an element of L^2. This should be the case since $v|_{\partial\mathbb{H}^d} \in H^{-1/2}\left(\mathbb{H}^d\right)$ while $\partial^\alpha w \in H^{1/2}\left(\mathbb{H}^d\right)$ with bounds. However we have not proven such statements so we will proceed by a different but closely related approach.

Proposition 52.12 (Local Tangential Boundary Regularity). *Let* $a_{\alpha,\beta} \in C^\infty\left(\bar{N}_t\right)$ *with* $a_{ij}\,\xi_i\,\xi_j \geq 2\epsilon|\xi|^2$,

$$Q(u,v) = \int_{N_t} \sum_{|\alpha|,|\beta|\leq 1} a_{\alpha\beta}\partial^\alpha u \cdot \partial^\beta v \; dx, \tag{52.15}$$

$X = H^1(N_t)$ or X be the closed subspace of $H^1(N_t)$ defined in Eq. (52.8) of Notation 52.8. Suppose $k \in \mathbb{N}_0$, $u \in X$ and $f \in H^k(N_t)$ satisfy,

$$Q(u,v) = \int_{N_t} fv \; dx \text{ for all } v \in X_t. \tag{52.16}$$

Given $\rho < t$, there exists $C < \infty$ such that for all $\gamma \in \mathbb{N}_0^{d-1} \times \{0\}$ with $|\gamma| \leq k+1$, $\partial^\gamma u \in H^1(N_\rho)$ and

$$\|\partial^\gamma u\|_{H^1(N_\rho)} \leq C(\|f\|_{H^k(N_t)} + \|u\|_{H^1(N_t)}) \tag{52.17}$$

Proof. Let $\rho < r < s < t$ and consider the half nested balls as in Figure 52.4 below. The proof will be by induction on $j = |\gamma|$. When $j = 0$ the

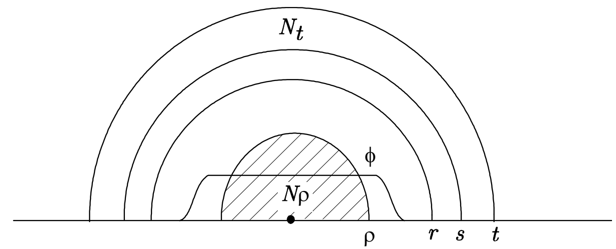


Fig. 52.4. A collection of nested half balls along with the cutoff function ϕ.

assertion is trivial. Assume now there exists $j \in [1, k+1] \cap \mathbb{N}$ such that $\partial^\gamma u \in H^1(N_s)$ for all $|\gamma| < j$ with $\gamma_d = 0$ and

$$\|\partial^\gamma u\|_{H^1(N_s)} \leq C(\|f\|_{H^k(N_t)} + \|u\|_{H^1(N_t)}).$$

Fix $\phi \in C_c^\infty(\overline{N}_t)$ such that $\phi = 0$ on $\bar{N}_t \setminus \bar{N}_r$ and $\phi = 1$ in a neighborhood of $\overline{N}_\rho$. Suppose γ is a multi-index such that $|\gamma| = j$ and $\gamma_d = 0$. Then $\partial_h^\gamma(\phi u) \in X_r$ for h sufficiently small.

With out loss of generality we may assume $\gamma_1 > 0$ and write $\gamma = e_1 + \gamma'$ and $\partial_h^\gamma = \partial_h^1 \partial_h^{\gamma'}$. For $v \in X_r$,

$$\begin{aligned}
Q(\partial_h^\gamma(\phi u),v) &= \int_{N_t}\sum_{|\alpha|,|\beta|\le 1} a_{\alpha\beta}\partial^\alpha\partial_h^\gamma(\phi u)\cdot\partial^\beta v = \int_{N_t}\sum_{|\alpha|,|\beta|\le 1} a_{\alpha\beta}\partial_h^\gamma\partial^\alpha(\phi u)\cdot\partial^\beta v \\
&= \sum_{|\alpha|,|\beta|\le 1}\int_{N_t}\partial_h^\gamma(a_{\alpha\beta}\partial^\alpha(\phi u))\cdot\partial^\beta v \\
&\quad + \overbrace{\sum_{|\alpha|,|\beta|\le 1}\int_{N_t}[a_{\alpha\beta},\partial_h^\gamma]\partial^\alpha(\phi u)\cdot\partial^\beta v}^{E_1} \\
&= \sum_{|\alpha|,|\beta|\le 1}\int_{N_t}\partial_h^\gamma(a_{\alpha\beta}\phi\partial^\alpha u)\cdot\partial^\beta v \\
&\quad + \overbrace{\sum_{|\alpha|,|\beta|\le 1}\int_{N_t}\partial_h^\gamma(a_{\alpha\beta}[\partial^\alpha,\phi]u)\cdot\partial^\beta v}^{E_2} + E_1 \\
&= \sum_{|\alpha|,|\beta|\le 1}(-1)^{|\gamma|}\int_{N_t} a_{\alpha\beta}\partial^\alpha u\cdot\phi\partial^\beta\partial_{-h}^\gamma v + E_1 + E_2 \\
&= \sum_{|\alpha|,|\beta|\le 1}(-1)^{|\gamma|}\int_{N_t} a_{\alpha\beta}\partial^\alpha u\cdot\partial^\beta\left[\phi\partial_{-h}^\gamma v\right] + E_1 + E_2 \\
&\qquad + \overbrace{\sum_{|\alpha|,|\beta|\le 1}(-1)^{|\gamma|}\int_{N_t} a_{\alpha\beta}\partial^\alpha u\cdot\left[\phi,\partial^\beta\right]\partial_{-h}^\gamma v}^{E_3} \\
&= (-1)^{|\gamma|}Q(u,\phi\partial_{-h}^\gamma v) + E_1 + E_2 + E_3 \\
&= (-1)^{|\gamma|}\int_{N_t}\phi f\partial_{-h}^\gamma v + E_1 + E_2 + E_3 \\
&= E_1 + E_2 + E_3 \overbrace{- \int_{N_t}\partial_h^{\gamma'}[\phi f]\cdot\partial_{-h}^1 v}^{E_4}. \\
&= E_1 + E_2 + E_3 + E_4.
\end{aligned}$$

To summarize,

$$Q(\partial_h^\gamma(\phi u),v) = E_1 + E_2 + E_3 + E_4$$

where

$$\begin{aligned}
E_1 &:= \sum_{|\alpha|,|\beta|\le 1}\int_{N_t}[a_{\alpha\beta},\partial_h^\gamma]\partial^\alpha(\phi u)\cdot\partial^\beta v \\
E_2 &:= \sum_{|\alpha|,|\beta|\le 1}\int_{N_t}\partial_h^\gamma(a_{\alpha\beta}[\partial^\alpha,\phi]u)\cdot\partial^\beta v \\
E_3 &:= \sum_{|\alpha|,|\beta|\le 1}(-1)^{|\gamma|}\int_{N_t} a_{\alpha\beta}\partial^\alpha u\cdot\left[\phi,\partial^\beta\right]\partial_{-h}^\gamma v \text{ and} \\
E_4 &:= -\int_{N_t}\partial_h^{\gamma'}[\phi f]\cdot\partial_{-h}^1 v.
\end{aligned}$$

To finish the proof we will estimate each of the terms E_i for $i = 1,\dots,4$. Using Lemma 52.10,

$$\begin{aligned}
|E_1| &\le \sum_{|\alpha|,|\beta|\le 1}\int_{N_t}\left|[a_{\alpha\beta},\partial_h^\gamma]\partial^\alpha(\phi u)\cdot\partial^\beta v\right| \\
&\le \|v\|_{H^1(N_r)}\sum_{|\alpha|,|\beta|\le 1}\|[a_{\alpha\beta},\partial_h^\gamma]\partial^\alpha(\phi u)\|_{L^2(N_r)} \\
&\le \|v\|_{H^1(N_r)}\sum_{|\alpha|,|\beta|\le 1}\sum_{\delta<\gamma} C_\gamma(a_{\alpha\beta})\|\partial^\delta\partial^\alpha(\phi u)\|_{L^2(N_r)} \\
&\lesssim \|v\|_{H^1(N_r)}\sum_{\delta<\gamma}\|\partial^\delta u\|_{H^1(N_r)} \\
&\lesssim \|v\|_{H^1(N_s)}\left(\|f\|_{H^k(N_s)} + \|u\|_{H^1(N_s)}\right) \text{ (by induction).}
\end{aligned}$$

For E_2,

$$\begin{aligned}
|E_2| &= \left|\sum_{|\beta|\le 1,|\alpha|=1}\int_{N_t}\partial_h^\gamma(a_{\alpha\beta}(\partial^\alpha\phi)u)\cdot\partial^\beta v\right| \\
&\le \|v\|_{H^1(N_r)}\sum_{|\beta|\le 1,|\alpha|=1}\|\partial_h^\gamma[a_{\alpha\beta}(\partial^\alpha\phi)u]\|_{L^2(N_r)} \\
&\le C\|v\|_{H^1(N_r)}\sum_{|\beta|\le 1,|\alpha|=1}\|\partial^\gamma[a_{\alpha\beta}(\partial^\alpha\phi)u]\|_{L^2(N_r)} \\
&\le C\|v\|_{H^1(N_r)}\sum_{\delta\le\gamma}\|\partial^\delta u\|_{L^2(N_r)} \\
&\le C\|v\|_{H^1(N_r)}\sum_{|\delta|\le j-1,\delta_n=0}\|\partial^\delta u\|_{H^1(N_r)} \\
&\le C\|v\|_{H^1(N_s)}(\|f\|_{H^k(N_s)} + \|u\|_{H^1(N_s)}) \text{ (by induction).}
\end{aligned}$$

For E_3,

$$\begin{aligned}|E_3| &\le \sum_{|\alpha|\le 1, |\beta|=1} \left| \int_{N_t} a_{\alpha\beta}\partial^\alpha u \cdot (\partial^\beta \phi)\, \partial^\gamma_{-h} v \right| \\ &= \sum_{|\alpha|\le 1, |\beta|=1} \left| \int_{N_t} \partial_h^{\gamma'} \left[a_{\alpha\beta}\partial^\alpha u \cdot (\partial^\beta \phi)\right] \cdot \partial^i_{-h} v \right| \\ &\le \sum_{|\alpha|\le 1, |\beta|=1} \|v\|_{H^1(N_r)} \|\partial^{\gamma'} \left[a_{\alpha\beta}\partial^\alpha u \cdot (\partial^\beta \phi)\right] \|_{L^2(N_r)} \\ &\le C\|v\|_{H^1(N_r)} \sum_{|\alpha|\le 1} \sum_{\delta \le \gamma'} \|\partial^{\delta+\alpha} u\|_{L^2(N_r)} \\ &\le C\|v\|_{H^1(N_r)} \sum_{\delta\le\gamma'} \|\partial^\delta u\|_{H^1(N_r)} \\ &\le C\|v\|_{H^1(N_s)} (\|f\|_{H^k(N_s)} + \|u\|_{H^1(N_s)}) \text{ (by the induction hypothesis).}\end{aligned}$$

Finally for E_4,

$$\begin{aligned}|E_4| &= \left| \int_{N_t} \partial_h^{\gamma'} [\phi f] \cdot \partial^1_{-h} v \right| \le \|\partial^1_{-h} v\|_{L^2(N_r)}\ \|\partial_h^{\gamma'}(\phi f)\|_{L^2(N_r)} \\ &\le \|v\|_{H^1(N_r)}\ \|\partial^{\gamma'}(\phi f)\|_{L^2(N_r)} \\ &\le \|v\|_{H^1(N_s)}\ \|\phi f\|_{H^{j-1}(N_r)} \le C\|v\|_{H^1(N_s)} \|f\|_{H^k(N_s)}.\end{aligned}$$

Putting all of these estimates together proves, whenever $|\gamma| = j$,

$$|Q(\partial_h^\gamma(\phi u), v)| \le C\|v\|_{H^1(N_s)} (\|f\|_{H^k(N_s)} + \|u\|_{H^1(N_s)}) \tag{52.18}$$

for all $v \in X_s$. In particular we may take $v = \partial_h^\gamma(\phi u) \in X_s$ in the above inequality to learn

$$Q\left(\partial_h^\gamma(\phi u), \partial_h^\gamma(\phi u)\right) \le C\|\partial_h^\gamma(\phi u)\|_{H^1(N_s)} (\|f\|_{H^k(N_s)} + \|u\|_{H^1(N_s)}). \tag{52.19}$$

But by coercivity of Q,

$$\begin{aligned}\|\partial_h^\gamma(\phi u)\|^2_{H^1(N_s)} &\le C\left[Q(\partial_h^\gamma(\phi u), \partial_h^\gamma(\phi u)) + \|\partial_h^\gamma(\phi u)\|^2_{L^2(N_s)}\right] \\ &\lesssim \|\partial_h^\gamma(\phi u)\|_{H^1(N_s)} (\|f\|_{H^k(N_s)} + \|u\|_{H^1(N_s)}) \\ &\qquad + \|\partial_h^\gamma(\phi u)\|_{H^1(N_s)} \|\partial_h^\gamma(\phi u)\|_{L^2(N_s)} \\ &\lesssim \|\partial_h^\gamma(\phi u)\|_{H^1(N_s)} \begin{pmatrix} \|f\|_{H^k(N_s)} + \|u\|_{H^1(N_s)} \\ + \|\partial_h^\gamma(\phi u)\|_{L^2(N_s)} \end{pmatrix}\end{aligned} \tag{52.20}$$

and hence

$$\|\partial_h^\gamma(\phi u)\|_{H^1(N_s)} \lesssim \|f\|_{H^k(N_s)} + \|u\|_{H^1(N_s)} + \|\partial_h^\gamma(\phi u)\|_{L^2(N_s)}. \tag{52.21}$$

Now

$$\begin{aligned}\|\partial_h^\gamma(\phi u)\|_{L^2(N_s)} &= \|\partial_h^i \partial_h^{\gamma'}(\phi u)\|_{L^2(N_s)} \\ &\le \|\partial_h^{\gamma'}(\phi u)\|_{H^1(N_s)} \le \|\partial^{\gamma'}(\phi u)\|_{H^1(N_s)} \\ &\le C \sum_{\alpha \le \gamma'} \|\partial^\alpha u\|_{H^1(N_s)} \text{ by the chain rule} \\ &\le C(\|f\|_{H^k(N_t)} + \|u\|_{H^1(N_t)}) \text{ by induction.}\end{aligned}$$

This last estimated combined with Eq. (52.21) shows

$$\|\partial_h^\gamma(\phi u)\|_{H^1(N_s)} \lesssim \|f\|_{H^k(N_t)} + \|u\|_{H^1(N_t)}$$

and therefore $\partial^\gamma(\phi u) \in H^1(N_s)$ and

$$\|\partial^\gamma(\phi u)\|_{H^1(N_s)} \lesssim \|f\|_{H^k(N_t)} + \|u\|_{H^1(N_t)}.$$

This proves the proposition since $\phi \equiv 1$ on N_ρ so

$$\begin{aligned}\|\partial^\gamma u\|_{H^1(N_\rho)} &= \|\partial^\gamma(\phi u)\|_{H^1(N_\rho)} \le \|\partial^\gamma(\phi u)\|_{H^1(N_s)} \\ &\lesssim \|f\|_{H^k(N_t)} + \|u\|_{H^1(N_t)}.\end{aligned}$$

■

Theorem 52.13 (Local Boundary Regularly). *As in Proposition 52.12, let $a_{\alpha,\beta} \in C^\infty(\bar{N}_t)$ with $a_{ij}\,\xi_i\,\xi_j \ge 2\epsilon|\xi|^2$,*

$$Q(u,v) = \sum_{|\alpha|,|\beta|\le 1} \int_{N_t} a_{\alpha\beta}\partial^\alpha u \cdot \partial^\beta v\ dx$$

and $X = H^1_0(N_t)$ or $X \subset H^1(N_t)$ as in Eq. (52.8) . If $f \in H^k(N_t)$ for some $k \ge 0$ and $u \in X$ solves Q

$$Q(u,v) = (f,v) \text{ for all } v \in X_t$$

then for all $\rho < t$, $u \in H^{k+2}(N_\rho)$ and there exists $C < \infty$ such that

$$\|u\|_{H^{k+2}(N_\rho)} \le C(\|f\|_{H^k(N_t)} + \|u\|_{H^1(N_t)}).$$

Proof. The theorem will be proved by showing $\partial^\gamma u \in L^2(N_\rho)$ for all $|\gamma| \le k+2$ and

$$\|\partial^\gamma u\|_{L^2(N_\rho)} \lesssim \|f\|_{H^k(N_t)} + \|u\|_{H^1(N_t)}. \tag{52.22}$$

The proof of Eq. (52.22) will be by induction on $j = \gamma_d$. The case $j = 0, 1$ follows from Proposition 52.12. Suppose $j = \gamma_d \ge 2$ and $\gamma' = \gamma - 2e_d$ so $\partial^\gamma = \partial^{\gamma'}\partial_d^2$. Now letting

$$L = \sum_{|\alpha|,|\beta|\le 1} (-1)^{|\beta|}\partial^\beta a_{\alpha\beta}\partial_\alpha = \sum_{|\alpha|\le 2} A_\alpha \partial^\alpha,$$

then $Lu = f$ in the distributional sense. Writing $\tilde{A}$ for $A_{(0,0,\dots,0,2)}$,

$$f = \tilde{A}\partial_d^2 u + \sum_{|\alpha|\le 2, \alpha_d<2} A_\alpha \partial^\alpha u$$

so that

$$\partial_d^2 u = \frac{1}{\tilde{A}}(f - \sum_{|\alpha|\le 2, \alpha_d<2} A_\alpha \partial^\alpha u)$$

and

$$\partial^\gamma u = \partial^{\gamma'}\partial_d^2 u = \partial^{\gamma'}\left(\frac{1}{\tilde{A}}f - \sum_{|\alpha|\le 2, \alpha_d<2} \frac{A_\alpha}{\tilde{A}}\partial^\alpha u\right). \tag{52.23}$$

Now by the product rule

$$\sum_{|\alpha|\le 2, \alpha_d<2} \partial^{\gamma'}\left(\frac{A_\alpha}{\tilde{A}}\partial^\alpha u\right) := \sum_{|\alpha|\le 2, \alpha_d<2, \delta\le\gamma'} \binom{\gamma'}{\delta}\partial^{(\gamma'-\delta+\alpha)}\left(\frac{A_\alpha}{\tilde{A}}\right)\cdot\partial^{(\delta+\alpha)}u. \tag{52.24}$$

Since $(\gamma' + \alpha)_d < j$, the induction hypothesis (i.e. Eq. (52.22) is valid for $|\gamma| < j$) shows the right member of Eq. (52.24) is in $L^2(N_\rho)$ and gives the estimate

$$\begin{aligned}\left\|\sum_{|\alpha|\le 2, \alpha_d<2} \partial^{\gamma'}\left(\frac{A_\alpha}{\tilde{A}}\partial^\alpha u\right)\right\|_{L^2(N_\rho)} &\lesssim \sum_{|\alpha|\le 2, \alpha_d<2, \delta\le\gamma'} \left\|\partial^{(\delta+\alpha)}u\right\|_{L^2(N_\rho)} \\ &\lesssim \|f\|_{H^k(N_t)} + \|u\|_{H^1(N_t)}.\end{aligned}$$

Combining this with Eq. (52.23) gives $\partial^\gamma u \in L^2(N_\rho)$ and

$$\begin{aligned}\|\partial^\gamma u\|_{L^2(N_\rho)} &\lesssim \|f\|_{H^{|\gamma'|}(N_t)} + \left\|\sum_{|\alpha|\le 2, \alpha_d<2} \partial^{\gamma'}\left(\frac{A_\alpha}{\tilde{A}}\partial^\alpha u\right)\right\|_{L^2(N_\rho)} \\ &\lesssim \|f\|_{H^k(N_t)} + \|f\|_{H^k(N_t)} + \|u\|_{H^1(N_t)} \lesssim \|f\|_{H^k(N_t)} + \|u\|_{H^1(N_t)}.\end{aligned} \tag{52.25}$$

■

The following assumptions an notation will be in force for the remainder of this chapter.

Assumption 5 *Let Ω be a bounded open subset such that $\bar{\Omega}^o = \Omega$ and $\bar{\Omega}$ is a C^∞ – manifold with boundary, X be either $H_0^1(\Omega)$ or $H^1(\Omega)$ and $\mathcal{E}$ be a Dirichlet form as in Notation 51.1 which is assumed to be elliptic. Also if W is an open subset of $\mathbb{R}^d$ let*

$$X_W := \{v \in X : \operatorname{supp}(v) \sqsubset\sqsubset W \cap \bar{\Omega}\}.$$

Lemma 52.14. *For each $p \in \partial\Omega$ there exists precompact open neighborhoods V and W in $\mathbb{R}^d$ such that $\bar{V} \subset W$, for each $k \in \mathbb{N}$ there is a constant $C_k < \infty$ such that if $u \in X$ and $f \in H^k(\Omega)$ satisfies*

$$\mathcal{E}(u, v) = \int_\Omega fv\, dx \text{ for all } v \in X_W \tag{52.26}$$

then $u \in H^{k+2}(V \cap \Omega)$ and

$$\|u\|_{H^{k+2}(V\cap\Omega)} \le C(\|f\|_{H^k(\Omega)} + \|u\|_{H^1(\Omega)}) \tag{52.27}$$

Proof. Let W be an open neighborhood of p such that there exists a chart $\psi : W \to B(0, r)$ with inverse $\phi := \psi^{-1} : B(0, r) \to W$ satisfying:

1. The maps ψ and ϕ has bounded derivatives to all orders.
2. $\psi(W \cap \Omega) = B(0, r) \cap \mathbb{H}^d = N_r$ and $\psi(W \cap \operatorname{bd}(\Omega)) = B(0, r) \cap \operatorname{bd}(\mathbb{H}^d)$.

Now let $\rho < r$ and define $V := \phi(B(0, \rho))$, see Figure 52.5.

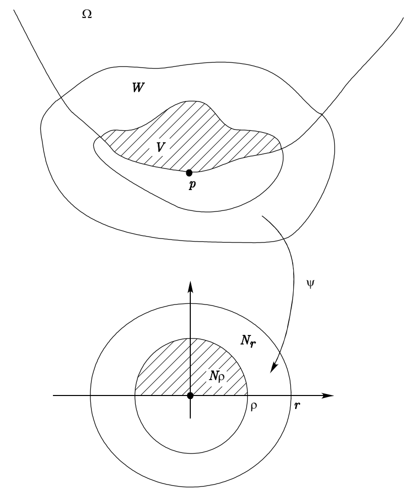


Fig. 52.5. Flattening out the boundary of Ω in a neighborhood of p.

Suppose that $u \in X$ satisfies Eq. (52.26) and $v \in X_W$. Then making the change of variables $x = \phi(y)$,

$$\int_\Omega fv dm = \int_{N_r} f(\phi(y))\, v(\phi(y))\, J(y) dy = \int_{N_r} \tilde{f}(y)\tilde{v}(y) dy$$

where $J(y) := |\det \phi'(y)|$, $\tilde{f}(y) := J(y) f(\phi(y))$ and $\tilde{v}(y) = v(\phi(y))$. By the change of variables theorem, $\phi^* v := v \circ \phi$ is the generic element of $X_r(N_r)$ and $\tilde{f} \in H^k(N_r)$. We also define a quadratic form on $X(N_r)$ by

$$Q(\tilde{u}, \tilde{v}) := \sum_{|\alpha|,|\beta| \le 1} \int_W a_{\alpha\beta} \partial^\alpha (\tilde{u} \circ \psi) \cdot \partial^\beta (\tilde{v} \circ \psi)\, dm.$$

Again by making change of variables (using Theorem 48.16 along with the change of variables theorem for integrals) this quadratic form may be written in the standard form,

$$Q(\tilde{u}, \tilde{v}) = \sum_{|\alpha|,|\beta| \le 1} \int_{N_r} \tilde{a}_{\alpha,\beta} \partial^\alpha \tilde{u} \cdot \partial^\beta \tilde{v}\, dm.$$

This new form is still elliptic. To see this let Γ denote the matrix (a_{ij}), then

$$\begin{aligned}\sum_{i,j=1}^d a_{ij} \partial_i (\tilde{u} \circ \psi) \cdot \partial_j (\tilde{v} \circ \psi) &= \Gamma \nabla (\tilde{u} \circ \psi) \cdot \nabla (\tilde{u} \circ \psi) \\ &= \Gamma [\psi']^{tr} \nabla \tilde{u} \circ \psi \cdot [\psi']^{tr} \nabla \tilde{v} \circ \psi\end{aligned}$$

which shows

$$\tilde{a}_{ij} = \Gamma [\psi']^{tr} e_i \cdot [\psi']^{tr} e_i$$

and

$$\sum_{i,j=1}^d \tilde{a}_{ij} \xi_i \xi_j = \Gamma [\psi']^{tr} \xi \cdot [\psi']^{tr} \xi \ge \epsilon \left| [\psi']^{tr} \xi \right|^2 \ge \epsilon\delta |\xi|^2$$

where

$$\delta = \inf \left\{ \left| [\psi'(x)]^{tr} \xi \right|^2 : |\xi| = 1 \ \&\ x \in W \right\} > 0.$$

Then Eq. (52.26) implies

$$Q(\tilde{u}, \tilde{v}) = \int_{N_r} \tilde{f}(y) \tilde{v}(y) dy \text{ for all } \tilde{v} \in X_r.$$

Therefore by local boundary regularity Theorem 52.13, $\tilde{u} \in H^{k+2}(N_\rho)$ and there exists $C < \infty$ such that

$$\|\tilde{u}\|_{H^{k+2}(N_\rho)} \le C(\|\tilde{f}\|_{H^k(N_t)} + \|\tilde{u}\|_{H^1(N_t)}). \tag{52.28}$$

Invoking the change of variables Theorem 48.16 again shows $u \in H^k(V)$ and the estimate in Eq. (52.28) implies the estimated in Eq. (52.27). ■

Theorem 52.15 (Elliptic Regularity). *Let Ω be a bounded open subset such that $\bar{\Omega}^o = \Omega$ and $\bar{\Omega}$ is a C^∞ – manifold with boundary, X be either $H_0^1(\Omega)$ or $H^1(\Omega)$ and $\mathcal{E}$ be a Dirichlet form as in Notation 51.1. If $k \in \mathbb{N}$ and $u \in X$ such that $\mathcal{L}_\mathcal{E} u \in H^k(\Omega)$ then $u \in H^{k+2}(\Omega)$ and*

$$\|u\|_{H^{k+2}(\Omega)} \le C(\|f\|_{H^k(\Omega)} + \|u\|_{X^*}) \le C(\|f\|_{H^k(\Omega)} + \|u\|_{L^2(\Omega)}). \tag{52.29}$$

Proof. Cover $\partial\Omega$ with $\{V_i\}_{i=1}^N$ and $\{W_i\}_{i=1}^N$ as in the above Lemma 52.14 such that $\bar{V}_i \sqsubset\sqsubset W_i$. Also choose a precompact open subset V_0 contained in Ω such that $\{V_i\}_{i=0}^N$ covers $\overline{\Omega}$. Choose W_0 such that $\overline{V}_0 \subset W_0$ and $\bar{W}_0 \subset \Omega$. If $\mathcal{L}_\mathcal{E} u =: f \in H^k(\Omega)$, then by Lemma 52.14 for $i \ge 1$ and Theorem 52.5 for $i = 0$, $u \in H^{k+2}(V_i)$ and there exist $C_i < \infty$ such that

$$\|u\|_{H^{k+2}(V_i \cap \Omega)} \le C_i(\|f\|_{H^k(W_i \cap \Omega)} + \|u\|_{H^1(W_i \cap \Omega)}). \tag{52.30}$$

Summing Eq. (52.30) on i implies $u \in H^{k+2}(\Omega)$ and

$$\|u\|_{H^{k+2}(\Omega)} \le C(\|f\|_{H^k(\Omega)} + \|u\|_X). \tag{52.31}$$

Finally

$$\begin{aligned}\|u\|_X^2 &\le C(\mathcal{E}(u,u) + \|u\|_{H^{-1}(\Omega)}^2) \\ &= C((f,u)_{L^2(\Omega)} + \|u\|_{H^{-1}(\Omega)}^2) \\ &\le C(\|f\|_{L^2(\Omega)} \|u\|_{L^2(\Omega)} + \|u\|_{H^{-1}(\Omega)}^2) \\ &\le C(\frac{1}{2\delta}\|f\|_{L^2(\Omega)}^2 + \frac{\delta}{2}\|u\|_X^2 + \|u\|_{H^{-1}(\Omega)}^2)\end{aligned}$$

for any $\delta > 0$. Choosing δ so that $C\delta = 1$, we find

$$\frac{1}{2}\|u\|_X^2 \le C(\frac{1}{2\delta}\|f\|_{L^2(\Omega)}^2 + \|u\|_{H^{-1}(\Omega)}^2)$$

which implies with a new constant C that

$$\|u\|_X \le C\left(\|f\|_{L^2(\Omega)} + \|u\|_{H^{-1}(\Omega)}\right). \tag{52.32}$$

Combining Eqs. (52.31) and (52.32) implies Eq. (52.29). ■

53

Unbounded operators and quadratic forms

53.1 Unbounded operator basics

Definition 53.1. *If X and Y are Banach spaces and D is a subspace of X, then a linear transformation T from D into Y is called a linear transformation (or operator) from X to Y with domain D. We will sometimes wr If D is dense in X, T is said to be* densely defined.

Notation 53.2 *If S and T are operators from X to Y with domains $D(S)$ and $D(T)$ and if $D(S) \subset D(T)$ and $Sx = Tx$ for $x \in D(S)$, then we say T is an* extension *of S and write $S \subset T$.*

We note that $X \times Y$ is a Banach space in the norm

$$\|\langle x, y\rangle\| = \sqrt{\|x\|^2 + \|y\|^2}.$$

If H and K are Hilbert spaces, then $H \times K$ and $K \times H$ become Hilbert spaces by defining

$$(\langle x, y\rangle, \langle x', y'\rangle)_{H\times K} := (x, x')_H + (y, y')_K$$

and

$$(\langle y, x\rangle, \langle y', x'\rangle)_{K\times H} := (x, x')_H + (y, y')_K.$$

Definition 53.3. *If T is an operator from X to Y with domain D, the* graph *of T is*

$$\Gamma(T) := \{\langle x, Dx\rangle : x \in D(T)\} \subset H \times K.$$

Note that $\Gamma(T)$ is a subspace of $X \times Y$.

Definition 53.4. *An operator $T : X \to Y$ is* closed *if $\Gamma(T)$ is closed in $X \times Y$.*

Remark 53.5. It is easy to see that T is closed iff for all sequences $x_n \in D$ such that there exists $x \in X$ and $y \in Y$ such that $x_n \to x$ and $Tx_n \to y$ implies that $x \in \mathcal{D}$ and $Tx = y$.

Let H be a Hilbert space with inner product $(\cdot, \cdot)$ and norm $\|v\| := \sqrt{(v, v)}$. As usual we will write H^* for the continuous dual of H and $\overline{H^*}$ for the continuous conjugate linear functionals on H. Our convention will be that $(\cdot, v) \in H^*$ is linear while $(v, \cdot) \in \overline{H^*}$ is conjugate linear for all $v \in H$.

Lemma 53.6. *Suppose that $T : H \to K$ is a densely defined operator between two Hilbert spaces H and K. Then*

1. *T^* is always a closed but not necessarily densely defined operator.*
2. *If T is closable, then $\bar{T}^* = T^*$.*
3. *T is closable iff $T^* : K \to H$ is densely defined.*
4. *If T is closable then $\bar{T} = T^{**}$.*

Proof. Suppose $\{v_n\} \subset D(T)$ is a sequence such that $v_n \to 0$ in H and $Tv_n \to k$ in K as $n \to \infty$. Then for $l \in D(T^*)$, by passing to the limit in the equality, $(Tv_n, l) = (v_n, T^*l)$ we learn $(k, l) = (0, T^*l) = 0$. Hence if T^* is densely defined, this implies $k = 0$ and hence T is closable. This proves one direction in item 3. To prove the other direction and the remaining items of the Lemma it will be useful to express the graph of T^* in terms of the graph of T. We do this now.

Recall that $k \in D(T^*)$ and $T^*k = h$ iff $(k, Tx)_K = (h, x)_H$ for all $x \in D(T)$. This last condition may be written as $(k, y)_K - (h, x)_H = 0$ for all $\langle x, y\rangle \in \Gamma(T)$.

Let $V : H \times K \to K \times H$ be the unitary map defined by $V\langle x, y\rangle = \langle -y, x\rangle$. With this notation, we have $\langle k, h\rangle \in \Gamma(T^*)$ iff $\langle k, h\rangle \perp V\Gamma(T)$, i.e.

$$\Gamma(T^*) = (V\Gamma(T))^{\perp} = V(\Gamma(T)^{\perp}), \tag{53.1}$$

where the last equality is a consequence of V being unitary. As a consequence of Eq. (53.1), $\Gamma(T^*)$ is always closed and hence T^* is always a closed operator, and this proves item 1. Moreover if T is closable, then

$$\Gamma(T^*) = V\Gamma(T)^{\perp} = V\overline{\Gamma(T)}^{\perp} = V\Gamma(\bar{T})^{\perp} = \Gamma(\bar{T}^*)$$

which proves item 2.

Now suppose T is closable and $k \perp \mathcal{D}(T^*)$. Then

$$\langle k, 0\rangle \in \Gamma(T^*)^{\perp} = V\Gamma(T)^{\perp\perp} = V\overline{\Gamma(T)} = V\Gamma(\bar{T}),$$

where $\bar{T}$ denotes the closure of T. This implies that $\langle 0, k\rangle \in \Gamma(\bar{T})$. But $\bar{T}$ is a well defined operator (by the assumption that T is closable) and hence $k = \bar{T}0 = 0$. Hence we have shown $\mathcal{D}(T^*)^{\perp} = \{0\}$ which implies $\mathcal{D}(T^*)$ is dense in K. This completes the proof of item 3.

4. Now assume T is closable so that T^* is densely defined. Using the obvious analogue of Eq. (53.1) for T^* we learn $\Gamma(T^{**}) = U\Gamma(T^*)^{\perp}$ where $U\langle y, x\rangle = \langle -x, y\rangle = -V^{-1}\langle y, x\rangle$. Therefore,

$$\Gamma(T^{**}) = UV(\Gamma(T)^{\perp})^{\perp} = -\overline{\Gamma(T)} = \overline{\Gamma(T)} = \Gamma(\bar{T})$$

and hence $\bar{T} = T^{**}$. ∎

Lemma 53.7. *Suppose that H and K are Hilbert spaces, $T : H \to K$ is a densely defined operator which has a densely defined adjoint T^*. Then* $\mathrm{Nul}(T^*) = \mathrm{Ran}(T)^\perp$ *and* $\mathrm{Nul}(\bar{T}) = \mathrm{Ran}(T^*)^\perp$ *where $\bar{T}$ denotes the closure of T.*

Proof. Suppose that $k \in \mathrm{Nul}(T^*)$ and $h \in \mathcal{D}(T)$, then $(k, Th) = (T^*k, h) = 0$. Since $h \in \mathcal{D}(T)$ is arbitrary, this proves that $\mathrm{Nul}(T^*) \subset \mathrm{Ran}(T)^\perp$. Now suppose that $k \in \mathrm{Ran}(T)^\perp$. Then $0 = (k, Th)$ for all $h \in D(T)$. This shows that $k \in \mathcal{D}(T^*)$ and that $T^*k = 0$. The assertion $\mathrm{Nul}(\bar{T}) = \mathrm{Ran}(T^*)^\perp$ follows by replacing T by T^* in the equality, $\mathrm{Nul}(T^*) = \mathrm{Ran}(T)^\perp$. ■

Definition 53.8. *A quadratic form q on H is a dense subspace $\mathcal{D}(q) \subset H$ called the domain of q and a sesquilinear form $q : \mathcal{D}(q) \times \mathcal{D}(q) \to \mathbb{C}$. (**Sesquilinear** means that $q(\cdot, v)$ is linear while $q(v, \cdot)$ is conjugate linear on $\mathcal{D}(q)$ for all $v \in \mathcal{D}(q)$.) The form q is **symmetric** if $q(v, w) = \overline{q(w, v)}$ for all $v, w \in \mathcal{D}(q)$, q is **positive** if $q(v) \geq 0$ (here $q(v) = q(v, v)$) for all $v \in \mathcal{D}(q)$, and q is **semi-bounded** if there exists $M_0 \in (0, \infty)$ such that $q(v, v) \geq -M_0 \|v\|^2$ for all $v \in \mathcal{D}(q)$.*

53.2 Lax-Milgram Methods

For the rest of this section q will be a sesquilinear form on H and to simplify notation we will write X for $\mathcal{D}(q)$.

Theorem 53.9 (Lax-Milgram). *Let $q : X \times X \to \mathbb{C}$ be a sesquilinear form and suppose the following added assumptions hold.*

1. *X is equipped with a Hilbertian inner product $(\cdot, \cdot)_X$.*
2. *The form q is **bounded** on X, i.e. there exists a constant $C < \infty$ such that $|q(v, w)| \leq C\|v\|_X \cdot \|w\|_X$ for all $v, w \in X$.*
3. *The form q is **coercive,** i.e. there exists $\epsilon > 0$ such that $|q(v, v)| \geq \epsilon \|v\|_X^2$ for all $v \in X$.*

Then the maps $\mathcal{L} : X \to \overline{X^*}$ and $\mathcal{L}^\dagger : X \to X^*$ defined by $\mathcal{L}v := q(v, \cdot)$ and $\mathcal{L}^\dagger v := q(\cdot, v)$ are linear and (respectively) conjugate linear isomorphisms of Hilbert spaces. Moreover

$$\|\mathcal{L}^{-1}\| \leq \epsilon^{-1} \text{ and } \|(\mathcal{L}^\dagger)^{-1}\| \leq \epsilon^{-1}.$$

Proof. The operator $\mathcal{L}$ is bounded because

$$\|\mathcal{L}v\|_{X^*} = \sup_{w \neq 0} \frac{|q(v, w)|}{\|w\|_X} \leq C\|v\|_X. \tag{53.2}$$

Similarly $\mathcal{L}^\dagger$ is bounded with $\left\|\mathcal{L}^\dagger\right\| \leq C$.

Let $\beta : X \to \overline{X^*}$ denote the linear Riesz isomorphism defined by $\beta(x) = (x, \cdot)_X$ for $x \in X$. Define $R := \beta^{-1}\mathcal{L} : X \to X$ so that $\mathcal{L} = \beta R$, i.e.

$$\mathcal{L}v = q(v, \cdot) = (Rv, \cdot)_X \text{ for all } v \in X.$$

Notice that R is a bounded **linear** map with operator bound less than C by Eq. (53.2). Since

$$\left(\mathcal{L}^\dagger v\right)(w) = q(w, v) = (Rw, v)_X = (w, R^*v)_X \text{ for all } v, w \in X,$$

we see that $\mathcal{L}^\dagger v = (\cdot, R^*v)_X$, i.e. $R^* = \bar{\beta}^{-1}\mathcal{L}^\dagger$, where $\bar{\beta}(x) := \overline{(x, \cdot)_X} = (\cdot, x)_X$. Since β and $\bar{\beta}$ are linear and conjugate linear isometric isomorphisms, to finish the proof it suffices to show R is invertible and that $\|R^{-1}\|_X \leq \epsilon^{-1}$.

Since

$$|(v, R^*v)_X| = |(Rv, v)_X| = |q(v, v)| \geq \epsilon \|v\|_X^2, \tag{53.3}$$

one easily concludes that $\mathrm{Nul}(R) = \{0\} = \mathrm{Nul}(R^*)$. By Lemma 53.7, $\overline{\mathrm{Ran}(R)} = \mathrm{Nul}(R^*)^\perp = \{0\}^\perp = X$ and so we have shown $R : X \to X$ is injective and has a dense range. From Eq. (53.3) and the Schwarz inequality, $\epsilon \|v\|_X^2 \leq \|Rv\|_X \|v\|_X$, i.e.

$$\|Rv\|_X \geq \epsilon \|v\|_X \text{ for all } v \in X. \tag{53.4}$$

This inequality proves the range of R is closed. Indeed if $\{v_n\}$ is a sequence in X such that $Rv_n \to w \in X$ as $n \to \infty$ then Eq. (53.4) implies

$$\epsilon \|v_n - v_m\|_X \leq \|Rv_n - Rv_m\|_X \to 0 \text{ as } m, n \to \infty.$$

Thus $v := \lim_{n \to \infty} v_n$ exists in X and hence $w = Rv \in \mathrm{Ran}(R)$ and so $\mathrm{Ran}(R) = \overline{\mathrm{Ran}(R)}^X = X$. So $R : X \to X$ is a bijective map and hence invertible. By replacing v by $R^{-1}v$ in Eq. (53.4) we learn R^{-1} is bounded with operator norm no larger than ϵ^{-1}. ■

Theorem 53.10. *Let q be a bounded coercive sesquilinear form on X as in Theorem 53.9. Further assume that the inclusion map $i : X \to H$ is bounded and let L and $L^\dagger$ be the unbounded linear operators on H defined by:*

$$\mathcal{D}(L) := \{v \in X : w \in X \to q(v, w) \text{ is } H \text{ - continuous}\},$$
$$\mathcal{D}(L^\dagger) := \{w \in X : v \in X \to q(v, w) \text{ is } H \text{ - continuous}\}$$

and for $v \in \mathcal{D}(L)$ and $w \in \mathcal{D}(L^\dagger)$ define $Lv \in H$ and $L^\dagger w \in H$ by requiring

$$q(v, \cdot) = (Lv, \cdot) \text{ and } q(\cdot, w) = (\cdot, L^\dagger w).$$

Then $\mathcal{D}(L)$ and $\mathcal{D}(L^\dagger)$ are dense subspaces of X and hence of H. The operators $L^{-1} : H \to \mathcal{D}(L) \subset H$ and $(L^\dagger)^{-1} : H \to \mathcal{D}(L^\dagger) \subset H$ are bounded when viewed as operators from H to H with norms less than or equal to $\epsilon^{-1} \|i\|_{L(X,H)}^2$. Furthermore, $L^* = L^\dagger$ and $(L^\dagger)^* = L$ and in particular both L and $L^\dagger = L^*$ are closed operators.

Proof. Let $\alpha : H \to \overline{X^*}$ be defined by $\alpha(v) = (v, \cdot)|_X$. If $(v,\cdot)_X$ is perpendicular to $\alpha(H) = \overline{i^*(H^*)} \subset \overline{X^*}$, then

$$0 = ((v,\cdot)_X, \alpha(w))_{\overline{X^*}} = ((v,\cdot)_X, (w,\cdot))_{\overline{X^*}} = (v,w) \text{ for all } w \in H.$$

Taking $w = v$ in this equation shows $v = 0$ and hence the orthogonal complement of $\alpha(H)$ in $\overline{X^*}$ is $\{0\}$ which implies $\alpha(H) = \overline{i^*(H^*)}$ is dense in $\overline{X^*}$.

Using the notation in Theorem 53.9, we have $v \in \mathcal{D}(L)$ iff $\mathcal{L}v \in \overline{i^*(H^*)} = \alpha(H)$ iff $v \in \mathcal{L}^{-1}(\alpha(H))$ and for $v \in \mathcal{D}(L)$, $\mathcal{L}v = (Lv,\cdot)|_X = \alpha(Lv)$. This and a similar computation shows

$$\mathcal{D}(L) = \mathcal{L}^{-1}(\overline{i^*(H^*)}) = \mathcal{L}^{-1}(\alpha(H)) \text{ and } \mathcal{D}(L^\dagger) := (\mathcal{L}^\dagger)^{-1}(i^*(H^*)) = (\mathcal{L}^\dagger)^{-1}(\bar{\alpha}(I$$

and for $v \in \mathcal{D}(L)$ and $w \in \mathcal{D}(L^\dagger)$ we have $\mathcal{L}v = (Lv,\cdot)|_X = \alpha(Lv)$ and $\mathcal{L}^\dagger w = (\cdot, L^\dagger w)|_X = \bar{\alpha}(L^\dagger w)$. The following commutative diagrams summarizes the relationships of L and $\mathcal{L}$ and $L^\dagger$ and $\mathcal{L}^\dagger$,

$$\begin{array}{ccc} & X & \overset{\mathcal{L}}{\to} \overline{X^*} \\ i & \uparrow & \uparrow \ \alpha \\ & D(L) & \overset{L}{\to} H \end{array} \text{ and } \begin{array}{ccc} & X & \overset{\mathcal{L}^\dagger}{\to} X^* \\ i & \uparrow & \uparrow \ \bar{\alpha} \\ & D(L^\dagger) & \overset{L^\dagger}{\to} H \end{array}$$

where in each diagram i denotes an inclusion map. Because $\mathcal{L}$ and $\mathcal{L}^\dagger$ are invertible, $L : D(L) \to H$ and $L^\dagger : D(L^\dagger) \to H$ are invertible as well. Because both $\mathcal{L}$ and $\mathcal{L}^\dagger$ are isomorphisms of X onto $\overline{X^*}$ and X^* respectively and $\alpha(H)$ is dense in $\overline{X^*}$ and $\bar{\alpha}(H)$ is dense in X^*, the spaces $\mathcal{D}(L)$ and $\mathcal{D}(L^\dagger)$ are dense subspaces of X, and hence also of H.

For the norm bound assertions let $v \in \mathcal{D}(L) \subset X$ and use the coercivity estimate on q to find

$$\begin{aligned} \epsilon\|v\|_H^2 &\le \epsilon \|i\|_{L(X,H)}^2 \|v\|_X^2 \le \|i\|_{L(X,H)}^2 |q(v,v)| = \|i\|_{L(X,H)}^2 |(Lv,v)_H| \\ &\le \|i\|_{L(X,H)}^2 \|Lv\|_H \|v\|_H. \end{aligned}$$

Hence $\epsilon\|v\|_H \le \|i\|_{L(X,H)}^2 \|Lv\|_H$ for all $v \in \mathcal{D}(L)$. By replacing v by $L^{-1}v$ (for $v \in H$) in this last inequality, we find

$$\|L^{-1}v\|_H \le \frac{\|i\|_{L(X,H)}^2}{\epsilon} \|v\|_H, \text{ i..e } \|L^{-1}\|_{B(H)} \le \epsilon^{-1} \|i\|_{L(X,H)}^2.$$

Similarly one shows that $\|(L^\dagger)^{-1}\|_{B(H)} \le \epsilon^{-1} \|i\|_{L(X,H)}^2$ as well.

For $v \in \mathcal{D}(L)$ and $w \in \mathcal{D}(L^\dagger)$,

$$(Lv, w) = q(v,w) = (v, L^\dagger w) \tag{53.5}$$

which shows $L^\dagger \subset L^*$. Now suppose that $w \in \mathcal{D}(L^*)$, then

$$q(v,w) = (Lv,w) = (v, L^* w) \text{ for all } v \in \mathcal{D}(L).$$

By continuity if follows that

$$q(v,w) = (v, L^* w) \text{ for all } v \in X$$

and therefore by the definition of $L^\dagger$, $w \in \mathcal{D}(L^\dagger)$ and $L^\dagger w = L^* w$, i.e. $L^* \subset L^\dagger$. Since we have shown $L^\dagger \subset L^*$ and $L^* \subset L^\dagger$, $L^\dagger = L^*$. A similar argument shows that $\left(L^\dagger\right)^* = L$. Because the adjoints of operators are always closed, both $L = \left(L^\dagger\right)^*$ and $L^\dagger = L^*$ are closed operators. ■

Corollary 53.11. *If q in Theorem 53.10 is further assumed to be symmetric then L is self-adjoint, i.e. $L^* = L$.*

Proof. This simply follows from Theorem 53.10 upon observing that $L = L^\dagger$ when q is symmetric. ■

53.3 Close, symmetric, semi-bounded quadratic forms and self-adjoint operators

Definition 53.12. *A symmetric, sesquilinear quadratic form $q : X \times X \to \mathbb{C}$ is **closed** if whenever $\{v_n\}_{n=1}^\infty \subset X$ is a sequence such that $v_n \to v$ in H and*

$$q(v_n - v_m) := q(v_n - v_m, v_n - v_m) \to 0 \text{ as } m, n \to \infty$$

*implies that $v \in X$ and $\lim_{n\to\infty} q(v - v_n) = 0$. The form q is said to be **closable** iff for all $\{v_n\} \subset X$ such that $v_n \to 0 \in H$ and $q(v_n - v_m) \to 0$ as $m, n \to \infty$ implies that $q(v_n) \to 0$ as $n \to \infty$.*

Example 53.13. Let H and K be Hilbert spaces and $T : H \to K$ be a densely defined operator. Set $q(v,w) := (Tv, Tw)_K$ for $v, w \in X := \mathcal{D}(q) := \mathcal{D}(T)$. Then q is a positive symmetric quadratic form on H which is closed iff T is closed and is closable iff T is closable.

For the remainder of this section let $q : X \times X \to \mathbb{C}$ be a symmetric, sesquilinear quadratic form which is semi-bounded and satisfies $q(v) \ge -M_0 \|v\|^2$ for all $v \in X$ and some $M_0 < \infty$.

Notation 53.14 *For $v, w \in X$ and $M > M_0$ let $(v,w)_M := q(v,w) + M(v,w)$. Notice that*

$$\begin{aligned} \|v\|_M^2 &= q(v) + M\|v\|^2 = q(v) + M_0\|v\|^2 + (M - M_0)\|v\|^2 \\ &\ge (M - M_0)\|v\|^2, \end{aligned} \tag{53.6}$$

from which it follows that $(\cdot,\cdot)_M$ is an inner product on X and $i : X \to H$ is bounded by $(M - M_0)^{-1/2}$. Let H_M denote the Hilbert space completion of $(X, (\cdot,\cdot)_M)$.

Formally, $H_M = \mathcal{C}/\sim$, where $\mathcal{C}$ denotes the collection of $\|\cdot\|_M$–Cauchy sequences in X and $\sim$ is the equivalence relation, $\{v_n\} \sim \{u_n\}$ iff $\lim_{n\to\infty} \|v_n - u_n\|_M = 0$. For $v \in X$, let $i(v)$ be the equivalence class of the constant sequence with elements v. Notice that if $\{v_n\}$ and $\{u_n\}$ are in $\mathcal{C}$, then $\lim_{m,n\to\infty}(v_n, u_m)_M$ exists. Indeed, let C be a finite upper bound for $\|u_n\|_M$ and $\|v_n\|_M$. (Why does this bound exist?) Then

$$\begin{aligned}|(v_n, u_m)_M - (v_k, u_l)_M| &= |(v_n - v_k, u_m)_M + (v_k, u_m - u_l)_M| \\ &\leq C\{\|v_n - v_k\|_M + \|u_m - u_l\|_M\} \end{aligned} \tag{53.7}$$

and this last expression tends to zero as $m, n, k, l \to \infty$. Therefore, if $\bar{v}$ and $\bar{u}$ denote the equivalence class of $\{v_n\}$ and $\{u_n\}$ in $\mathcal{C}$ respectively, we may define $(\bar{v}, \bar{u})_M := \lim_{m,n\to\infty}(v_n, u_m)_M$. It is easily checked that H_M with this inner product is a Hilbert space and that $i : X \to H_M$ is an isometry.

Remark 53.15. The reader should verify that all of the norms, $\{\|\cdot\|_M : M > M_0\}$, on X are equivalent so that H_M is independent of $M > M_0$.

Lemma 53.16. *The inclusion map $i : X \to H$ extends by continuity to a continuous linear map $\hat{\imath}$ from H_M into H. Similarly, the quadratic form $q : X \times X \to \mathbb{C}$ extends by continuity to a continuous quadratic form $\hat{q} : H_M \times H_M \to \mathbb{C}$. Explicitly, if $\bar{v}$ and $\bar{u}$ denote the equivalence class of $\{v_n\}$ and $\{u_n\}$ in $\mathcal{C}$ respectively, then $\hat{\imath}(\bar{v}) = H - \lim_{n\to\infty} v_n$ and $\hat{q}(\bar{v}, \bar{u}) = \lim_{m,n\to\infty} q(v_n, u_n)$.*

Proof. This routine verification is left to the reader. ■

Lemma 53.17. *Let q be as above and $M > M_0$ be given.*

1. *The quadratic form q is closed iff $(X, (\cdot,\cdot)_M)$ is a Hilbert space.*
2. *The quadratic form q is closable iff the map $\hat{\imath} : H_M \to H$ is injective. In this case we identify H_M with $\hat{\imath}(H_M) \subset H$ and therefore we may view $\hat{q}$ as a quadratic form on H. The form $\hat{q}$ is called the* ***closure*** *of q and as the notation suggests is a closed quadratic form on H.*

A more explicit description of $\hat{q}$ is as follows. The domain $\mathcal{D}(\hat{q})$ consists of those $v \in H$ such that there exists $\{v_n\} \subset X$ such that $v_n \to v$ in H and $q(v_n - v_m) \to 0$ as $m, n \to \infty$. If $v, w \in \mathcal{D}(\hat{q})$ and $v_n \to v$ and $w_n \to w$ as just described, then $\hat{q}(v, w) := \lim_{n\to\infty} q(v_n, w_n)$.

Proof. 1. Suppose q is closed and $\{v_n\}_{n=1}^\infty \subset X$ is a $\|\cdot\|_M$ – Cauchy sequence. By the inequality in Eq. (53.6), $\{v_n\}_{n=1}^\infty$ is $\|\cdot\|_H$ – Cauchy and hence $v := \lim_{n\to\infty} v_n$ exists in H. Moreover,

$$q(v_n - v_m) = \|v_n - v_m\|_M^2 - M\|v_n - v_m\|_H^2 \to 0$$

and therefore, because q is closed, $v \in \mathcal{D}(q) = X$ and $\lim_{n\to\infty} q(v - v_n) = 0$ and hence $\lim_{n\to\infty} \|v_n - v\|_M^2 = 0$. The converse direction is simpler and will be left to the reader.

2. The proof that q is closable iff the map $\hat{\imath} : H_M \to H$ is injective will be complete once the reader verifies that the following assertions are equivalent. 1) $\hat{\imath} : H_1 \to H$ is injective, 2) $\hat{\imath}(\bar{v}) = 0$ implies $\bar{v} = 0$, 3) if $v_n \overset{H}{\to} 0$ and $q(v_n - v_m) \to 0$ as $m, n \to \infty$ implies that $q(v_n) \to 0$ as $n \to \infty$.

By construction H_M equipped with the inner product $(\cdot,\cdot)_M := \hat{q}(\cdot,\cdot) + M(\cdot,\cdot)$ is complete. So by item 1. it follows that $\hat{q}$ is a closed quadratic form on H if q is closable. ■

Example 53.18. Suppose $H = L^2([-1,1])$, $\mathcal{D}(q) = C([-1,1])$ and $q(f,g) := f(0)\bar{g}(0)$ for all $f, g \in \mathcal{D}(q)$. The form q is not closable. Indeed, let $f_n(x) = (1+x^2)^{-n}$, then $f_n \to 0 \in L^2$ as $n \to \infty$ and $q(f_n - f_m) = 0$ for all m, n while $q(f_n - 0) = q(f_n) = 1 \nrightarrow 0$ as $n \to \infty$. This example also shows the operator $T : H \to \mathbb{C}$ defined by $\mathcal{D}(T) = C([-1,1])$ with $Tf = f(0)$ is not closable.

Let us also compute T^* for this example. By definition $\lambda \in D(T^*)$ and $T^*\lambda = f$ iff $(f, g) = \lambda\overline{Tg} = \lambda\overline{g(0)}$ for all $g \in C([-1,1])$. In particular this implies $(f, g) = 0$ for all $g \in C([-1,1])$ such that $g(0) = 0$. However these functions are dense in H and therefore we conclude that $f = 0$ and hence $\mathcal{D}(T^*) = \{0\}$!!

Exercise 53.19. Keeping the notation in Example 53.18, show $\overline{\Gamma(T)} = H \times \mathbb{C}$ which is clearly not the graph of a linear operator $S : H \to \mathbb{C}$.

Proposition 53.20. *Suppose that $A : H \to H$ is a densely defined positive symmetric operator, i.e. $(Av, w) = (v, Aw)$ for all $v, w \in \mathcal{D}(A)$ and $(v, Av) \geq 0$ for all $v \in \mathcal{D}(A)$. Define $q_A(v, w) := (v, Aw)$ for $v, w \in \mathcal{D}(A)$. Then q_A is closable and the closure $\hat{q}_A$ is a non-negative, symmetric closed quadratic form on H.*

Proof. Let $(\cdot,\cdot)_1 = (\cdot,\cdot) + q_A(\cdot,\cdot)$ on $\mathcal{D}(A) \times \mathcal{D}(A)$, $v_n \in \mathcal{D}(A)$ such that H-$\lim_{n\to\infty} v_n = 0$ and

$$q_A(v_n - v_m) = (A(v_n - v_m), (v_n - v_m)) \to 0 \text{ as } m, n \to \infty.$$

Then

$$\limsup_{n\to\infty} q_A(v_n) \leq \lim_{n\to\infty} \|v_n\|_1^2 = \lim_{m,n\to\infty} (v_m, v_n)_1 = \lim_{m,n\to\infty} \{(v_m, v_n) + (v_m, Av_n)\} =$$

where the last equality follows by first letting $m \to \infty$ and then $n \to \infty$. Notice that the above limits exist because of Eq. (53.7). ■

Lemma 53.21. *Let A be a positive self-adjoint operator on H and define $q_A(v, w) := (v, Aw)$ for $v, w \in \mathcal{D}(A) = \mathcal{D}(q_A)$. Then q_A is closable and the closure of q_A is*

$$\hat{q}_A(v, w) = (\sqrt{A}v, \sqrt{A}w) \text{ for } v, w \in X := \mathcal{D}(\hat{q}_A) = \mathcal{D}(\sqrt{A}).$$

Proof. Let $\hat{q}(v,w) = (\sqrt{A}v, \sqrt{A}w)$ for $v, w \in X = \mathcal{D}(\sqrt{A})$. Since $\sqrt{A}$ is self-adjoint and hence closed, it follows from Example 53.13 that $\hat{q}$ is closed. Moreover, $\hat{q}$ extends q_A because if $v, w \in \mathcal{D}(A)$, then $v, w \in \mathcal{D}(A) = \mathcal{D}((\sqrt{A})^2)$ and $\hat{q}(v,w) = (\sqrt{A}v, \sqrt{A}w) = (v, Aw) = q_A(v,w)$. Thus to show $\hat{q}$ is the closure of q_A it suffices to show $\mathcal{D}(A)$ is dense in $X = \mathcal{D}(\sqrt{A})$ when equipped with the Hilbertian norm, $\|w\|_1^2 = \|w\|^2 + \hat{q}(w)$.

Let $v \in \mathcal{D}(\sqrt{A})$ and define $v_m := 1_{[0,m]}(A)v$. Then using the spectral theorem along with the dominated convergence theorem one easily shows that $v_m \in X = \mathcal{D}(A)$, $\lim_{m\to\infty} v_m = v$ and $\lim_{m\to\infty} \sqrt{A}v_m = \sqrt{A}v$. But this is equivalent to showing that $\lim_{m\to\infty} \|v - v_m\|_1 = 0$. ■

Theorem 53.22. *Suppose $q : X \times X \to \mathbb{C}$ is a symmetric, closed, semi-bounded (say $q(v,v) \geq -M_0\|v\|^2$) sesquilinear form. Let $L : H \to H$ be the possibly unbounded operator defined by*

$$D(L) := \{v \in X : q(v, \cdot) \text{ is } H - \text{continuous}\}$$

and for $v \in D(L)$ let $Lv \in H$ be the unique element such that $q(v,\cdot) = (Lv, \cdot)|_X$. Then

1. *L is a densely defined self-adjoint operator on H and $L \geq -M_0 I$.*
2. *$D(L)$ is a* ***form core*** *for q, i.e. the closure of $D(L)$ is a dense subspace in $(X, \|\cdot\|_M)$. More explicitly, for all $v \in X$ there exists $v_n \in D(L)$ such that $v_n \to v$ in H and $q(v - v_n) \to 0$ as $n \to \infty$.*
3. *For and $M \geq M_0$, $D(q) = D\left(\sqrt{L + MI}\right)$.*
4. *Letting $q_L(v,w) := (Lv, w)$ for all $v, w \in D(L)$, we have q_L is closable and $\hat{q}_L = q$.*

Proof. 1. From Lemma 53.17, $(X, (\cdot,\cdot)_X := (\cdot,\cdot)_M)$ is a Hilbert space for any $M > M_0$. Applying Theorem 53.10 and Corollary 53.11 with q being $(\cdot,\cdot)_X$ gives a self-adjoint operator $L_M : H \to H$ such that

$$D(L_M) := \{v \in X : (v,\cdot)_X \text{ is } H - \text{continuous}\}$$

and for $v \in D(L_M)$,

$$(L_M v, w)_H = (v,w)_X = q(v,w) + M(v,w) \text{ for all } w \in X. \tag{53.8}$$

Since $(v,\cdot)_X$ is H – continuous iff $q(v,\cdot)$ is H – continuous it follows that $D(L_M) = D(L)$ and moreover Eq. (53.8) is equivalent to

$$((L_M - MI)\, v, w)_H = q(v,w) \text{ for all } w \in X.$$

Hence it follows that $L := L_M - MI$ and so L is self-adjoint. Since $(Lv, v) = q(v,v) \geq -M_0 \|v\|^2$, we see that $L \geq -M_0 I$.

2. The density of $\mathcal{D}(L) = \mathcal{D}(L_M)$ in $(X, (\cdot,\cdot)_M)$ is a direct consequence of Theorem 53.10.

3. For

$$v, w \in \mathcal{D}(Q) := \mathcal{D}\left(\sqrt{L_M}\right) = \mathcal{D}\left(\sqrt{L + MI}\right) = \mathcal{D}\left(\sqrt{L + M_0 I}\right)$$

let $Q(v,w) := \left(\sqrt{L_M}v, \sqrt{L_M}w\right)$. For $v, w \in D(L)$ we have

$$Q(v,w) = (L_M v, w) = (Lv, w) + M\,(v,w) = q(v,w) + M\,(v,w) = (v,w)_M\,.$$

By Lemma 53.21, Q is a closed, non-negative symmetric form on H and $\mathcal{D}(L) = \mathcal{D}(L_M)$ is dense in $(\mathcal{D}(Q), Q)$. Hence if $v \in \mathcal{D}(Q)$ there exists $v_n \in \mathcal{D}(L)$ such that $Q(v - v_n) \to 0$ and this implies $q(v_m - v_n) \to 0$ as $m, n \to \infty$. Since q is closed, this implies $v \in \mathcal{D}(q)$ and furthermore that $Q(v,w) = (v,w)_M$ for all $v, w \in \mathcal{D}(Q)$.

Conversely, by item 2., if $v \in X = \mathcal{D}(q)$, there exists $v_n \in \mathcal{D}(L)$ such that $\|v - v_m\|_M \to 0$. From this it follows that $Q(v_m - v_n) \to 0$ as $m, n \to \infty$ and therefore since Q is closed, $v \in \mathcal{D}(Q)$ and again $Q(v,w) = (v,w)_M$ for all $v, w \in \mathcal{D}(q)$. This proves item 3. and also shows that

$$q(v,w) = \left(\sqrt{L + MI}v, \sqrt{L + MI}w\right) - M(v,w) \text{ for all } v, w \in X$$

where $X := \mathcal{D}\left(\sqrt{L_M}\right)$.

4. Since $q_L \subset q$, q_L is closable and the closure of q_L is still contained in q. Since $q_L = Q - L(\cdot,\cdot)$ on $D(L)$ and the closure of $Q|_{D(L)} = (\cdot,\cdot)_M$, it is easy to conclude that the closure of q_L is q as well. ■

Notation 53.23 *Let $\mathcal{P}$ denote the collection of positive self-adjoint operators on H and $\mathcal{Q}$ denote the collection of positive and closed symmetric forms on H.*

Theorem 53.24. *The map $A \in \mathcal{P} \to \hat{q}_A \in \mathcal{Q}$ is bijective, where $\hat{q}_A(v,w) := (\sqrt{A}v, \sqrt{A}w)$ with $\mathcal{D}(\hat{q}_A) = \mathcal{D}(\sqrt{A})$ is the closure of the quadratic form $q_A(v,w) := (Av, w)$ for $v, w \in \mathcal{D}(q) := \mathcal{D}(A)$. The inverse map is given by $q \in \mathcal{Q} \to A_q \in \mathcal{P}$ where A_q is uniquely determined by*

$$\mathcal{D}(A_q) = \{v \in \mathcal{D}(q) : q(v,\cdot) \text{ is } H \text{ - continuous}\} \text{ and}$$
$$(A_q v, w) = q(v,w) \text{ for } v \in \mathcal{D}(A_q) \text{ and } w \in \mathcal{D}(q).$$

Proof. From Lemma 53.21, $\hat{q}_A \in \mathcal{Q}$ and $\hat{q}_A$ is the closure of q_A. From Theorem 53.22 $A_q \in \mathcal{P}$ and

$$q(\cdot,\cdot) = \left(\sqrt{A_q}\cdot, \sqrt{A_q}\cdot\right) = \hat{q}_{A_q}.$$

So to finish the proof it suffices to show $A \in \mathcal{P} \to \hat{q}_A \in \mathcal{Q}$ is injective. However, again by Theorem 53.22, if $q \in \mathcal{Q}$ and $A \in \mathcal{P}$ such that $q = \hat{q}_A$, then $v \in \mathcal{D}(A_q)$ and $A_q v = w$ iff

$$(\sqrt{A}v, \sqrt{A}\cdot) = q(v,\cdot) = (A_q v, \cdot)|_X.$$

But this implies $\sqrt{A}v \in \mathcal{D}\left(\sqrt{A}\right)$ and $A_q v = \sqrt{A}\sqrt{A}v = Av$. But by the spectral theorem, $D\left(\sqrt{A}\sqrt{A}\right) = D(A)$ and so we have proved $A_q = A$. ■

53.4 Construction of positive self-adjoint operators

The main theorem concerning closed symmetric semi-bounded quadratic forms q is Friederich's extension theorem.

Corollary 53.25 (The Friederich's extension). *Suppose that $A : H \to H$ is a densely defined positive symmetric operator. Then A has a positive self-adjoint extension $\hat{A}$. Moreover, $\hat{A}$ is the only self-adjoint extension of A such that $\mathcal{D}(\hat{A}) \subset \mathcal{D}(\hat{q}_A)$.*

Proof. By Proposition 53.20, $q := \hat{q}_A$ exists in $\mathcal{Q}$. By Theorem 53.24, there exists a unique positive self-adjoint operator B on H such that $\hat{q}_B = q$. Since for $v \in \mathcal{D}(A)$, $q(v, w) = (Av, w)$ for all $w \in X$, it follows from Eq. (**??**) and (**??**) that $v \in \mathcal{D}(B)$ and $Bv = Av$. Therefore $\hat{A} := B$ is a self-adjoint extension of A.

Suppose that C is another self-adjoint extension of A such that $\mathcal{D}(C) \subset X$. Then $\hat{q}_C$ is a closed extension of q_A. Thus $q = \hat{q}_A \subset \hat{q}_C$, i.e. $\mathcal{D}(\hat{q}_A) \subset \mathcal{D}(\hat{q}_C)$ and $\hat{q}_A = \hat{q}_C$ on $\mathcal{D}(\hat{q}_A) \times \mathcal{D}(\hat{q}_A)$. For $v \in \mathcal{D}(C)$ and $w \in \mathcal{D}(A)$, we have that

$$\hat{q}_C(v, w) = (Cv, w) = (v, Cw) = (v, Aw) = (v, Bw) = q(v, w).$$

By continuity it follows that

$$\hat{q}_C(v, w) = (Cv, w) = (v, Bw) = q(v, w)$$

for all $w \in \mathcal{D}(B)$. Therefore, $v \in \mathcal{D}(B^*) = \mathcal{D}(B)$ and $Bv = B^*v = Cv$. That is $C \subset B$. Taking adjoints of this equation shows that $B = B^* \subset C^* = C$. Thus $C = B$. ■

Corollary 53.26 (von Neumann). *Suppose that $D : H \to K$ is a closed operator, then $A = D^*D$ is a positive self-adjoint operator on H.*

Proof. The operator D^* is densely defined by Lemma 53.6. The quadratic form $q(v, w) := (Dv, Dw)$ for $v, w \in X := \mathcal{D}(D)$ is closed (Example 53.13) and positive. Hence by Theorem 53.24 there exists an $A \in \mathcal{P}$ such that $q = \hat{q}_A$, i.e.

$$(Dv, Dw) = \left(\sqrt{A}v, \sqrt{A}w\right) \text{ for all } v, w \in X = \mathcal{D}(D) = \mathcal{D}(\sqrt{A}). \tag{53.9}$$

Recalling that $v \in D(A) \subset X$ and $Av = g$ happens iff

$$(Dv, Dw) = q(v, w) = (g, w) \text{ for all } w \in X$$

and this happens iff $Dv \in D(D^*)$ and $D^*Dv = g$. Thus we have shown $D^*D = A$ which is self-adjoint and positive. ■

53.5 Applications to partial differential equations

Let $U \subset \mathbb{R}^n$ be an open set, $\rho \in C^1(U \to (0, \infty))$ and for $i, j = 1, 2, \ldots, n$ let $a_{ij} \in C^1(U, \mathbb{R})$. Take $H = L^2(U, \rho dx)$ and define

$$q(f, g) := \int_U \sum_{i,j=1}^n a_{ij}(x)\partial_i f(x)\partial_j g(x)\, \rho(x)\, dx$$

for $f, g \in X = C_c^2(U)$.

Proposition 53.27. *Suppose that $a_{ij} = a_{ji}$ and that $\sum_{i,j=1}^n a_{ij}(x)\xi_i\xi_j \geq 0$ for all $\xi \in \mathbb{R}^n$. Then q is a symmetric closable quadratic form on H. Hence there exists a unique self-adjoint operator $\hat{A}$ on H such that $\hat{q} = \hat{q}_{\hat{A}}$. Moreover $\hat{A}$ is an extension of the operator*

$$Af(x) = -\frac{1}{\rho(x)} \sum_{i,j=1}^n \partial_j(\rho(x)a_{ij}(x)\partial_i f(x))$$

for $f \in \mathcal{D}(A) = C_c^2(U)$.

Proof. A simple integration by parts argument shows that $q(f, g) = (Af, g)_H = (f, Ag)_H$ for all $f, g \in \mathcal{D}(A) = C_c^2(U)$. Thus by Proposition 53.20, q is closable. The existence of $\hat{A}$ is a result of Theorem 53.24. In fact $\hat{A}$ is the Friederich's extension of A as in Corollary 53.25. ■

Given the above proposition and the spectral theorem, we now know that (at least in some weak sense) we may solve the general heat and wave equations: $u_t = -Au$ for $t \geq 0$ and $u_{tt} = -Au$ for $t \in \mathbb{R}$. Namely, we will take

$$u(t, \cdot) := e^{-t\hat{A}}u(0, \cdot)$$

and

$$u(t, \cdot) = \cos(t\sqrt{\hat{A}})u(0, \cdot) + \frac{\sin(t\sqrt{\hat{A}})}{\sqrt{\hat{A}}}u_t(0, \cdot)$$

respectively. In order to get classical solutions to the equations we would have to better understand the operator $\hat{A}$ and in particular its domain and the domains of the powers of $\hat{A}$. This will be one of the topics of the next part of the course dealing with Sobolev spaces.

Remark 53.28. By choosing $\mathcal{D}(A) = C_c^2(U)$ we are essentially using Dirichlet boundary conditions for A and $\hat{A}$. If U is a bounded region with C^2–boundary, we could have chosen (for example VERIFY THIS EXAMPLE)

$$\mathcal{D}(A) = \{f \in C^2(U) \cap C^1(\bar{U}) : \text{ with } \partial u/\partial n = 0 \text{ on } \partial U\}.$$

This would correspond to Neumann boundary conditions. Proposition 53.27 would be valid with this domain as well provided we assume that $a_{i,j}$ and ρ are in $C^1(\bar{U})$.

For a second application let $H = L^2(U, \rho dx; \mathbb{R}^N)$ and for $j = 1, 2, \ldots, n$, let $A_j : U \to \mathcal{M}_{N\times N}$ (the $N \times N$ matrices) be a C^1 function. Set $\mathcal{D}(D) := C_c^1(U \to \mathbb{R}^N)$ and for $S \in \mathcal{D}(D)$ let $DS(x) = \sum_{i=1}^n A_i(x)\partial_i S(x)$.

Proposition 53.29 ("Dirac Like Operators"). *The operator D on H defined above is closable. Hence $A := D^*\bar{D}$ is a self-adjoint operator on H, where $\bar{D}$ is the closure of D.*

Proof. Again a simple integration by parts argument shows that $\mathcal{D}(D) \subset \mathcal{D}(D^*)$ and that for $S \in \mathcal{D}(D)$,

$$D^*S(x) = \frac{1}{\rho(x)} \sum_{i=1}^n -\partial_i(\rho(x)A_i(x)S(x)).$$

In particular D^* is a densely defined operator and hence D is closable. The result now follows from Corollary 53.26. ■

54

L^2 – operators associated to $\mathcal{E}$

Let $\bar{\Omega}$ be a C^2 – manifold with boundary such that $\Omega = \bar{\Omega}^o$, $\rho \in C^\infty(\overline{\Omega})$ with $\rho > 0$ on $\overline{\Omega}$, $a_{\alpha,\beta} \in BC^\infty(\bar{\Omega})$ such that $(a_{ij}) \geq \epsilon I$ for some $\epsilon > 0$ and $\mathcal{E}$ be the elliptic Dirichlet form given by

$$\mathcal{E}(u,v) = \int_\Omega \sum a_{\alpha\beta}\partial^\alpha u \partial^\beta v \, d\mu$$

where $d\mu := \rho dm$. Let $H = L^2(\mu) \cong L^2(m)$ and $X = H^1(\Omega)$ or $H_0^1(\Omega)$.

Definition 54.1. *Let*

$$D(L_\mathcal{E}) = \{u \in X : \mathcal{L}_\mathcal{E} u := \mathcal{E}(u,\cdot) \in L^2(\rho\, dx)\}$$
$$D(L_\mathcal{E}^\dagger) = \{u \in X : \mathcal{L}_\mathcal{E}^\dagger u := \mathcal{E}(\cdot,u) \in L^2(\rho\, dx)\}$$

and for $u \in D(L_\mathcal{E})$ ($u \in D(L_\mathcal{E}^\dagger)$) define $L_\mathcal{E} u$ ($L_\mathcal{E}^\dagger u$) to be the unique element in $L^2(\mu)$ such that

$$\mathcal{E}(u,v) = (L_\mathcal{E} u, v)_{L^2(\mu)} = \int_\Omega L_\mathcal{E} u \cdot v \, \rho \, dm$$
$$\mathcal{E}(v,u) = (v, L_\mathcal{E} u)_{L^2(\mu)} = \int_\Omega v \cdot L_\mathcal{E}^\dagger u \, \rho \, dm$$

for all $v \in X$.

Theorem 54.2. *If $X = H_0^1(\Omega)$ let $B = B^\dagger \equiv 0$ and if $X = H^1(\Omega)$ let B and $B^\dagger$ be given (as in Proposition 51.4) by*

$$Bu := \sum a_{\alpha j}\partial^\alpha u \, n_j = n \cdot a\nabla u + (n \cdot a_{0,\cdot})\, u$$
$$B^\dagger u := \sum a_{i\beta}\partial^\beta u \, n_i = an \cdot \nabla u + (n \cdot a_{\cdot,0})\, u.$$

Then

$$D(L_\mathcal{E}) = \{u \in H^2(\Omega) \cap X : Bu|_{\partial\Omega} = 0\}$$
$$D(L_\mathcal{E}^\dagger) = \{u \in H^2(\Omega) \cap X : B^\dagger u|_{\partial\Omega} = 0\}$$

and

$$L_\mathcal{E} u = \frac{1}{\rho}\sum((-\partial)^\beta \rho a_{\alpha\beta}\partial^\alpha u) =: Lu$$
$$L_\mathcal{E}^\dagger u = \frac{1}{\rho}\sum(-\partial)^\alpha \rho a_{\alpha\beta}\partial^\beta u) =: L^\dagger u.$$

Moreover $L_\mathcal{E} = (L_\mathcal{E}^\dagger)^*$ and $L_\mathcal{E}^* = L_\mathcal{E}^\dagger$.

Proof. By replacing $\mathcal{E}(u,v)$ by $\mathcal{E}(u,v) + C(u,v)$, $L_\mathcal{E} \to L_\mathcal{E} + C$ and $L_\mathcal{E}^\dagger \to L_\mathcal{E}^\dagger + C$ for a sufficiently large constant C, we may assume that $\mathcal{E}(u,v)$ satisfies

$$\epsilon \|u\|_{H^1}^2 \leq \mathcal{E}(u,u) \text{ for all } u \in H^1.$$

Then by Theorem 53.10, $L_\mathcal{E}^* = L_\mathcal{E}^\dagger$, $(L_\mathcal{E}^\dagger)^* = L_\mathcal{E}$ and

$$L_\mathcal{E} : D(L_\mathcal{E}) \to L^2(\Omega) \text{ and } L_\mathcal{E}^\dagger : D(L_\mathcal{E}^\dagger) \to L^2(\Omega)$$

are linear isomorphisms. By the elliptic regularity Theorem 52.15, both $D(L_\mathcal{E})$ and $D(L_\mathcal{E}^\dagger)$ are subspaces of $H^2(\Omega)$ and moreover there is a constant $C < \infty$ such that

$$\|u\|_{H^2(\Omega)} \leq C \|L_\mathcal{E} u\|_{L^2(\Omega)}. \tag{54.1}$$

From Proposition 51.4 (integration by parts), for $u \in H^2(\Omega)$ and $v \in X$,

$$\mathcal{E}(u,v) = (Lu,v)_{L^2(\mu)} + \int_{\partial\Omega} Bu\big|_{\partial\Omega} \cdot v\big|_{\partial\Omega} \rho \, d\sigma \tag{54.2}$$

while, by definition, if $u \in D(L_\mathcal{E})$ then

$$\mathcal{E}(u,v) = (L_\mathcal{E} u, v)_{L^2(\mu)} \text{ for all } v \in X. \tag{54.3}$$

Choosing $v \in H_0^1(\Omega) \subset X$, comparing Eqs. (54.2) and (54.3) shows that $Lu = L_\mathcal{E} u$. So for $u \in D(L_\mathcal{E})$, $L_\mathcal{E} u = Lu$ and moreover we must have $Bu\big|_{\partial\Omega} = 0$ as well. Therefore

$$D(L_\mathcal{E}) \subset H^2(\Omega) \cap \{u : Bu\big|_{\partial\Omega} = 0\}.$$

Conversely if $u \in H^2(\Omega)$ with $Bu\big|_{\partial\Omega} = 0$, $\mathcal{E}(u,v) = (Lu,v)_{L^2(\mu)}$ for all $v \in X$ and therefore by definition of $\mathcal{L}_\mathcal{E}$, $u \in D(L_\mathcal{E})$ and $L_\mathcal{E} u = Lu$. The assertions involving $L_\mathcal{E}^\dagger$ are proved in the same way. ■

54.1 Compact perturbations of the identity and the Fredholm Alternative

Definition 54.3. *A bounded operator $F : H \to B$ is* ***Fredholm*** *iff the* $\dim \mathrm{Nul}(F) < \infty$, $\dim$ *coker*$(F) < \infty$ *and Ran(F) is closed in B. (Recall: coker$(F) := B/\mathrm{Ran}(F)$.) The* ***index*** *of F is the integer,*

$$index(F) = \dim \mathrm{Nul}(F) - \dim coker(F) \tag{54.4}$$
$$= \dim \mathrm{Nul}(F) - \dim \mathrm{Nul}(F^*). \tag{54.5}$$

Example 54.4. Suppose that H and B are finite dimensional Hilbert spaces and $F : H \to B$ is a linear operator. In this case, the rank nullity theorem implies

$$\begin{aligned}\operatorname{index}(F) &= \dim \operatorname{Nul}(F) - \dim \operatorname{coker}(F) \\ &= \dim \operatorname{Nul}(F) - [\dim B - \dim \operatorname{Ran}(F)] \\ &= \dim \operatorname{Nul}(F) + \dim \operatorname{Ran}(F) - \dim B \\ &= \dim H - \dim B.\end{aligned}$$

Theorem 54.5. *If $R : H \to H$ is finite rank, then $F = I + R$ is Fredholm and index$(F) = 0$.*

Proof. Let $H_1 = \operatorname{Nul}(R)$, $H_2 = \operatorname{Ran}(R)$, $P_i : H \to H_i$ be orthogonal projection, $\{\psi_i\}_{i=1}^n$ be an orthonormal basis for $\operatorname{Ran}(R)$ and $\phi_i := R^*\psi_i$ for $i = 1, 2, \dots, n$. Then for $h \in H$,

$$Rh = \sum_{i=1}^n (Rh, \psi_i)\, \psi_i = \sum_{i=1}^n (h, R^*\psi_i)\, \psi_i = \sum_{i=1}^n (h, \phi_i)\, \psi_i$$

and hence $\{\phi_1, \dots, \phi_n\}^\perp \subset \operatorname{Nul}(R)$. Therefore $H_2 = \operatorname{Nul}(R)^\perp \subset \operatorname{span}\{\phi_1, \dots, \phi_n\}$ is finite dimensional. For $h = h_1 + h_2 \in H_1 \oplus H_2$,

$$\begin{aligned}Fh &= (P_1 + P_2)(h_1 + h_2 + Rh_2) = (h_1 + P_1 R h_2) + (h_2 + P_2 R P_2 h_2) \\ &= (h_1 + P_1 R h_2) + (I_{H_2} + P_2 R P_2)\, h_2. \end{aligned} \tag{54.6}$$

From Eq. (54.6) we see that $h = h_1 + h_2 \in \operatorname{Nul}(F)$ iff $h_2 \in \operatorname{Nul}(I_{H_2} + P_2 R P_2)$ and $h_1 = -P_1 R h_2$ and hence

$$\operatorname{Nul}(F) \cong \operatorname{Nul}(I_{H_2} + P_2 R P_2). \tag{54.7}$$

It is also easily seen from Eq. (54.6) that

$$\operatorname{Ran}(F) = H_1 \oplus \operatorname{Ran}(I_{H_2} + P_2 R P_2). \tag{54.8}$$

Since H_2 is finite dimensional, $\operatorname{Ran}(I_{H_2} + P_2 R P_2)$ is a closed subspace of H_2 and so $\operatorname{Ran}(F)$ is closed. Moreover

$$\begin{aligned}\operatorname{coker}(F) &= H / \operatorname{Ran}(F) = [H_1 \oplus H_2] / [H_1 \oplus \operatorname{Ran}(I_{H_2} + P_2 R P_2)] \\ &\cong H_2 / \operatorname{Ran}(I_{H_2} + P_2 R P_2) = \operatorname{coker}(I_{H_2} + P_2 R P_2). \end{aligned} \tag{54.9}$$

So by Eqs. (54.7), (54.9) and Example 54.4,

$$\begin{aligned}\operatorname{index}(F) &= \dim \operatorname{Nul}(F) - \dim \operatorname{coker}(F) \\ &= \dim \operatorname{Nul}(I_{H_2} + P_2 R P_2) - \dim \operatorname{coker}(I_{H_2} + P_2 R P_2) \\ &= \operatorname{index}(I_{H_2} + P_2 R P_2) = 0.\end{aligned}$$

■

Corollary 54.6. *If $K : H \to H$ compact then $F = I + K$ is Fredholm and* index$(F) = 0$.

Proof. Choose $R_1 : H \to H$ finite rank such that $\varepsilon := K - R_1$ is a bounded operator with operator norm less than one. Then

$$F = I + K = I + \varepsilon + R_1 = (I + \varepsilon)(I + (I + \varepsilon)^{-1} R_1) = U(I + R),$$

where $U := (I + \varepsilon) : H \to H$ is invertible and $R := (I + \varepsilon)^{-1} R_1 : H \to H$ is finite rank. Therefore, $\operatorname{Ran}(F) = U(\operatorname{Ran}(I + R))$ is closed,

$$\dim \operatorname{coker}(F) = \dim \operatorname{coker}(I + R) < \infty, \tag{54.10}$$

$\operatorname{Nul}(F) = \operatorname{Nul}(I + R)$, and

$$\dim \operatorname{Nul}(F) = \dim \operatorname{Nul}(I + R). \tag{54.11}$$

From this it follows that F is Fredholm and $\operatorname{index}(F) = \operatorname{index}(I + R) = 0$. ■

54.2 Solvability of $Lu = f$ and properties of the solution

Theorem 54.7. *Let $\bar{\Omega} \subset \mathbb{R}^d$ be a C^∞ – manifold with boundary such that $\Omega = \bar{\Omega}^o$. Let $\mathcal{E}$ be an elliptic Dirichlet form, $L := L_{\mathcal{E}}$ be the associated operator.*

1. *For $C > 0$ sufficiently large, $(L + C) : D(L) \to L^2(\Omega)$ is a linear isomorphism and*
$$(L + C)^{-1} : L^2(\Omega) \to D(L) \subset H^2(\Omega)$$
is a bounded operator and $D(L)$ is a closed subspace of $H^2(\Omega)$.
2. *$(L + C)^{-1}$ as viewed as an operator from $L^2(\Omega)$ to $L^2(\Omega)$ is compact.*
3. *If $u \in D(L)$ and $Lu \in H^k(\Omega)$ then $u \in H^{k+2}(\Omega)$.*
4. *If $u \in D(L)$ and $Lu \in C^\infty(\overline{\Omega})$ then $u \in C^\infty(\overline{\Omega}) := \bigcap_{k=0}^\infty C^k(\overline{\Omega})$.*
5. *If $u \in D(L)$ is an eigenfunction of L, i.e. $Lu = \lambda u$ for some $\lambda \in \mathbb{C}$, then $u \in C^\infty(\overline{\Omega})$.*

Proof. 1. It was already shown in the proof of Theorem 54.2 that $(L+C) : D(L) \to L^2(\Omega)$ is bijective. Moreover the bound in Eq. (54.1) shows that $(L + C)^{-1} : L^2(\Omega) \to H^2(\Omega)$ is bounded. If $\{u_n\}_{n=1}^\infty \subset D(L) \subset H^2(\Omega)$ is a sequence such that $u_n \to u \in H^2(\Omega)$, then $\{(L + C)u_n\}_{n=1}^\infty$ is convergent in $L^2(\Omega)$ since $(L + C) : H^2(\Omega) \to L^2(\Omega)$ is bounded. Because L is a closed operator, it follows that $u \in D(L)$ and so $D(L)$ is a closed subspace of $H^2(\Omega)$.

2. This follows from item 1. and the Rellich - Kondrachov Compactness Theorem 49.25 which implies the embedding $H^2(\Omega) \hookrightarrow L^2(\Omega)$ is compact.

3. If $f \in H^k(\Omega)$ and $u \in D(L)$ such that $Lu = f \in H^k(\Omega)$ then $\mathcal{L}_{\mathcal{E}} u = f$ and hence the elliptic regularity Theorem 53.10 gives the result.

4. Since $C^\infty(\overline{\Omega}) \subset H^k(\Omega)$ for all k, it follows by item 1. that $u \in \bigcap_{k=0}^{\infty} H^k(\Omega)$. But $\bigcap_{k=0}^{\infty} H^k(\Omega) \subset C^\infty(\overline{\Omega})$ by the Sobolev embedding Theorem 49.18.

5. If $u \in D(L) \subset H^2(\Omega)$ and $Lu = \lambda u \in H^2(\Omega)$ for some $\lambda \in \mathbb{C}$, then by item 3., $u \in H^4(\Omega)$ and then reapplying item 3. we learn $u \in H^6(\Omega)$. This process may be repeated and so by induction, $u \in \bigcap_{k=0}^{\infty} H^k(\Omega) \subset C^\infty(\overline{\Omega})$. ∎

Theorem 54.8 (Fredholm Alternative). *Let $\bar{\Omega} \subset \mathbb{R}^d$ be a C^∞ – manifold with boundary such that $\Omega = \bar{\Omega}^o$. Let $\mathcal{E}$ be an elliptic Dirichlet form, $L := L_{\mathcal{E}}$ be the associated operator. Then*

1. $L : D(L) \to L^2(\Omega)$ *and* $L^* : D(L^*) \to L^2(\Omega)$ *are Fredholm operators.*
2. $\operatorname{index}(L) = \operatorname{index}(L^*) = 0$.
3. $\dim \operatorname{Nul}(L) = \dim \operatorname{Nul}(L^*)$.
4. $\operatorname{Ran}(L) = \operatorname{Nul}(L^*)^\perp$.
5. $\operatorname{Ran}(L) = L^2(\Omega)$ *iff* $\operatorname{Nul}(L) = \{0\}$.

Proof. Choose $C > 0$ such that $(L + C) : D(L) \to L^2(\Omega)$ is a invertible map and let

$$K := C(L + C)^{-1} : L^2(\Omega) \to D(L)$$

which by Theorem 54.7 is compact when viewed as an operator from $L^2(\Omega)$ to $L^2(\Omega)$. With this notation we have

$$(L + C)^{-1}L = (L + C)^{-1}(L + C - C) = I_{D(L)} - K$$

and

$$L(L + C)^{-1} = (L + C)(L + C)^{-1} - C(L + C)^{-1} = I_{L^2(\Omega)} - K.$$

By Corollary 54.6 orProposition 16.35, $I_{L^2(\Omega)} - K$ is a Fredholm operator with $\operatorname{index}(I_{L^2(\Omega)} - K) = 0$. Since $\operatorname{Ran}(L) = \operatorname{Ran}(L(L + C)^{-1}) = \operatorname{Ran}(I - K)$ it follows that $\operatorname{Ran}(L)$ is a closed and finite codimension subspace of $L^2(\Omega)$ and

$$\dim \operatorname{coker}(L) = \dim \operatorname{coker}(I - K).$$

Since

$$u \in \operatorname{Nul}(L) \to (L + C)\, u \in \operatorname{Nul}(L(L + C)^{-1}) = \operatorname{Nul}(I - K)$$

is an isomorphism of vector spaces

$$\dim \operatorname{Nul}(L) = \dim \operatorname{Nul}(I - K) < \infty.$$

Combining the above assertions shows that L is a Fredholm operator and

$$\begin{aligned}\operatorname{index}(L) &= \dim \operatorname{coker}(L) - \dim \operatorname{Nul}(L)\\ &= \dim \operatorname{coker}(I - K) - \dim \operatorname{Nul}(I - K)\\ &= \operatorname{index}(I - K) = 0.\end{aligned}$$

The same argument applies to L^* to show L^* is Fredholm and $\operatorname{index}(L^*) = 0$. Because $\operatorname{Ran}(L)$ is closed and $\operatorname{Ran}(L)^\perp = \operatorname{Nul}(L^*)$, $\operatorname{Ran}(L) = \operatorname{Nul}(L^*)^\perp$ and

$$L^2(\Omega) = \operatorname{Nul}(L^*)^\perp \oplus \operatorname{Nul}(L^*) = \operatorname{Ran}(L) \oplus \operatorname{Nul}(L^*).$$

Thus $\dim \operatorname{coker}(L) = \dim \operatorname{Nul}(L^*)$ and so

$$\begin{aligned}0 = \operatorname{index}(L) &= \dim \operatorname{Nul}(L) - \dim \operatorname{coker}(L)\\ &= \dim \operatorname{Nul}(L) - \dim \operatorname{Nul}(L^*).\end{aligned}$$

This proves items 1-4 and finishes the proof of the theorem since item 5. is a direct consequence of items 3 and 4. ∎

Example 54.9 (Dirichlet Boundary Conditions). Let Δ denote the Laplacian with Dirichlet boundary conditions, i.e. $D(\Delta) = H_0^1(\Omega) \cap H^2(\Omega)$. If $u \in D(\Delta)$ then

$$\int_\Omega \nabla u \cdot \nabla v dm = (-\Delta u, v) \text{ for all } v \in H_0^1(\Omega) \tag{54.12}$$

and in particular for $u \in \operatorname{Nul}(\Delta)$ we have

$$\int_\Omega |\nabla u|^2 \, dm = (-\Delta u, u) = 0.$$

By the Poincaré inequality in Theorem 49.31 (or by more direct means) this implies $u = 0$ and therefore $\operatorname{Nul}(\Delta) = \{0\}$. It now follows by the Fredholm alternative in Theorem 54.8 that there exists a unique solution $u \in D(\Delta)$ to $\Delta u = f$ for any $f \in L^2(\Omega)$.

Example 54.10 (Neuwmann Boundary Conditions). Suppose Δ is the Laplacian on Ω with Neuwmann boundary conditions, i.e.

$$D(\Delta) = \{u \in H^2(\Omega) : \frac{\partial u}{\partial n} = 0\}.$$

If $u \in D(\Delta)$ then

$$\int_\Omega \nabla u \cdot \nabla v dm = (-\Delta u, v) \text{ for all } v \in H^1(\Omega). \tag{54.13}$$

so that the Dirichlet form associated to Δ is symmetric and hence $\Delta = \Delta^*$. Moreover if $u \in \operatorname{Nul}(\Delta)$, then

$$0 = (-\Delta u, v) = \int_\Omega |\nabla u|^2 \, dm,$$

i.e. $\nabla u = 0$. As in the proof of the Poincaré Lemma 49.30 (or using the Poincaré Lemma itself), u is constant on each connected component of Ω. Assuming, for simplicity, that Ω is connected, we have shown

$$\mathrm{Nul}(\Delta_N) = \mathrm{span}\,\{1\}.$$

The Fredholm alternative in Theorem 54.8 implies that there exists a (non unique solution) $u \in D(\Delta)$ to $\Delta u = f$ for precisely those $f \in L^2(\Omega)$ such that $f \perp 1$, i.e. $\int_\Omega f 1\, dm = 0$.

Remark 54.11. Suppose $\mathcal{E}$ is an elliptic Dirichlet form and $L = L_{\mathcal{E}}$ is the associated operator on $L^2(\Omega)$. If $\mathcal{E}$ has the property that the only solution to $\mathcal{E}(u,u) = 0$ is $u = 0$, then the equation $Lu = f$ always has a unique solution for any $f \in L^2(\Omega)$.

Example 54.12. Let $A_{ij} = A_{ji}, A_i$ and A_0 be in $C^\infty\left(\bar{\Omega}\right)$ with $A_0 \geq 0$ and $(A_{ij}) \geq \epsilon I$ for some $\epsilon > 0$. For $u, v \in H_0^1(\Omega)$ let

$$\mathcal{E}(u,v) = \int_\Omega (\sum A_{ij}\partial_i u \partial_j v + A_0 uv) dm, \tag{54.14}$$

and $L = L_{\mathcal{E}}$, then

$$L = -\partial_j A_{ij} \partial_i u + A_0 u, \tag{54.15}$$

with $D(L) := H^2(\Omega) \cap H_0^1(\Omega)$. If $u \in \mathrm{Nul}(L)$, then $0 = (Lu, u) = \mathcal{E}(u,u) = 0$ implies $\partial_i u = 0$ a.e. and hence u is constant on each connected component of Ω. Since $u \in H_0^1(\Omega)$, $u\Big|_{\partial\Omega} = 0$ from which we learn that $u \equiv 0$. Therefore $Lu = f$ has a (unique) solution for all $f \in L^2(\Omega)$.

Example 54.13. Keeping the same notation as Example 54.12, except now we view $\mathcal{E}$ as a Dirichlet form on $H^1(\Omega)$. Now $L = L_{\mathcal{E}}$ is the operator given in Eq. (54.15) but now

$$D(L) = \{u \in H^2(\Omega) : Bu = 0 \text{ on } \partial\Omega\}$$

where $Bu = n_j A_{ij} \partial_i u$. Again if $u \in \mathrm{Nul}(L)$ it follows that u is constant on each connected component of Ω. If we further assume that $A_0 > 0$ at some point in each connected component of Ω, we could then conclude from

$$0 = \mathcal{E}(u,u) = \int_\Omega A_0 u^2 dm,$$

that $u = 0$. So again $\mathrm{Nul}(L) = \{0\}$ and $\mathrm{Ran}\,(L) = L^2\,(\Omega)$.

54.3 Interior Regularity Revisited

Theorem 54.14 (Jazzed up interior regularity). *Let L be a second order elliptic differential operator on Ω. If $u \in L^2_{Loc}(\Omega)$ such that $Lu \in H^k_{Loc}(\Omega)$ then $u \in H^{k+2}_{Loc}(\Omega)$ and for any open precompact open sets Ω_0 and Ω_1 contained in Ω such that $\bar{\Omega}_0 \subset \Omega_1 \subset \bar{\Omega}_1 \subset \Omega$ there is a constant $C < \infty$ independent of u such that*

$$\|u\|_{H^{k+2}(\Omega_0)} \leq C(\|Lu\|_{H^k(\Omega_1)} + \|u\|_{L^2(\Omega_1)}).$$

Proof. When $k > 0$ the theorem follows from Theorem 52.5. So it suffices to consider the case, $k = 0$, i.e. $u \in L^2_{Loc}(\Omega)$ such that $Lu \in L^2_{Loc}(\Omega)$. To finish the proof, again because of Theorem 52.5, it suffices to show $u \in H^1_{loc}(\Omega)$. By replacing Ω by a precompact open subset of Ω which contains Ω_1 we may further assume that $u \in L^2(\Omega)$ and $Lu \in L^2(\Omega)$. Further, by replacing L by $L + C$ for some constant $C > 0$, the Lax-Milgram method implies we may assume $L : H_0^1(\Omega) \to H^{-1}(\Omega)$ is an isomorphism of Banach spaces. We will now finish the proof by showing $u \in H^1_{loc}(\Omega)$ under the above assumptions.

If $\chi \in C_c^\infty(\Omega)$, then by Lemma 50.7 $[L, M_\chi]$ is a first order operator so,

$$L(\chi u) = \chi Lu + [L, M_\chi]u =: f_\chi \in L^2(\Omega) + H^{-1}(\Omega) = H^{-1}(\Omega).$$

Let $u_0 = L^{-1} f_\chi \in H_0^1(\Omega)$, $\psi \in C_c^\infty(\Omega)$ such that $\psi = 1$ on a neighborhood of $\mathrm{supp}(\chi)$ and

$$v := \psi\,(\chi u - u_0) = \chi u - \psi u_0 \in L^2_c(\Omega).$$

Then, because $\mathrm{supp}(f_\chi) \subset \mathrm{supp}(\chi)$, we have $f_\chi = \psi f_\chi$ and

$$\begin{aligned} Lv &= L\,[\chi u - \psi u_0] = f_\chi - \psi L u_0 - [L, M_\psi]u_0 = f_\chi - \psi f_\chi - [L, M_\psi]u_0 \\ &= -[L, M_\psi]u_0 =: g \in L^2_c(\Omega). \end{aligned}$$

Let $D\,(L_D) := H^2(\Omega) \cap H_0^1(\Omega)$ and $L_D u = Lu$ for all $u \in D(L_D)$ so that L_D is L with Dirichlet boundary conditions on Ω. I now claim that $v \in D(L_D) \subset H_0^1(\Omega)$. To prove this suppose

$$\xi \in D(L_D^\dagger) = \left\{\xi \in H_0^1(\Omega) : L_D^\dagger \xi \in L^2(\Omega)\right\} = H_0^1(\Omega) \cap H^2(\Omega)$$

and let $\xi_m := \eta_m * \xi$ where η_m is an approximate δ – sequence so that $\xi_m \to \xi$ in $H^2_{loc}(\Omega)$. Choose $\phi \in C_c^\infty(\Omega)$ such that $\phi = 1$ on a neighborhood of $\mathrm{supp}(v) \supset \mathrm{supp}(g)$, then

$$\begin{aligned} (g, \xi) &= \lim_{m\to\infty}(g, \xi_m) = \lim_{m\to\infty}(Lv, \phi\xi_m) = \lim_{m\to\infty}(v, L^\dagger(\phi\xi_m)) \\ &= \lim_{m\to\infty}(v, L^\dagger \xi_m) = (v, L^\dagger \xi) = (v, L_D^\dagger \xi). \end{aligned}$$

Since this holds for all $\xi \in D(L_D^\dagger)$ we see that

$$v \in D\left(\left(L_D^\dagger\right)^*\right) = D(L_D) = H_0^1(\Omega) \cap H^2(\Omega) \subset H^1(\Omega),$$

where the first equality is a consequence of Theorem 54.2 which states $L_D = \left(L_D^\dagger\right)^*$. Therefore, $\chi u = \psi u_0 + v \in H^1(\Omega)$ and since $\chi \in C_c^\infty(\Omega)$ was arbitrary we learn that $u \in H^1_{loc}(\Omega)$. ■

54.4 Classical Dirichlet Problem

Let $\bar{\Omega}$ be a C^∞ – manifold with boundary such that $\Omega = \bar{\Omega}^o$ and let $L = \Delta$ with Dirichlet boundary conditions, so $D(\Delta) := H^1(\Omega) \cap H^2(\Omega)$.

Theorem 54.15. *To each $f \in C(\partial\Omega)$, there exists a unique solution $u \in C^\infty(\Omega) \cap C(\bar{\Omega})$ to the equation*

$$\Delta u = 0 \text{ with } u = f \text{ on } \partial\Omega.$$

Proof. Choose $f_n \in C^\infty(\overline{\Omega})$ such that $\lim_{n\to\infty} \|f_n|_{\partial\Omega} - f\|_{L^\infty(\partial\Omega)} = 0$. We will now show that there exists $u_n \in C^\infty(\overline{\Omega})$ such that

$$\Delta u_n = 0 \text{ with } u_n = f_n \text{ on } \partial\Omega. \tag{54.16}$$

To prove this let us write the desired solution as $u_n = v_n + f_n$ in which case $v_n = 0$ on $\partial\Omega$ and $0 = \Delta u_n = \Delta v_n + \Delta f_n$. Hence if v_n solves $\Delta v_n = -\Delta f_n$ on Ω with $v_n = 0$ on $\partial\Omega$ then $u_n = v_n + f_n$ solves the Dirichlet problem in Eq. (54.16).

By the maximum principle,

$$\|u_n - u_m\|_{L^\infty(\bar{\Omega})} \leq \|f_n - f_m\|_{L^\infty(\partial\Omega)} \to 0 \text{ as } m, n \to \infty$$

and so $\{u_n\}_{n=1}^\infty \subset C(\overline{\Omega})$ is uniformly convergent sequence. Let $u := \lim_{n\to\infty} u_n \in C(\overline{\Omega})$. The proof will be completed by showing $u \in C^\infty(\Omega)$ and $\Delta u = 0$. This can be done in one stroke by showing u satisfies the mean value property. This it the case since each function u_n satisfies the mean value property and this property is preserved under uniform limits. ■

Remark 54.16. Theorem 54.15 is more generally valid in the case Δ is replaced by an elliptic operator of the form $L = -\sum_{ij} \frac{1}{\rho}\partial^i(\rho a_{ij}\partial^j)$ with $\rho \in C^\infty(\overline{\Omega}, (0,\infty))$ and $a_{ij} \in C^\infty(\overline{\Omega})$ such that $\sum_{i,j=1}^n a_{ij}(x)\xi_i\xi_j \geq \epsilon|\xi|^2$ for all $x \in \overline{\Omega}, \xi \in \mathbb{R}^d$. Then again for all $f \in C(\partial\Omega)$ there exist a solution $u \in C^\infty(\Omega) \cap C(\overline{\Omega})$ such that

$$Lu = 0 \text{ with } u = f \text{ on } \partial\Omega.$$

The proof is the same as that of Theorem 54.15 except the last step needs to be changed as follows. As above, we construct solution to $Lu_n = 0$ with $u_n = f_n$ on $\partial\Omega$ and we then still have $u_n \to u \in C(\overline{\Omega})$ via the maximum principle. To finish the proof, because of Theorem 54.14, it suffices to show $Lu = 0$ in the sense of distributions. This is valid becasue if $\phi \in \mathcal{D}(\Omega)$ then

$$0 = \lim_{n\to\infty}\langle Lu_n, \phi\rangle = \lim_{n\to\infty}\langle u_n, L^\dagger\phi\rangle = \langle u, L^\dagger\phi\rangle = \langle Lu, \phi\rangle,$$

i.e. $Lu = 0$ in the sense of distribution.

54.5 Some Non-Compact Considerations

In this section we will make use of the results from Section 53.3. Let $\rho \in C^\infty(\mathbb{R}^d, (0,\infty))$, $A_{ij} \in BC^\infty(\mathbb{R}^d)$ such that $\sum A_{ij}\,\xi_i\,\xi_j > 0$ for all $|\xi| \neq 0$ and define

$$Q(u,v) = \int_{\mathbb{R}^d} \sum A_{ij}\partial^i u\,\partial^j v\,\rho\,dx \text{ for } u,v \in C_c^\infty\left(\mathbb{R}^d\right).$$

Then as we have seen Q has a closed extension $\hat{Q}$ and unique self adjoint operator L on $L^2(\rho\,dx)$ such that $u \in D(L)$ iff $v \to Q(u,v)$ is $L^2(\rho dx)$ continuous on $\mathcal{D}(\hat{Q})$ and in which case $Q(u,v) = (Lu,v)_{L^2(\rho dx)}$. Standard integration by parts shows

$$u \in C_c^2(\mathbb{R}^d) \subset \mathcal{D}(L) \text{ and } Lu = -\frac{1}{\rho}\sum \partial_j(\rho A_{ij}\partial_i u).$$

Proposition 54.17. *Let $\bar{Q}(u,u) := \int_{\mathbb{R}^d} \sum A_{ij}\partial^i u \cdot \partial^j u\,\rho\,dx$, then*

$$\mathcal{D}(\hat{Q}) \subset \{u \in L^2(\rho\,dx) \cap H^1_{loc}(\mathbb{R}^d) : \bar{Q}(u,u) < \infty\}.$$

Proof. By definition of the closure of Q, $C_c^\infty(\mathbb{R}^d)$ is dense in $(\mathcal{D}(\hat{Q}), \hat{Q}_1)$. Since for all $\Omega \subset\subset \mathbb{R}^d$ there exists an $\epsilon = \epsilon(\Omega)$ such that $\rho\sum A_{ij}\,\xi_i\,\xi_j \geq \epsilon|\xi|^2$ on Ω, we learn

$$Q(u,u) \geq \epsilon\|u\|^2_{H^1(\Omega)} \text{ for all } u \in C_c^\infty(\mathbb{R}^d). \tag{54.17}$$

Therefore if $u \in \mathcal{D}(\hat{Q})$ and $u_n \in C_c^\infty\left(\mathbb{R}^d\right)$ such that $\hat{Q}_1(u-u_n) \to 0$ as $n \to \infty$ then $\|u_n - u_m\|_{H^1(\Omega)} \to 0$, i.e. $u = \lim_{n\to\infty} u_n$ in $H^1(\Omega)$. Hence a simple limiting argument shows Eq. (54.17) holds for all $u \in \mathcal{D}(\hat{Q})$:

$$Q_1(u,u) \geq \epsilon\|u\|^2_{H^1(\Omega)} \text{ for all } u \in \mathcal{D}(\hat{Q}).$$

This shows $\mathcal{D}(\hat{Q}) \subset H^1_{Loc}(\mathbb{R}^d)$. Moreover

$$\begin{array}{ccc} \hat{Q}(u_n,u_n) & \geq & \int_\Omega A_{ij}\partial_i u_n\,\partial_j u_n\,\rho\,dx \\ \downarrow & n\to\infty & \downarrow \\ \hat{Q}(u,u) & \geq & \int_\Omega A_{ij}\partial_i u\partial_j u\,\rho\,dx. \end{array}$$

Since $\Omega \subset\subset \mathbb{R}^d$ is arbitrary this implies that

$$\hat{Q}(u,u) \geq \int_\Omega A_{ij}\partial_i u\, \partial_i u\, \rho\, dx.$$

■

Proposition 54.18. *Suppose $u \in D(L^k)$ then $u \in H^{2k}_{Loc}(\mathbb{R}^d)$ and for all $\Omega \subset\subset \Omega_1 \subset\subset \mathbb{R}^d$ there exist $C = C_k(\Omega)$ such that*

$$\|u\|_{H^{2k}(\Omega)} \leq C(\|L^k u\|_{L^2(\rho\, dx)} + \|u\|_{L^2(\rho\, dx)}) \tag{54.18}$$

Proof. Suppose $u \in D(L)$ and $Lu = f$. Then for all $\phi \in C_c^\infty(\mathbb{R}^d)$, $(Lu,\phi) = (f,\phi)_{L^2(\rho)}$. Therefore

$$\int \sum A_{ij}\partial_i u\, \partial_j u\, \rho\, dx = \int f\phi\, \rho\, dx$$

so $-\partial_j(\rho A_{ij}\,\partial_i u) = \rho f$ in the sense if distributions and hence

$$-\sum A_{ij}\,\partial_j\partial_i u + L.O.T. = f$$

in the distributional sense. Since $D(L) \subset D(\hat{Q}) \subset H^1_{Loc}(\mathbb{R}^d)$, by local elliptic regularity it follows that $D(L) \subset H^2_{Loc}(\mathbb{R}^d)$ and for all $\Omega_1 \supset \Omega_0$

$$\|u\|_{H^2(\Omega_0)} \leq C(\|Lu\|_{L^2(\Omega_1)} + \|u\|_{L^2(\Omega_1)}).$$

Now suppose $u \in D(L^2)$, then $u \in D(L) \subset H^2_{Loc}(\mathbb{R}^d)$ and $Lu \in D(L) \subset H^2_{Loc}(\mathbb{R}^d)$ implies $u \in H^4_{Loc}(\mathbb{R}^d)$ and

$$\begin{aligned}\|u\|_{H^4(\Omega)} &\leq C(\|Lu\|_{H^2(\Omega_1)} + \|Lu\|_{L^2(\Omega_1)})\\ &\leq C(L^2u\|_{L^2(\Omega_2)} + \|u\|_{L^2(\Omega_2)} + \|u\|_{H^2(\Omega_1)})\\ &\leq C(\|L^2u\|_{L^2(\Omega_2)} + \|Lu\|_{L^2(\Omega_2)} + \|u\|_{L^2(\Omega_2)}).\end{aligned}$$

If $u \in D(L^3)$ then $u \in H^4_{Loc}(\mathbb{R}^d)$ and $Lu \in H^4_{Loc}(\mathbb{R}^d)$ implies $u \in H^6_{Loc}(\mathbb{R}^d)$ and

$$\begin{aligned}\|u\|_{H^6(\Omega_0)} &\leq C(\|Lu\|_{H^4(\Omega_1)} + \|u\|_{L^2(\Omega_1)})\\ &\leq C(\|L^3u\|_{L^2(\Omega_2)} + \|L^2u\|_{L^2(\Omega_2)} + \|Lu\|_{L^2(\Omega_2)} + \|u\|_{L^2}).\end{aligned}$$

$u \in D(L^k)$ implies $u \in H^{2k}_{Loc}(\mathbb{R}^d)$ and

$$\begin{aligned}\|u\|_{H^2(\Omega_0)} &\leq C\sum_{j=0}^k \|L^j u\|_{L^2(\tilde{\Omega})} \leq C\sum_{j=0}^k \|L^j u\|_{L^2(\rho)}\\ &\leq C(\|L^k u\|_{L^2(\rho\, ds)} + \|u\|_{L^2(\rho\, dx)})\end{aligned}$$

by the spectral theorem. ■

54.5.1 Heat Equation

Let $u(t) = e^{-tL}u_0$ where $u_0 \in L^2(\rho)$. Then $u(t) \in D(L^k)$ for all k when $t > 0$ and hence $u(t) \in H^{2k}_{Loc}(\mathbb{R}^d)$ for all k. But this implies for each $t > 0$ that $u(t)$ has a continuous in fact C^∞- version because $H^{2k}(\Omega) \hookrightarrow C^{2k-2/d}(\bar{\Omega})$ for $k > \frac{1}{d}$. Moreover

$$\left\| L^k\left[\frac{u(t+h)-u(t)}{h} - Lu(t)\right]\right\|_{L^2(\Omega)} \to 0 \text{ as } h \to 0$$

for all $k = 0, 1, 2, \ldots$ and therefore,

$$\left\|\frac{u(t+h)-u(t)}{h} - Lu(t)\right\|_{C^{2k-2/d}(\bar{\Omega})} \to 0 \text{ as } h \to 0$$

when $k > \frac{1}{d}$. This shows $t \to u(t)$ is differentiable and in $C^{2k-\frac{2}{n}}(\overline{\Omega})$ for all $k > \frac{1}{d}$. Thus we conclude that $u(t,x)$ is in $C^{1,\infty}((0,\infty)\times\mathbb{R}^d)$ and $\frac{\partial u}{\partial t}(t,x) = L\,u(t,x)$, i.e. u is a classical solution to the heat equation.

54.5.2 Wave Equation

Now consider the generalized solution to the wave equation

$$u(t) = \cos(\sqrt{L}t)f + \frac{\sin(\sqrt{L}t)}{\sqrt{L}}g$$

where $f, g \in L^2(\rho)$. If $f, g \in C_c^\infty(\mathbb{R}^d)$, then $f, g \in D(L^k)$ for all k and hence $u(t) \in D(L^k)$ for all k. It now follows that $u(t)$ is C^∞-differentiable in t relative to the norm $\|f\|_k := \|f\|_{L^2(\rho dx)} + \|L^k f\|_{L^2(\rho dx)}$ for all $k \in \mathbb{N}$. So by the above ideas $u(t,x) \in C^\infty(\mathbb{R}\times\mathbb{R}^d)$ and

$$\begin{aligned}\ddot{u}(t,x) + Lu(t,x) &= 0 \text{ with}\\ u(0,x) &= f_0(x) \text{ and}\\ \dot{u}(0,x) &= g_0(x).\end{aligned}$$

55

Spectral Considerations

For this section, let Ω be a bounded open subset of $\mathbb{R}^d$ such that $\bar{\Omega}^o = \Omega$ and $\bar{\Omega}$ is C^∞ – manifold with boundary. Also let $\mathcal{E}$ be a symmetric Dirichlet form with domain being either $X = H^1(\Omega)$ or $X = H_0^1(\Omega)$ and let $L := L_\mathcal{E}$ be the corresponding self adjoint operator.

Theorem 55.1. *There exist $\{\lambda_i\}_{i=1}^\infty \subset \mathbb{R}$ such that $\lambda_1 \le \lambda_2 \le \lambda_3 \le \dots \to \infty$ and $\{\phi_n\} \subset D(L) \subset L^2(\Omega)$ such that $\{\phi_n\}$ is an orthonormal basis for $L^2(\Omega)$ and $L\phi_n = \lambda_n \phi_n$ for all n.*

Proof. Choose $C > 0$ such that $(L + C) : D(L) \to L^2(\Omega)$ is invertible and let $T := (L+C)^{-1}$ which is a compact operator (see Theorem 54.7) when viewed as an operator from $L^2(\Omega)$ to $L^2(\Omega)$. Since $L = L^*$,

$$((L+C)u, v) = (u, (L+C)v) \text{ for all } u, v \in D(L)$$

and using this equation with u, v being replaced by Tu, Tv respectively shows $(u, Tv) = (Tu, v)$ for all $u, v \in L^2$. Moreover if $u \in L^2(\Omega)$ and $v = Tu \in D(L)$,

$$(Tu, u) = (T(L+C)v, (L+C)v) = ((L+C)v, v) \ge 0$$

and so we have shown $T = T^*$ and $T > 0$. By the spectral Theorem 16.17 for self-adjoint compact operators, there exist $\{\mu_n\}_{n=1}^\infty \subset \mathbb{R}_+$ and an orthonormal basis $\{\phi_n\}_{n=1}^\infty$ of $L^2(\Omega)$ such that $T\phi_n = \mu_n \phi_n$ and $\mu_1 \ge \mu_2 \ge \mu_3 \ge \dots$ and $\lim \mu_i = 0$. Since $T\phi_n = \mu_n \phi_n$ iff $(L+C)^{-1}\phi_n = \mu_n \phi_n \in D(L)$ iff $\mu_n (L+C)\phi_n = \phi_n$ iff

$$L\phi_n = \frac{1}{\mu_n}(1 - C\mu_n)\phi_n = \lambda_n \phi_n$$

where $\lambda_n := \left(\frac{1}{\mu_n} - C\right) \uparrow \infty$ as $n \to \infty$. ∎

Corollary 55.2. *Let L be as above, then*

$$D(L) = \{u \in L^2(\Omega) : \sum_{n=1}^\infty \lambda_n^2 (u, \phi_n)^2 < \infty\}.$$

Moreover $Lu = \sum_{n=1}^\infty \lambda_n (u, \phi_n)\phi_n$ for all $u \in D(L)$, i.e. L is unitarily equivalent to the operator $\Lambda : \ell^2 \to \ell^2$ defined by $(\Lambda x)_n = \lambda_n x_n$ for all $n \in \mathbb{N}$.

Proof. Suppose $u \in D(L)$, then

$$Lu = \sum (Lu, \phi_n)\phi_n = \sum (u, L\phi_n)\phi_n = -\sum \lambda_n (u, \phi_n)\phi_n$$

with the above sums being L^2 convergent and hence

$$\sum \lambda_n^2 (u, \phi_n)^2 = \|Lu\|_{L^2}^2 < \infty.$$

Conversely if $\sum \lambda_n^2 (u, \phi_n)^2 < \infty$, let

$$u_N := \sum_{n=1}^N (u, \phi_n)\phi_n \in D(L).$$

Then $u_N \to u$ in $L^2(\Omega)$ and

$$Lu_N = \sum_1^N (u, \phi_n)\lambda_n \phi_n \to \sum_1^\infty (u, \phi_n)\lambda_n \phi_n \text{ in } L^2(\Omega).$$

Since L is a closed operator, $u \in D(L)$ and $Lu = \sum_1^\infty (u, \phi_n)\lambda_n \phi_n$. ∎

55.1 Growth of Eigenvalues I

Example 55.3. Let $\Omega = (0, \ell)$.

1. Suppose $L = -\frac{d^2}{dx^2}$ with $D(L) = H^2(\Omega) \cap H_0^1(\Omega)$, i.e. we impose Dirichlet boundary conditions. Because $L = L^*$ and $L \ge 0$ (in fact $L \ge \epsilon I$ for some $\epsilon > 0$ by the by the Poincaré Lemma in Theorem 49.31) if $Lu = \lambda u$ then $\lambda > 0$. Let $\lambda = \omega^2 > 0$, then the general solution to $Lu = \omega^2 u$ is given by

$$u(x) = A\cos(\omega x) + B\sin(\omega x)$$

where $A, B \in \mathbb{C}$. Because we want $u(0) = 0 = u(\ell)$ we must require $A = 0$ and $\ell\omega = n\pi$. Hence we have $\lambda_n = \frac{n^2\pi^2}{\ell^2}$ and $u_n(x) = \sqrt{\frac{2}{\ell}} \sin\left(\frac{n\pi}{\ell}x\right)$ for $n \in \mathbb{N}$ is an orthonormal basis of eigenvectors for L.

2. The reader is invited to show that if $L = \frac{-d^2}{dx^2}$ with Neuwmann boundary conditions then $u_0(x) := \frac{1}{\sqrt{\ell}}$ and $u_n(x) = \sqrt{\frac{2}{\ell}} \cos\left(\frac{n\pi}{\ell} x\right)$ for $n \in \mathbb{N}$ forms an orthonormal basis of eigenfunctions of L with eigenvalues given by $\lambda_n = \frac{n^2\pi^2}{\ell^2}$ for $n \in \mathbb{N}_0$.
3. Suppose that $L = -\Delta$ on Ω^d with Dirichlet boundary conditions and for $m \in \mathbb{N}^d$ let
$$U_m(x) = u_{m_1}(x_1)\dots u_{m_d}(x_d)$$
where each u_i is given as in Item 1. Then $\{U_m : m \in \mathbb{N}^d\}$ is an orthonormal basis of eigenfunctions of L with eigenvalues given
$$\lambda_m = \frac{\pi^2}{\ell^2}\sum_{i=1}^d m_i^2 = \frac{\pi^2}{\ell^2}|m|^2_{\mathbb{R}^d} \text{ for all } m \in \mathbb{N}^d.$$

Remark 55.4. Keeping the notation of item 3. of Example 55.3, for $\lambda > 0$ let
$$E_\lambda = \operatorname{span}\{\phi_m : \lambda_m \le \lambda\}.$$
Then
$$\dim(E_\lambda) = \#\left\{m \in \mathbb{N}^d : \lambda_m \le \lambda\right\} = \#\left\{m \in \mathbb{N}^d : |m|^2_{\mathbb{R}^d} \le \frac{\lambda\ell^2}{\pi^2}\right\}$$
from which it follows that
$$\dim(E_\lambda) \asymp m_d\left(B\left(0, \sqrt{\frac{\lambda\ell^2}{\pi^2}}\right)\right) = \omega_d\left(\frac{\lambda\ell^2}{\pi^2}\right)^{d/2} = Cm_d\left(\Omega^\ell\right)\lambda^{d/2}.$$

Lemma 55.5. *Let Ω, $\mathcal{E}$ and L be as described at the beginning of this section. Given $k \in \mathbb{N}$, there exists $C = C_k < \infty$ such that*
$$\|u\|_{H^{2k}(\Omega)} \le C(\|u\|_{L^2(\Omega)} + \left\|L^k u\right\|_{L^2(\Omega)}) \text{ for all } u \in D(L^k). \tag{55.1}$$

Proof. We first claim that for $u \in D(L^k)$.
$$\|u\|_{H^{2k}(\Omega)} \le C_k(\|u\|_{L^2(\Omega)} + \|Lu\|_{L^2(\Omega)} + \dots + \|L^k u\|_{L^2(\Omega)}). \tag{55.2}$$
We prove Eq. (55.2) by induction. When $k = 0$, Eq. (55.2) is trivial. Consider $u \in D(L^{(k+1)}) \subset D(L^k)$. By elliptic regularity (Theorem 52.15) and then using the induction hypothesis,
$$\|u\|_{H^{2(k+1)}(\Omega)} \le C(\|u\|_{L^2(\Omega)} + \|Lu\|_{H^k(\Omega)}) \le C_{k+1}\left(\sum_{j=0}^{k+1}\|L^j u\|_{L^2(\Omega)}\right).$$
This proves Eq. (55.2). Because $\sup\left\{\frac{\lambda^{2m}}{(1+\lambda^{2p})} : \lambda \ge 0\right\}$ for any $p > m$, it follows by the spectral theorem that
$$\|L^m u\|^2_{L^2(\Omega)} \le C\left(\|u\|^2_{L^2(\Omega)} + \|L^p u\|^2_{L^2(\Omega)}\right) \text{ for all } p > m.$$
Combining this fact with Eq. (55.2) implies Eq. (55.1). ■

Theorem 55.6. *Continue the notation in Lemma 55.5 and let $\{\phi_n\}_{n=1}^\infty$ be the orthonormal basis described in Theorem 55.1 and for $\lambda \in \mathbb{R}$ let*
$$E_\lambda := \operatorname{span}\{\phi_n : \lambda_n \le \lambda\}.$$
If k is the smallest integer such that $k > d/4$, there exist $C < \infty$ such that
$$\dim(E_\lambda) \le C(1 + \lambda^{2k}) \tag{55.3}$$
for all $\lambda \ge \inf\sigma(L)$.

Proof. By the Sobolev embedding Theorem 49.18, $H^{2k}(\Omega) \hookrightarrow C^{2k-\frac{d}{2}}(\Omega) \subset C\left(\bar{\Omega}\right)$. Combining this with Lemma 55.5 implies $D\left(L^k\right) \hookrightarrow C\left(\bar{\Omega}\right)$ and for $u \in D(L^k)$,
$$\|u\|_{C^0(\Omega)} \le \|u\|_{C^{2k-\frac{d}{2}}(\Omega)} \le C\|u\|_{H^{2k}(\Omega)} \le C(\|u\|_{L^2(\Omega)} + \|L^k u\|_{L^2(\Omega)}). \tag{55.4}$$
Let $\lambda \ge \inf\sigma(L)$ and $u \in E_\lambda \subset D(L^k)$. Since
$$u = \sum_{n:\lambda_n\le\lambda}(u,\phi_n)\phi_n \text{ and } L^k u = \sum_{\lambda_n\le\lambda}\lambda_n^k(u,\phi_n)\phi_n,$$
$$\|L^k u\|^2_{L^2(\Omega)} = \sum_{\lambda_n\le\lambda}|\lambda_n|^{2k}|(u,\phi_n)|^2 \le |\lambda|^{2k}\|u\|^2_{L^2(\Omega)}. \tag{55.5}$$
Combining Eqs. (55.4) and (55.5) implies
$$\|u\|_{C^0} \le C(1+\lambda^k)\|u\|_{L^2(\Omega)} = C(1+\lambda^k)\sqrt{\sum_{\lambda_n\le\lambda}|(u,\phi_n)|^2}. \tag{55.6}$$

Let $N = \dim(E_\lambda)$, $y \in \Omega$ and take $u(x) := \sum_{n=1}^N \phi_n(y)\phi_n(x)$ in Eq. (55.6) to find
$$\sum_{n=1}^N |\phi_n(y)|^2 \le \sup_{x\in\Omega}\left|\sum_{n=1}^N \phi_n(y)\phi_n(x)\right| \le C(1+\lambda^k)\sqrt{\sum|\phi_n(y)|^2}$$
from which it follows that
$$\sum_{n=1}^N |\phi_n(y)|^2 \le C^2(1+\lambda^k)^2.$$
Integrating this estimate over $y \in \Omega$ then shows
$$\dim(E_\lambda) = \sum_1^N 1 = \sum_1^N \int_\Omega |\phi_n(y)|^2\,dy \le C^2(1+\lambda^k)^2|\Omega|$$
which implies Eq. (55.3). ■

Corollary 55.7. *Let k be the smallest integer larger than $d/4$. Then there exists $\epsilon > 0$ such that $\lambda_n \geq \epsilon n^{1/2k}$ for n sufficiently large. Noting that $k \sim d/4$ this says roughly that $\lambda_n \sim n^{d/2}$, which is the correct result.*

Proof. Since[1]

$$n \leq \dim(E_{\lambda_n}) \leq c(1 + \lambda_n^{2k}),$$

$\frac{n}{c} - 1 \leq \lambda_n^{2k}$ or $\lambda_n \geq \left(\frac{n}{c} - 1\right)^{\frac{1}{2k}}$. ■

[1] If $\lambda = \lambda_n$ has multiplicity larger than one, then $n < \dim E_{\lambda_n}$ otherwise $n = \dim E_{\lambda_n}$.

Part XVI

Heat Kernel Properties

56

Construction of Heat Kernels by Spectral Methods

A couple of references for this and later sections are Davies [3, 4] and L. Saloff-Coste [12].

For this section, again let Ω be a bounded open subset of $\mathbb{R}^d$ such that $\bar{\Omega}^o = \Omega$ and $\bar{\Omega}$ is C^∞ – manifold with boundary. Also let $\mathcal{E}$ be a symmetric Dirichlet form with domain being either $X = H^1(\Omega)$ or $X = H_0^1(\Omega)$ and let $L := L_{\mathcal{E}}$ be the corresponding self adjoint operator. Let $\{\phi_n\}_{n=1}^\infty$ be the orthonormal basis of eigenvectors of L as described in Theorem 55.1 and $\{\lambda_n\}_{n=1}^\infty$ denote the corresponding eigenvalues, i.e. $L\phi_n = \lambda_n\phi_n$.

As we have seen abstractly before,

$$\begin{aligned} u(t) &= \cos(\sqrt{L}t)f + \frac{\sin(\sqrt{L}t)}{\sqrt{L}}\, g \\ &:= \sum_{n=1}^\infty \{\cos(\sqrt{\lambda_n}t)(f,\phi_n) + \frac{\sin(\sqrt{\lambda_n}t)}{\sqrt{\lambda_n}}\,(g,\phi_n)\}\phi_n \end{aligned}$$

solves the wave equation

$$\frac{\partial^2 u}{\partial t^2} + Lu = 0 \text{ with } u(0,x) = f(x) \text{ and } \dot{u}(0,x) = g(x)$$

and

$$u(t) = e^{-tL}u_0 := \sum_{n=1}^\infty e^{-t\lambda_n}(u_0,\phi_n)\phi_n$$

solves the heat equation,

$$\frac{\partial u}{\partial t} = -Lu \text{ with } u(0,x) = u_0(x). \tag{56.1}$$

Here we will concentrate on some of the properties of the solutions to the heat equation (56.1). Let us begin by writing out $u(t,x)$ more explicitly as

$$u(t,x) = \sum_{n=1}^\infty e^{-t\lambda_n}\int_\Omega u_0(y)\phi_n(y)dy\,\phi_n(x) = \lim_{N\to\infty}\int_\Omega u_0(y)\sum_{n=1}^N e^{-t\lambda_n}\phi_n(x)\phi_n(y)\,d \tag{56.2}$$

Theorem 56.1. *Let $p_t(x,y)$ denote the **heat kernel** associated to L defined by*

$$p_t(x,y) := \sum_1^\infty e^{-t\lambda_n}\phi_n(x)\phi_n(y). \tag{56.3}$$

Then

1. *the sum in Eq. (56.3) is uniformly convergent for all $t > 0$*
2. *$(t,x,y) \to p_t(x,y) \in C^\infty(\mathbb{R}^+ \times \overline{\Omega} \times \overline{\Omega})$*
3. *$u(t,x) = \int_\Omega p_t(x,y)u_0(y)dy$ solves Eq. (56.1).*

Proof. Let

$$p_t^N(x,y) = \sum_{n=1}^N e^{-t\lambda_n}\phi_n(x)\phi_n(y),$$

then $(t,x,y) \to p_t^N(x,y) \in C^\infty(\mathbb{R} \times \overline{\Omega} \times \overline{\Omega})$. Since $L^k\phi_n(x) = \lambda_n^k\phi_n$, by the Elliptic regularity Theorem 52.15,

$$\|\phi_n\|_{H^{2k}(\Omega)} \le C(\|\phi_n\|_{L^2(\Omega)} + \|L^k\phi_n\|_{L^2(\Omega)}) \le C(1+\lambda_n^k).$$

Taking $k > d/4$, the Sobolev embedding Theorem 49.18 implies

$$\|\phi_n\|_{C^0(\bar{\Omega})} \le \|\phi_n\|_{C^{2k-d/2}(\bar{\Omega})} \le C\|\phi_n\|_{H^{2k}(\Omega)} \le C(1+\lambda_n^k).$$

Therefore $\sup_{x,y\in\Omega}|\phi_n(x)\phi_n(y)| \le C^2(1+\lambda_n^k)^2$ while by Corollary 55.7, $\lambda_n \ge n^{2/d}$ and therefore while $\sum_{n=1}^\infty e^{-t\lambda_n}(1+\lambda_n^k)^2 < \infty$. More generally if $|\alpha| = 2m$

$$\|\partial^\alpha\phi_n\|_{C^0} \le \|\partial^\alpha\phi_n\|_{H^{2k}} \le \|\phi_n\|_{H^{2(k+m)}} \le C(1+\lambda_n^{k+m})$$

and hence

$$\|\phi_n\otimes\phi_n\|_{C^{2m}(\bar{\Omega}\times\Omega)} \le C^2(1+\lambda_n^{k+m})^2$$

from which it follows that

$$\begin{aligned} \sum_{n=1}^\infty \sup_{t\ge\epsilon} e^{-t\lambda_n}\|\phi_n\otimes\phi_n\|_{C^{2m}(\bar{\Omega}\times\Omega)} &= \sum_{n=1}^\infty e^{-\epsilon\lambda_n}\|\phi_n\otimes\phi_n\|_{C^{2m}(\bar{\Omega}\times\Omega)} \\ &\le C^2\sum_{n=1}^\infty e^{-\epsilon\lambda_n}(1+\lambda_n^{k+m})^2 < \infty. \end{aligned}$$

So $p_t^N(x,y)$ and all of its derivatives converge uniformly in $t \ge \epsilon$ and $x,y \in \bar{\Omega}$ as $N \to \infty$. Therefore $p_t(x,y) := \lim_{N\to\infty} p_t^N(x,y)$ exists and $(t,x,y) \to p_t(x,y)$

is $C^\infty(\bar{\Omega})$ for $t>0$ and $x,y\in\bar{\Omega}$. It is now easy to justify passing the limit under the integral sign in Equation (56.2) to find $u(t,x)=\int_\Omega p_t(x,y)u_0(y)dy$. ■

Remark 56.2. $p_t(x,y)$ solves the following problem $\frac{\partial p_t}{\partial t}=-L_x p_t$, $p_t(\cdot,y)$ satisfies the boundary conditions and $\lim_{t\downarrow 0} p_t(\cdot,y)=\delta_y$.

Definition 56.3. *A bounded operator $T:L^2(\Omega)\to L^2(\Omega)$ is positivity preserving if for every $f\in L^2(\Omega)$ with $f\ge 0$ a.e. on Ω has the property that $Tf\ge 0$ a.e. on Ω.*

Proposition 56.4 (Positivity of heat kernel's). *Suppose $D(L)=H_0^1(\Omega)\cap H^2(\Omega)$, i.e. L has Dirichlet boundary conditions, then the operator e^{-tL} is positivity preserving for all $t>0$ and the associated heat kernel $p_t(x,y)$ is non-negative for all $t\in(0,\infty)$ and $x,y\in\Omega$.*

Proof. Since $e^{-t(L+C)}=e^{-tC}e^{-tL}$ is positivity preserving iff e^{-tL} is positivity preserving, we may assume

$$L=-\sum a_{ij}\partial_i\partial_j+\sum a_i\partial_i+a$$

with $a>0$. Let $f\in C^\infty(\bar{\Omega},(0,\infty))$ and $u(t,x):=e^{-tL}f(x)$, in which case u solves

$$\frac{\partial u}{\partial t}=-Lu \text{ with } u(0,x)=f(x)\ge 0 \text{ for } (t,x)\in[0,T]\times\Omega$$
$$\text{with } u(t,x)=0 \text{ for } x\in\partial\Omega.$$

If there exist $(t_0,x_0)\in(0,T]\times\Omega$ such that

$$u(t_0,x_0)=\min\left\{u(t,x):0\le t\le T, x\in\bar{\Omega}\right\}<0,$$

then $\frac{\partial u}{\partial t}(t_0,x_0)\le 0$, $\partial_i u(t_0,x_0)=0$ for all i and by ellipticity,

$$a_{ij}(x_0)\partial_i\partial_j u(t_0,x_0)\ge 0.$$

Therefore at (t_0,x_0),

$$\begin{aligned}0&=\frac{\partial u}{\partial t}+Lu=\frac{\partial u}{\partial t}-a_{ij}\partial_i\partial_j u+a_i\partial_i u+au\\&=\frac{\partial u}{\partial t}-a_{ij}\partial_i\partial_j u+au=(\le 0)-(\ge 0)+(<0)<0\end{aligned}$$

which is a contradiction. Hence we have shown

$$0\le u(t,x)=\int_\Omega p_t(x,y)f(y)dy \text{ for all } (t,x)\in[0,T]\times\bar{\Omega}. \tag{56.4}$$

By a simple limiting argument, Eq. (56.4) also holds for all non-negative bounded measurable functions f on Ω. Indeed, let $f_n:=\left(f1_\Omega+n^{-1}\right)*\eta_n$ where $\eta_n\in C_c^\infty(\mathbb{R}^d,(0,\infty))$ is a spherically symmetric approximate δ-sequence. Then $f_n\in C^\infty(\bar{\Omega},(0,\infty))$ and hence

$$0\le\int_\Omega p_t(x,y)f_n(y)dy\to\int_\Omega p_t(x,y)f(y)dy.$$

From this equation it follows that $p_t(x,y)\ge 0$ and that e^{-tL} is positivity preserving. ■

Lemma 56.5. *Suppose $f_n\ge 0$ on Ω and $f_n\to\delta_x$ as $n\to\infty$, then $\lim_{n\to\infty}\int_\Omega f_n^2(x)dx=\infty$.*

Proof. For sake of contradiction assume $\lim_{n\to\infty}\int f_n^2(x)dx\ne\infty$. By passing to a subsequence if necessary we may then assmume $M=:\sup_n\int_\Omega f_n^2(x)dx<\infty$ and that f_n converges weakly to some $g\in L^2(\Omega)$. In which case we would have

$$\phi(x)=\lim_{n\to\infty}\int_\Omega f_n(y)\phi(y)dy=\int_\Omega g(y)\phi(y)dy \text{ for all } \phi\in C_c^\infty(\Omega).$$

But this would imply that $g=\delta_x$ which is incomensurate with g being an $L^2(\Omega)$ function and we have reached the desired contradiction. ■

Theorem 56.6. *Let p_t be a Dirichlet heat kernel, then $\lim_{t\downarrow 0}p_t(x,x)=\infty$ and $p_t(x,y)>0$ for all $x,y\in\Omega$ and $t>0$.*

Proof. We have seen $p_t(x,\cdot)\to\delta_x$ for all $x\in\Omega$. Therefore by lemma

$$\lim_{t\downarrow 0}\int_\Omega p_t(x,y)^2dy=\infty.$$

Now

$$\int_\Omega p_t(x,y)^2dy=\int_\Omega p_t(x,y)p_t(y,x)dy=p_{2t}(x,x).$$

Therefore $\lim_{t\downarrow 0}p_{2t}(x,x)=\infty$ for all $x\in\Omega$.

Sketch of the rest: Choose a compact set K in Ω, then by continuity there exists $t_0>0$ and $\epsilon>0$ such that $p_t(x,y)\ge 1$ for all $x\in K$, $|x-y|<\epsilon$ and $0<t\le t_0$. (Note $\tanh p_t(x,x)$ is continuous on $[0,1]\times\Omega$, where $\tanh p_0(x,x)=1$.) Now if $x\in K$ and $y\in K$ and $|x-y|<2\epsilon$, we have for $t<t_0$ that

$$\begin{aligned}p_t(x,y)&=\int_\Omega p_{t/2}(x,z)p_{t/2}(z,y)dz\ge\int_{B(x,\epsilon)\cap B(y,\epsilon)\cap\Omega}p_{t/2}(x,z)p_{t/2}(z,y)dz\\&\ge m\left(B(x,\epsilon)\cap B(y,\epsilon)\cap\Omega\right)>0.\end{aligned}$$

Working inductively, we may use this same idea to prove that $p_t(x,y) > 0$ for all $t < t_0$ and $x, y \in K$. Moreover the same semi-group argument allows one to show $p_t(x,y) > 0$ for all $t > 0$ as well.

Second Proof using the strong maximum principle.

Now for y fixed $0 < t \leq T$, $f(t,x) = p_t(x,y)$ solves $\frac{\partial f}{\partial t} = L_x f(t,x), f(\epsilon, x) = p_\epsilon(x,y)$. With out loss of generality, assume

$$L = -\sum -\partial_i a_{ij} \partial_j + a \text{ with } a \geq 0.$$

For if a is not greater than 0, replace L by $L+\lambda$ and observe that $e^{-t(L+\lambda)} = e^{-t\lambda} e^{tL}$. Therefore $p_t^\lambda(x,y) = e^{-t\lambda} p_t(x,y)$ so $p_t(x,y) > 0$ iff $p_t^\lambda(x,y) \geq 0$. Then $\frac{\partial p_t}{\partial t} = L_x p_t(x,y)$ for all $T \geq t \geq \epsilon$ and $x \in \Omega$. By the strong maximum principle of Theorem 12 on page 340 if there exist $(x_0, t_0) \in \Omega \times (\epsilon, T]$ such that $p_{t_0}(x_0, y) = 0$ then $x \to p_t(x,y)$ is a constant on $(0 t_0) \times \Omega$ which is false because the constant would have to be 0, but $\int p_{t_0}(x,y)dx > 0$ for T small. ■

56.1 Positivity of Dirichlet Heat Kernel by Beurling Deny Methods

Assumption 6 *Suppose* $L = -\partial_i a_{ij} \partial_j + a$ *where* $a_{ij} = a_{ji} \geq \epsilon I$ *and* $a \in C^\infty(\overline{\Omega})$ *and* $D(L) = H_0^1(\Omega) \cap H^2(\Omega)$.

Theorem 56.7. *Let* $\lambda_0 = \max(-a)$ *i.e.* $-\lambda_0 = \min(a)$. *Then for all* $\lambda \geq \lambda_0$, $L_\lambda := L + \lambda I : D(L) \to L^2(\Omega)$ *invertible and if* $f \in L^2(\Omega)$, $f \geq 0$ *a.e then* $L_\lambda^{-1} f \geq 0$ *a.e., i.e.* L_λ^{-1} *is positivity preserving.*

Proof. If $\lambda \geq \lambda_0$ then $\lambda + a \geq 0$ and hence if $L_\lambda u = 0$ then

$$(L_\lambda u, u) = \int_\Omega (a_{ij} \partial_i u \partial_j u + (a+\lambda) u^2) dx = 0.$$

This implies $\nabla u = 0$ a.e. and so u is constant and hence $u = 0$ because $u \in H_0^1$. Therefore $\mathrm{Nul}(L_\lambda) = \{0\}$ and so L_λ is invertible by the Fredholm alternative. Now suppose $f \in C^\infty(\overline{\Omega})$ such that $f > 0$, then $u = L_\lambda^{-1} f \in C^\infty(\overline{\Omega})$ with $L_\lambda u = f > 0$ and $u = 0$ on $\partial\Omega$. We may now use the maximum principle idea in Theorem 45.16 to conclude that $u \geq 0$. Indeed if there exists $x_0 \in \Omega$ such that $u(x_0) = \min u < 0$ at x_0, then

$$0 < f(x_0) = (L_\lambda u)(x_0) = -\underbrace{(a_{ij} \partial_i \partial_j u)(x_0)}_{\geq 0} + \underbrace{(\partial_i a_{ij}) \partial_j u(x_0)}_{=0} + \underbrace{(a+\lambda) u(x_0)}_{\leq 0} \leq 0$$

which is a contradiction. Thus we have shown $u = L_\lambda^{-1} f \geq 0$ if $f \in C^\infty(\overline{\Omega}, (0,\infty))$. Given $f \in L^2(\Omega)$ such that $f \geq 0$ a.e. on $\overline{\Omega}$, choose $f_n \in C^\infty(\overline{\Omega})$ such that $f_n \geq \frac{1}{n}$ and $f_n \to f$ in $L^2(\Omega)$ and $f_n \to f$ a.e. on $\overline{\Omega}$. (For example, take $f_n = \eta_{\delta_n} * \left(f 1_{\overline{\Omega}} + \frac{1}{n}\right)$ say.) Then $u_n = L_\lambda^{-1} f_n \geq 0$ for all n and $u_n \to u = L_\lambda^{-1} f$ in $H^2(\Omega)$. By passing to a subsequence if necessary we may assume that $u_n \to u$ a.e. from which it follows that $u \geq 0$ a.e. on Ω. ■

Theorem 56.8. *Keeping* L *as above,* $e^{-tL} : L^2(\Omega) \to L^2(\Omega)$ *is positivity preserving for all* $t \geq 0$.

Proof. By the spectral theorem and the fact that $\left(1 + \frac{t\lambda}{n}\right)^{-1n} \to e^{-t\lambda}$ boundedly for and $\lambda \geq 0$,

$$e^{-tL} f = \lim_{n\to\infty} \left(1 + \frac{tL}{n}\right)^{-n} f = \lim_{n\to\infty} \left(\frac{t}{n}\right)^{-n} \left[\left(\frac{n}{t} + L\right)^{-1}\right]^n f.$$

Now $\left(\frac{n}{t} + L\right)^{-1}$ is positivity preserving operator on $L^2(\Omega)$. Where $\frac{n}{t} \geq \lambda_0$ and hence so is the n – fold product. Thus if

$$u_n := \left(\frac{t}{n}\right)^{-n} \left[\left(\frac{n}{t} + L\right)^{-1}\right]^n f$$

then $u_n \geq 0$ a.e. and $u_n \to e^{-tL} f$ in $L^2(\Omega)$ implies $e^{-tL} f \geq 0$ a.e. ■

Theorem 56.9. $p_t(x,y) \geq 0$ *for all* $x, y \in \Omega$.

Proof. $f \in L^2(\Omega)$ with $f \geq 0$ a.e. on Ω, $e^{-tL} f \in C^\infty(\overline{\Omega})$ and $e^{-tL} f \geq 0$ a.e. by above. Thus $e^{-tL} f \geq 0$ everywhere. Now

$$\int_\Omega p_t(x,y) f(y) dy = (e^{-tL} f)(x) \geq 0$$

for all $f \geq 0$. Since $p_t(x,y)$ is smooth this implies $p_t(x,y) \geq 0$ for all $y \in \bar{\Omega}$ and since $x \in \bar{\Omega}$ was arbitrary we learn $p_t(x,y) \geq 0$ for all $x, y \in \overline{\Omega}$. ■

???? BRUCE for $f \in C_c^\infty(\Omega) \subset \cap D(L^n) = C^\infty(L)$ we have $\|(e^{-tL} f) - f\|_{H^k} \to 0$ as $t \downarrow 0$ for all k. By Sobolev embedding this implies that $(e^{-tL} f) \to f(x)$ as $t \downarrow 0$ for all $x \in \Omega$ i.e. $\int_\Omega p_t(x,y) f(y) dy \to f(x)$.

57

Nash Type Inequalities and Their Consequences

Corollary 57.1. *Suppose $d > 2$, then there is a constant $C_d < \infty$ such that*

$$\|u\|_2^{2+4/d} \le C_d \|\nabla u\|_2^2 \; \|u\|_1^{4/d} \tag{57.1}$$

for all $u \in C_c^1\left(\mathbb{R}^d\right)$.

Proof. By Corollary 49.15, $\|u\|_{2^*} \le C\|\nabla u\|_2$ where $2^* = \frac{2d}{d-2}$ and by interpolation

$$\|u\|_2 \le \|u\|_p^\theta \; \|u\|_q^{1-\theta}$$

where $\frac{\theta}{p} + \frac{1-\theta}{q} = \frac{1}{2}$. Taking $p = 2^*$ and $q = 1$ implies $\frac{\theta}{2^*} + 1 - \theta = \frac{1}{2}$, i.e. $\theta\left(\frac{1}{2^*} - 1\right) = -\frac{1}{2}$ and hence

$$\theta = \frac{\frac{1}{2}}{1 - \frac{1}{2^*}} = \frac{2^*}{2(2^* - 1)} = \frac{d}{(d-2)} \cdot \frac{1}{\frac{2d}{d-2} - 1}$$
$$= \frac{d}{d-2} \frac{d-2}{d+2} = \frac{d}{d+2}$$

and $1 - \theta = \frac{2}{d+2}$. Hence

$$\|u\|_2 \le \|u\|_{2^*}^{\frac{d}{d+2}} \; \|u\|_1^{\frac{2}{d+2}} \le C^{\frac{d}{d+2}} \; \|\nabla u\|_2^{\frac{d}{d=2}} \; \|u\|_1^{\frac{2}{d+2}}$$

and therefore

$$\|u\|_2^{\frac{d+2}{d}} \le C\|\nabla u\|_2 \; \|u\|_1^{\frac{2}{d}}.$$

and squaring this equation then gives the estimate in Eq. (57.1). ■

Proposition 57.2 (Nash). *Corollary 57.1 holds for all d.*

Proof. Since the Fourier transform is unitary, for any $R > 0$ and $\|\hat{u}\|_\infty \le \|u\|_{L^1}$,

$$\begin{aligned}\|u\|_2^2 &= \int_{\mathbb{R}^d} |\hat{u}(\xi)|^2 d\xi = \int_{|\xi|\le R} |\hat{u}|^2 d\xi + \int_{|\xi|>R} |\hat{u}|^2 d\xi \\ &\le \sigma\left(S^{d-1}\right) R^d \|u\|_{L^1}^2 + \frac{1}{R^2} \int_{|\xi|>R} |\xi|^2 |\hat{u}|^2 d\xi \\ &\le \sigma\left(S^{d-1}\right) R^d \|u\|_{L^1}^2 + \frac{1}{R^2} \|Du\|_2^2 = f(R)\end{aligned}$$

where $f(R) = aR^d + \frac{b}{R^2}$ and $a = \sigma\left(S^{d-1}\right)\|u\|_{L^1}^2$ and $b = \|Du\|_2^2$. To minimize f, we set $f'(R) = 0$ to find $daR^{d-1} - 2bR^{-3} = 0$, i.e. $R^{d+2} = \frac{2b}{da}$ and hence $R = \left(\frac{2b}{da}\right)^{\frac{1}{d+2}}$. With this value of R, we find

$$\begin{aligned}f(R) &= \frac{1}{R^2}\left(b + aR^{d+2}\right) = \left(\frac{da}{2b}\right)^{\frac{2}{d+2}} \left(b + \frac{2b}{d}\right) \\ &= \left(\frac{d+2}{d}\right) b \left(\frac{da}{2b}\right)^{\frac{2}{d+2}} = C_d \; a^{\frac{2}{d+2}} b^{1-\frac{2}{d+2}} = C_d \; a^{\frac{2}{d+2}} \; b^{\frac{d}{d+2}}\end{aligned}$$

which gives the estimate

$$\|u\|_2^2 \le C_d \|Du\|_2^{\frac{2d}{d+2}} \|u\|_1^{\frac{4}{d+2}}$$

which is equivalent to

$$\|u\|_2^{2+4/d} = \|u\|_2^{2\left(\frac{d+2}{d}\right)} \le C_d \|Du\|_2^2 \; \|u\|_1^{4/d}.$$

■

Proposition 57.3. *Suppose $A(x) = \{a_{ij}(x)\}_{i,j=1}^d$ such that there exists $\epsilon > 0$ and $M < \infty$ such that $\epsilon I \le A(x) \le MI$ for all $x \in \mathbb{R}^d$. Define $\mathcal{D}(\mathcal{E}) = W^{1,2}(\mathbb{R}^d)$ and*

$$\mathcal{E}(u, v) = \sum_{ij} \int_{\mathbb{R}^d} a_{ij} D_i u D_i v \; dx.$$

Then $\mathcal{E}$ is a closed symmetric quadratic form. Moreover $C_c^\infty(\mathbb{R}^d)$ is a ***core*** *for $\mathcal{E}$.*

Proof. Clearly $\|u\|_{W^{1,2}} \le \frac{1}{\epsilon}(\|u\|_{L^2(\Omega)}^2 + \mathcal{E}(u, u))$ and

$$\mathcal{E}(u, u) \le M\|Du\|^2 \le M\|u\|_{W^{1,2}}.$$

and hence

$$\|\cdot\|_2^2 + \mathcal{E}(\cdot, \cdot) \asymp \|\cdot\|_{W^{1,2}}$$

so that $\left(\mathcal{D}(\mathcal{E}), \sqrt{\|\cdot\|_2^2 + \mathcal{E}(\cdot, \cdot)}\right)$ is complete. ■

Theorem 57.4. *Let $-L$ denote the positive self-adjoint operator on $L^2(\mathbb{R}^d)$ such that $\mathcal{E}(u,v) = (\sqrt{-L}\ u, \sqrt{-L}\ v)_{L^2(\mathbb{R}^d)}$ and define $T_t := e^{tL} : L^2 \to L^2$. (Notice that $\|T_t f\|_2 \le \|f\|_2$ for all $f \in L^2$ and $t > 0$.) Then*

1. $T_t f \ge 0$ *a. e. if* $f \ge 0$ *a.e.*
2. $T_t f \in L^1 \cap L^\infty$ *for all* $f \in L^1 \cap L^\infty$
3. $0 \le T_t f \le 1$ *if* $0 \le f \le 1$ *a.e.*
4. $\|T_t f\|_{L^p} \le \|f\|_{L^p}$ *for all* $f \in L^1 \cap L^\infty$.

Proof. (Fake proof but the spirit is correct.) Let $u(t,x) = T_t f(x)$ so that

$$u_t = Lu \text{ with } u(0,x) = f(x)$$

where $Lf = \sum \partial_i (a_{ij} \partial_j f)(x)$ – a second order elliptic operator. Therefore by the "maximum principle,"

$$-\|f\|_\infty \ge \inf f(x) \le u(t,x) \le \sup f(x) \le \|f\|_\infty$$

and hence $\|f\|_\infty \ge \|T_t f\|_\infty$. This implies items 1. and 3. of the theorem.

For $g, f \in L^1 \cap L^\infty$,

$$|(T_t f, g)| = |(f, T_t g)| \le \|f\|_1 \|T_t g\|_\infty \le \|f\|_1 \|g\|_\infty.$$

Taking sup over $g \in L^1 \cap L^\infty$ such that $\|g\|_\infty = 1$ implies $\|T_t f\|_{L^1} \le \|f\|_1$ and we have verified that

$$\|T_t f\|_p \le \|f\|_p \text{ for } p \in \{1, 2, \infty\}.$$

Hence by the Riesz Thorin interpolation theorem, $\|T_t f\|_p \le \|f\|_p$ for all $p \in [1, \infty)$. ■

Theorem 57.5 (Beurling - Deny). *Items 1. – 4. of Theorem 57.4 hold if for all $u \in W^{1,2}$, $|u| \in W^{1,2}$ and $0 \vee (u \wedge 1) \in W^{1,2}$ and*

$$\mathcal{E}(|u|) \le \mathcal{E}(u) \text{ and } \mathcal{E}(0 \vee (u \wedge 1)) \le \mathcal{E}(u). \tag{57.2}$$

Proposition 57.6. *Suppose $u \in W^{1,p}(\Omega)$ then $1_{\{u=0\}}\ Du = 0$ a.e.*

Proof. Let $\phi \in C_c^\infty(\Omega)$ such that $\phi(0) = 1$. For ϵ small set $\phi_\epsilon(x) = \phi(x/\epsilon)$ and

$$\psi_\epsilon(x) = \int_{-\infty}^x \phi_\epsilon(y) dy = \int_{-\infty}^x \phi\left(\frac{y}{\epsilon}\right) dy$$
$$= \epsilon \int_{-\infty}^{x/\epsilon} \phi(u) du = \epsilon \psi_1(x/\epsilon).$$

Then

$$D[\psi_\epsilon(u)] = \psi_\epsilon'(u) Du = \phi_\epsilon(u) Du$$

and hence for all $f \in C_c^\infty(\Omega)$,

$$\langle \phi_\epsilon(u) Du, f\rangle = \langle D[\psi_\epsilon(u)], f\rangle = \langle \psi_\epsilon(u), -Df\rangle = 0(\epsilon) \to 0.$$

Combining this with the observation that

$$\phi_\epsilon(u) Du \xrightarrow{L^p} 1_{\{u=0\}} Du \text{ as } \epsilon \downarrow 0$$

implies

$$\int_\Omega 1_{\{u=0\}} Du \cdot f\ dm = 0 \text{ for all } f \in C_c^\infty(\Omega)$$

which proves $1_{\{u=0\}} Du = 0$ a.e. ■

Exercise 57.7. Let $u \in W^{1,p}$. Show

1. If $\phi \in C^1(\mathbb{R})$, $\phi(0) = 0$ and $|\phi'| \le M < \infty$, then $\phi(u) \in W^{1,p}$ and $D\phi(u) = \phi'(u) Du$ a.e.
2. $|u| \in W^{1,p}$ and $D|u| = \operatorname{sgn}(u) Du$.
3. Eq. (57.2) holds.

Solution 57.8. Let $u_n \in C^\infty(\mathbb{R}^d) \cap W^{1,p}$ such that $u_n \to u$ in $W^{1,p}$. By passing to a subsequence if necessary we may further assuem that $u_n(x) \to u(x)$ for a.e. $x \in \mathbb{R}^d$. Since $|\phi(u)| \le M|u|$

$$|\phi(u) - \phi(u_n)| \le M|u - u_n|$$

it follows that $|\phi(u)|, |\phi(u_n)| \in L^p$ and $\phi(u_n) \to \phi(u)$ in L^p. Since $\phi'(u_n)$ is bounded, $\phi'(u_n) \to \phi'(u)$ a.e. and $\partial_i u_n \to \partial_i u$ in L^p, it follows that

$$\begin{aligned}
&\|\partial_i \phi(u_n) - \partial_i \phi(u)\|_p \\
&= \|\phi'(u_n)\partial_i u_n - \phi'(u)\partial_i u\|_p \\
&\le \|\phi'(u_n)[\partial_i u_n - \partial_i u]\|_p + \|[\phi'(u) - \phi'(u_n)]\partial_i u\|_p \\
&\le M\|\partial_i u_n - \partial_i u\|_p + \|[\phi'(u) - \phi'(u_n)]\partial_i u\|_p \to 0
\end{aligned}$$

where the second term is handled by the dominated convergence theorem. Therefore $\phi(u) \in W^{1,p}$ and $\partial_i \phi(u) = \phi'(u)\partial_i u$.

Let $\phi_\epsilon(x) := \sqrt{x^2 + \epsilon^2}$, then $\phi_\epsilon'(x) = \frac{x}{\sqrt{x^2+\epsilon^2}}$, $|\phi_\epsilon'(x)| \le 1$ for all x and

$$\lim_{\epsilon \downarrow 0} \phi_\epsilon'(x) = \begin{cases} 0 & \text{if } x = 0 \\ \operatorname{sgn}(x) & \text{if } x \ne 0. \end{cases}$$

From part 1., $\phi_\epsilon(u) \in W^{1,p}$ and $\partial_i \phi_\epsilon(u) = \phi_\epsilon'(u)\partial_i u$. Since $\phi_\epsilon(u) \to |u|$ in L^p and

$$\partial_i \phi_\epsilon(u) = \phi_\epsilon'(u)\partial_i u \to 1_{u \ne 0}\operatorname{sgn}(u)\partial_i u = \operatorname{sgn}(u)\partial_i u \text{ a.e.}$$

where the last equality is a consequence of Proposition 57.6. Hence we see that $|u| \in W^{1,p}$ and $D|u| = \operatorname{sgn}(u) Du$.

Remark: (BRUCE) I think using the absolute continuity of u along lines could be used to simplify and generalize the above exercise to the case where $\phi \in AC(\mathbb{R})$ with $|\phi'(x)| \le M < \infty$ for m – a.e. $x \in \mathbb{R}$.

Remark 57.9. T_t extends by continuity to L^p for all $1 \le p \le \infty$, denote the extension by T_t^p and then $\overline{T}_t f = T_t^p f$ if $f \in L^p$ for some p. In this way we view $\overline{T}_t$ as a linear operator on $\bigcup_{1\le p\le\infty} L^p$.

Theorem 57.10. *There is a constant $C < \infty$ such that*

$$\|T_t f\|_{L^\infty} \le \frac{C}{t^{d/4}} \|f\|_{L^2} \tag{57.3}$$

for all $f \in L^2$.

Proof. Ignoring certain technical details. Set $u(t) = T_t f$ and $v(t) := \|u(t)\|_2^2$ and recall that u solves

$$\dot{u} = Lu \text{ with } u(0) = f.$$

Then

$$\begin{aligned} -\dot{v}(t) &= -\frac{d}{dt} \|u(t)\|_2^2 = -2(u, \dot{u}) = -2(u, Lu) \\ &= 2\mathcal{E}(u,u) \ge \frac{2}{\epsilon}\|Du\|_{L^2}^2. \end{aligned}$$

Combining this with the Nash inequality from Eq. (57.1),

$$\|u\|_2^{2+\frac{d}{4}} \le C\|Du\|_{L^2}^2 \|u\|_1^{4/d},$$

implies

$$-\dot{v}(t) \ge \frac{2C}{\epsilon} \frac{\|u(t)\|_2^{2+4/d}}{\|u(t)\|_1^{4/d}} \ge \frac{2C}{\epsilon} \frac{\|u(t)\|_2^{2+4/d}}{\|f\|_1^{4/d}} = \frac{2C}{\epsilon} \frac{v(t)^{1+2/d}}{\|f\|_1^{4/d}}. \tag{57.4}$$

Since $\int \dot{v} v^{(-1-2/d)} dt = -\frac{d}{2} v^{-2/d}$, it Eq. (57.4) is equivalent to

$$\frac{d}{dt}\left(\frac{d}{2} v^{-2/d}\right) \ge \frac{2C}{\epsilon \|f\|_1^{4/d}}$$

and integrating this inequality gives

$$\|T_t f\|_2^{-4/d} = v^{-2/d}(t) \ge v^{-2/d}(t) - v^{-2/d}(0) \ge \frac{4Ct}{d\epsilon \|f\|_1^{4/d}}.$$

Some algebra then implies

$$\|T_t f\|_2^{4/d} \le \frac{d\epsilon \|f\|_1^{4/d}}{4Ct}$$

and hence

$$\|T_t f\|_2 \le \frac{C}{t^{d/4}} \|f\|_1$$

and by duality, Lemma 57.12 below, this implies Eq. (57.3). ∎

Remark 57.11. From Eq. (57.3),

$$\begin{aligned} \|e^{2tL} f\|_{L^\infty}^2 &= \|e^{tL} e^{tL} f\|_{L^\infty}^2 \le \frac{C}{t^{d/2}} \|e^{tL} f\|_{L^2}^2 = \frac{C}{t^{d/2}} |(e^{2tL} f, f)| \\ &\le \frac{C^2}{t^{d/2}} \|e^{2tL} f\|_{L^\infty} \|f\|_{L^1} \end{aligned}$$

and hence Eq. (57.3) implies the inequality,

$$\|e^{2tL} f\|_{L^\infty} \le \frac{C^2}{t^{d/2}} \|f\|_{L^1}. \tag{57.5}$$

Lemma 57.12 (Duality Lemma). *Let T be a linear operator on $\cup_{p\in[1,\infty]} L^p$ such that $T(L^1 \cap L^\infty) = (L^1 \cap L^\infty)$ and $T : L^2 \to L^2$ is self adjoint. If*

$$\|Tf\|_p \le C\|f\|_q \text{ for all } f \in L^1 \cap L^\infty$$

then

$$\|Tf\|_q \le C\|f\|_{p'} \text{ for all } f \in L^1 \cap L^\infty$$

where $\frac{1}{q'} + \frac{1}{q} = 1$ and $\frac{1}{p} + \frac{1}{p'} = 1$.

Proof.

$$\begin{aligned} \|Tf\|_{q'} &= \sup_{\|g\|_q = 1} |(Tf, g)| = \sup_{\|g\|_q=1} |(f, Tg)| \le \sup_{\|g\|_q=1} \|f\|_{p'} \|Tg\|_p \\ &\le C\|f\|_{p'} \sup_{\|g\|_q=1} \|g\|_q = C\|f\|_{p'}. \end{aligned}$$

∎

Proposition 57.13 (Converse of Theorem 57.10.). *If*

$$\left\|e^{tL} f\right\|_\infty \le Ct^{-d/4} \|f\|_2 \text{ for all } f \in L^2 \tag{57.6}$$

then

$$\|f\|_2^{2+4/d} \le \frac{4C}{d} \mathcal{E}(f,f) \|f\|_1.$$

Proof. By the duality, Lemma 57.12, (57.6) implies

$$\left\|e^{tL} f\right\|_2 \le Ct^{-d/4} \|f\|_1 \text{ for all } f \in L^2$$

and therefore

$$C^2 t^{-d/2} \|f\|_1^2 \ge \|e^{tL} f\|_2^2 = (e^{2tL} f, f) = (f,f) + \int_0^t 2(Le^{2\tau L} f, f) d\tau.$$

Since $xe^{2\tau x} \ge x$ for $x \le 0$, it follows from the spectral theorem that $Le^{2\tau L} \ge L$ for $L \le 0$. Using this in the above equation gives the estimate

$$C^2 t^{-d/2}\|f\|_1^2 \geq \|f\|_2^2 + 2\int_0^t (Lf,f)d\tau = \|f\|_2^2 + 2t(Lf,f)$$
$$\geq \|f\|_2^2 - 2t\mathcal{E}(f,f).$$

Optimizing this inequality over $t > 0$ by taking $t = \mathcal{E}(f,f)^{-2/d+2}\|f\|_1^{4/d+2}$ implies

$$\|f\|_2^2 \leq C^2 \left(\mathcal{E}(f,f)^{-2/d+2}\|f\|_1^{4/d+2}\right)^{-d/2}\|f\|_1^2 + 2\mathcal{E}(f,f)^{-2/d+2}\|f\|_1^{4/d+2}\mathcal{E}(f,f)$$
$$=$$

$$\|f\|_2^{2+4/d} \leq C\mathcal{E}(f,f)\|f\|_1^{4/d}.$$

■

58

T. Coulhon Lecture Notes

Notes from coulhon.tex.

Theorem 58.1. *Let (X, μ) be a measure space and f be a positive measurable function. Then for $1 \le p < \infty$,*

$$\|f\|_p^p = p \int_0^\infty \mu(f > t)\, t^{p-1} dt$$

Proof. We have

$$\|f\|_p^p = \int_X f^p d\mu = \int_X \left(\int_0^{f^p} p t^{p-1} dt \right) d\mu = \int_{X \times \mathbb{R}_+} 1_{f>t} p t^{p-1} d\mu dt$$
$$= \int_0^\infty \mu(f > t)\, p t^{p-1} dt$$

■

In these notes we are going to work in either one of the two following settings.

58.1 Weighted Riemannian Manifolds

Here we assume that M is a non-compact, connected Riemannian manifold with Riemannian metric g which is also equipped with a smooth measure μ. We let $|\nabla f|^2 = g(\nabla f, \nabla f)$ and

$$\mathcal{E}(f,f) := \int_M |\nabla f|^2 d\mu = (\Delta_\mu f, f) \text{ on } L^2(\mu).$$

Here $(f,g) := \int_M f\bar{g} d\mu$. We have the following general important facts. The heat kernel is the smooth integral kernel for the heat operator, $e^{-t\Delta_\mu}$.

Definition 58.2. *Given a smooth hypersurface $A \subset M$ let $|A|$ denote the surface measure, i.e.*

$$|A| = \int_A |\mu(N, -)|$$

where N is a normal vector to A.

Theorem 58.3 ($\|\nabla 1_\Omega\|_1 = |\partial\Omega|_s$)**.** *Let $\Omega \subset\subset M$ be a precompact domain with smooth boundary and let $f_\epsilon(x) = h_\epsilon(d_{\Omega^c}(x))$ where $h_\epsilon(x) = \left(\frac{x}{\epsilon}\right) \wedge 1$. Here $d_{\Omega^c}(x)$ denotes the Riemannian distance of x to Ω^c. Then*

$$\lim_{\epsilon \downarrow 0} \|\nabla f_\epsilon\|_1 = |\partial\Omega|_s$$

which we write heuristically as

$$\|\nabla 1_\Omega\|_1 = |\partial\Omega|_s .$$

Proof. We have

$$\nabla f_\epsilon(x) = h'_\epsilon(d_{\Omega^c}(x)) \nabla d_{\Omega^c}(x) \text{ for a.e. } x$$

and hence

$$|\nabla f_\epsilon(x)| = \frac{1}{\epsilon} 1_{d_{\Omega^c}(x) < \epsilon} \text{ for a.e. } x$$

and therefore,

$$\lim_{\epsilon \downarrow 0} \|\nabla f_\epsilon\|_1 = \lim_{\epsilon \downarrow 0} \frac{1}{\epsilon} \mu(d_{\Omega^c} < \epsilon) = |\partial\Omega| .$$

■

Theorem 58.4 (Coarea Formula). *Let (M, g) be a Riemannian manifold, $f : M \to [0, \infty)$ be a reasonable function and μ be a smooth volume form on M. Then*

$$\|\nabla f\|_{L^1(\mu)} = \int_0^\infty |\partial \{f > t\}| \, dt.$$

Proof. See [5, 1, 2] for a complete Rigorous proof. We will only give the idea here. Locally choose coordinates $x = (x^1, \dots, x^n)$ on M such that $x^1 = f$. This is possible in neighborhood of points where df is non-zero. Then $\frac{\partial}{\partial x^i}$ is tangential to the level surface $\{f = t\}$ for $i = 2, \dots, n$ and therefore

$$\mu(\frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, \dots, \frac{\partial}{\partial x^n}) = \mu(\left(\frac{\partial}{\partial x^1}, \frac{\nabla f}{|\nabla f|}\right) \frac{\nabla f}{|\nabla f|}, \frac{\partial}{\partial x^2}, \dots, \frac{\partial}{\partial x^n})$$
$$= \left(\frac{\partial}{\partial x^1}, \frac{\nabla f}{|\nabla f|}\right) \mu(N, \frac{\partial}{\partial x^2}, \dots, \frac{\partial}{\partial x^n}).$$

Now

$$\left(\frac{\partial}{\partial x^1}, \frac{\nabla f}{|\nabla f|}\right) = |\nabla f|^{-1}\left(\frac{\partial}{\partial x^1}, \nabla f\right)$$
$$= |\nabla f|^{-1}\frac{\partial}{\partial x^1}f = |\nabla f|^{-1}\frac{\partial}{\partial x^1}x^1 = |\nabla f|^{-1}$$

and therefore,

$$\left|\mu(N, \frac{\partial}{\partial x^2}, \dots, \frac{\partial}{\partial x^n})\right| = |\nabla f|\left|\mu(\frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, \dots, \frac{\partial}{\partial x^n})\right|.$$

Integrating this equation with respect to $dx^1 \dots dx^n$ then gives

$$\int \left|\mu(N, \frac{\partial}{\partial x^2}, \dots, \frac{\partial}{\partial x^n})\right| dx^1 \dots dx^n = \|\nabla f\|_{L^1(\mu)}$$

on one hand, on the other

$$\int \left|\mu(N, \frac{\partial}{\partial x^2}, \dots, \frac{\partial}{\partial x^n})\right| dx^1 \dots dx^n$$
$$= \int_{\mathbb{R}}\left[\int_{\mathbb{R}^{n-1}}\left|\mu(N, \frac{\partial}{\partial x^2}, \dots, \frac{\partial}{\partial x^n})\right| dx^2 \dots dx^n\right] dx^1$$
$$= \int_{\mathbb{R}}\left[\int_{\{f=x^1\}}|\mu(N, -)|\right] dx^1 = \int_0^\infty |\partial\{f > t\}|\, dt.$$

■

58.2 Graph Setting

Let $\Gamma = (V, E)$ be a non-oriented graph with vertices V and edges E. We assume the graph is connected and locally finite, in fact I think he assumes the graph is finitely ramified, i.e. there is a bound $K < \infty$ on the number of edges that are attached to any vertex. Let d denote the graph distance and

$$\mu := \{\mu_{yx} = \mu_{xy} \geq 0 : x, y \in V \ \ni\ xy \in E\}$$

be a measure on E. Extend μ to all pairs of points xy with $x, y \in V$ by setting $\mu_{xy} = 0$ if $xy \notin E$. Using this notation we let

$$\mu_x := \sum_{y \in V}\mu_{xy} = \sum_{y \in V: xy \in E}\mu_{xy}$$
$$p(x,y) := \frac{\mu_{xy}}{\mu_x} \text{ (the Markov Kernel)},$$
$$\mu(\Omega) := \sum_{x \in \Omega}\mu_x \text{ when } \Omega \subset\subset V,$$
$$\partial\Omega := \{xy \in E : x \in \Omega \text{ and } y \in \partial\Omega\},$$
$$|\partial\Omega|_s := \sum_{x \in \Omega, y \notin \Omega}\mu_{xy} \text{ and}$$
$$|\nabla f(x)|^p := \sum_y |f(x) - f(y)|^p\, p(x,y).$$

In this setting $Pf(x) := \sum_y p(x,y)f(y)$ corresponds to $e^{-t\Delta_\mu}$ and P is a self-adjoint operator on $L^2(V, \mu)$. Also recall that Δ_μ corresponds to $1 - P$. As in the manifold case we still have

$$\|\nabla 1_\Omega\|_1 = 2\,|\partial\Omega|_s \text{ and} \tag{58.1}$$
$$\int_0^\infty |\partial\{f > t\}|\, dt = \frac{1}{2}\|\nabla f\|_{L^1(\mu)} \tag{58.2}$$

as we will now verify. Using the definitions,

$$\|\nabla 1_\Omega\|_1 = \sum_x \mu_x \sum_y |1_\Omega(x) - 1_\Omega(y)|\, p(x,y) = \sum_{x,y}|1_\Omega(x) - 1_\Omega(y)|\,\mu_{xy}$$
$$= \sum_{x \in \Omega, y \notin \Omega}\mu_{xy} + \sum_{x \notin \Omega, y \in \Omega}\mu_{xy} = 2\,|\partial\Omega|_s$$

verifying Eq. (58.1). For the graph co-area formula (58.2):

$$\|\nabla f\|_{L^1(\mu)} = \sum_x \mu_x \sum_y |f(x) - f(y)|\, p(x,y) = \sum_{x,y}|f(x) - f(y)|\,\mu_{xy}.$$

while

$$|\partial\{f > t\}| = \sum_{x \in \{f>t\}, y \notin \{f>t\}}\mu_{xy} = \sum_{f(x)>t \text{ and } f(y)\leq t}\mu_{xy} = \sum 1_{f(y)\leq t<f(x)}\mu_{xy}$$

so that

$$\int_0^\infty |\partial\{f > t\}|\, dt = \int_0^\infty \sum 1_{f(y)\leq t<f(x)}\mu_{xy}dt = \sum \int_0^\infty 1_{f(y)\leq t<f(x)}\mu_{xy}dt$$
$$= \sum 1_{f(y)<f(x)}\,(f(x) - f(y))\,\mu_{xy}$$
$$= \frac{1}{2}\sum |f(x) - f(y)|\,\mu_{xy} = \frac{1}{2}\|\nabla f\|_{L^1(\mu)}.$$

58.3 Basic Inequalities

We begin with a couple of very simple Sobolev inequalities. Suppose that $f \in C_c^\infty(\mathbb{R})$, then

$$f(x) = \frac{1}{2}\left(\int_{-\infty}^{x} f'(t)dt - \int_{x}^{\infty} f'(t)dt\right)$$

from which it follows that

$$|f(x)| \le \frac{1}{2}\int_{\mathbb{R}} |f'(t)|\, dt.$$

By the Mean value inequality we have the oscillation inequality,

$$|f(y) - f(x)| \le |y - x| \, \|f'\|_\infty .$$

The first inequality is dimension dependent and probes the global structure of $\mathbb{R}$ while the second inequality works in arbitrary generality and hence does not probe the global structure of $\mathbb{R}$ in any way. Let us now list a number of inequalities which are true in $\mathbb{R}^n$ for all $f \in C_c^\infty(\mathbb{R}^n)$.

$$\|f\|_{\frac{2n}{n-2}} \le C \|\nabla f\|_2 \text{ (Sobolev Inequality)} \tag{58.3}$$

$$\|f\|_2^{1+2/n} \le C \|f\|_1^{2/n} \|\nabla f\|_2 \text{ (Nash Inequality)} \tag{58.4}$$

$$\|f\|_{2(1+2/n)}^{(1+2/n)} \le C \|f\|_2^{2/n} \cdot \|\nabla f\|_2 \text{ (Moser Inequality)} \tag{58.5}$$

$$|f(x) - f(y)| \le C_p |x - y|^{1-n/p} \|\nabla f\|_p \text{ for } p > n \text{ (oscillation inequality).} \tag{58.6}$$

The last inequality is valid for all $f \in C^\infty(\mathbb{R}^n)$ as well.

Let $W(f) = \|\nabla f\|_p$ and for a positive f let

$$f_k := (f - 2^k)_+ \wedge 2^k = \begin{cases} 2^k & \text{if } \quad f \ge 2^{k+1} \\ f - 2^k & \text{if } 2^k \le f \le 2^{k+1} \\ 0 & \text{if } \quad f \le 2^k. \end{cases}$$

Then

$$\nabla f_k = \nabla f 1_{2^k \le f \le 2^{k+1}}$$

and therefore

$$\|\nabla f\|_p^p = \int_M |\nabla f|^p \, d\mu = \int_M \sum_k |\nabla f|^p \, 1_{2^k \le f \le 2^{k+1}} d\mu = \sum_k \int_M |\nabla f_k|^p \, d\mu.$$

This shows that

$$\|\nabla f\|_p = \left(\sum_k \|\nabla f_k\|_p\right)^{1/p}.$$

This is a key truncation property for positive Lipschitz functions f. As an application we have the following theorem.

Theorem 58.5. *The Nash and Sobolev inequalities are equivalent.*

Proof. Sobolev $\Longrightarrow$ Nash. Recall the Hölder interpolation inequality,

$$\|f\|_{p_\theta} \le \|f\|_{p_0}^{1-\theta} \|f\|_{p_1}^{\theta}$$

where

$$\frac{1}{p_\theta} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}.$$

Taking $p_\theta = 2$, $p_0 = 1$ and $p_1 = \frac{2n}{n-2}$ we solve for θ to find

$$\frac{1}{2} = \frac{1-\theta}{1} + \theta\frac{n-2}{2n} = 1 + \theta\left(\frac{n-2}{2n} - 1\right) = 1 - \theta\left(\frac{n+2}{2n}\right)$$

and hence

$$\theta = \frac{n}{n+2} \text{ and } 1 - \theta = \frac{2}{n+2}$$

and therefore,

$$\|f\|_2 \le \|f\|_1^{\frac{2}{n+2}} \|f\|_{\frac{2n}{n-2}}^{\frac{n}{n+2}}.$$

This inequality along with the Sobolev inequality (58.3) shows,

$$\|f\|_2^{1+2/n} = \|f\|_2^{\frac{n+2}{n}} \le \|f\|_1^{\frac{2}{n}} \|f\|_{\frac{2n}{n-2}} \le C \|f\|_1^{\frac{2}{n}} \|\nabla f\|_2$$

which is the Nash inequality (58.4) with the same constant.

Nash $\Longrightarrow$ Sobolev. Conversely suppose the Nash inequality (58.4) is valid. Applying Nash to f_k we find

$$\left(\int f_k^2\right)^{1+2/n} = \left(\|f_k\|_2^{1+2/n}\right)^2$$
$$\le C^2 \|f_k\|_1^{4/n} \|\nabla f_k\|_2^2 = C^2 \|f_k\|_1^{4/n} \int_{B_k} |\nabla f|^2 \, d\mu$$

where $B_k := \{2^k \le f < 2^{k+1}\}$. Combining this inequality with the following two elementary inequalities;

$$\int f_k^2 \ge \int_{\{f \ge 2^{k+1}\}} f_k^2 = 2^{2k} \mu\left(\{f \ge 2^{k+1}\}\right) \text{ and}$$
$$\|f_k\|_1 \le 2^k \mu\left(f \ge 2^k\right)$$

gives

$$\left(2^{2k} \mu\left(\{f \ge 2^{k+1}\}\right)\right)^{1+2/n} \le C^2 \left(2^k \mu\left(f \ge 2^k\right)\right)^{4/n} \int_{B_k} |\nabla f|^2 \, d\mu.$$

Let $q = \frac{2n}{n-2}$ be the exponent in the Sobolev inequality and $\nu = \frac{n}{n+2} < 1$, $a_k := 2^{qk} \mu\left(f \ge 2^k\right)$ and $b_k := \int_{B_k} |\nabla f|^2 \, d\mu$. Then the above inequality says,

$$a_{k+1} \leq C' b_k^\nu a_k^{2(1-\nu)}$$

and summing this equation of $k \in \mathbb{Z}$ then gives

$$\sum_k a_k = \sum_k a_{k+1} \leq C' \sum_k b_k^\nu a_k^{2(1-\nu)} \leq C' \left(\sum_k b_k\right)^\nu \left(\sum_k a_k^2\right)^{(1-\nu)}$$

wherein we have used Holder's inequalities with the conjugate indices $1/\nu$ and $1/(1-\nu)$ in the last inequality. Since we are using counting measure, we also have

$$\|a\|_2^2 \leq \sum_k \|a\|_1 a_k \leq \|a\|_1^2$$

which combined with the previous inequality gives

$$\sum_k a_k \leq C' \left(\sum_k b_k\right)^\nu \left(\sum_k a_k\right)^{2(1-\nu)}$$
$$= C' \left(\int |\nabla f|^2 d\mu\right)^\nu \left(\sum_k a_k\right)^{2(1-\nu)}$$

which then shows

$$\left(\sum_k a_k\right)^{2\nu-1} \leq C' \left(\int |\nabla f|^2 d\mu\right)^\nu$$

and hence

$$\sum_k a_k \leq C' \left(\int |\nabla f|^2 d\mu\right)^{\frac{\nu}{2\nu-1}}.$$

We also have

$$\|f\|_q^q = \sum_k \int_{2^k < f \leq 2^{k+1}} f^q d\mu \leq \sum_k 2^{q(k+1)} \mu\left(f \geq 2^k\right) = 2^q \sum_k a_k$$

and it then follows that

$$\|f\|_q^q \leq 2^q C' \left(\int |\nabla f|^2 d\mu\right)^{\frac{\nu}{2\nu-1}}$$

which proves the Sobolev inequality (58.3). ■

58.4 A Scale of Inequalities

In this section let $\phi : \mathbb{R}_+ \to \mathbb{R}_+$ be an increasing function, for example $\phi(t) = ct^{1/n}$ and $\phi(t) = c\log(t)$ for $t \geq 2$.

Definition 58.6 (S_ϕ^p). *Given $p \in [1,\infty]$ and ϕ as above, we say S_ϕ^p holds provided*

$$\|f\|_p \leq \phi(|\Omega|) \|\nabla f\|_p \text{ for all } \Omega \ni |\Omega| < \infty \text{ and } f \in Lip_o(\Omega) \qquad (58.7)$$

where $Lip_o(\Omega)$ denotes those functions f on M or V such that f is Lipschitz and $\text{supp}(f) \subset \bar{\Omega}$ *and $f = 0$ on $\partial\Omega$.*

Proposition 58.7. *Suppose $1 \leq p < q < \infty$ and S_ϕ^p holds then $S_{\frac{q}{p}\phi}^q$ holds.*

Proof. Apply S_ϕ^p to the function $f^{q/p}$ to find

$$\|f\|_q^{q/p} = \left\|f^{q/p}\right\|_p \leq \phi(|\Omega|) \left\|\nabla f^{q/p}\right\|_p = \phi(|\Omega|) \frac{q}{p} \left\|f^{(q/p-1)} \nabla f\right\|_p$$

where we should check that the approximate chain rule holds in the graph case here. Now apply Hölder's inequality with

$$\frac{1}{p} = \frac{1}{q} + \frac{1}{pq/(q-p)}$$

to the last expression to find

$$\left\|f^{(q/p-1)} \nabla f\right\|_p \leq \|\nabla f\|_q \left\|f^{(q/p-1)}\right\|_{pq/(q-p)} = \|\nabla f\|_q \left(\int \left(f^{(q-p)/p}\right)^{pq/(q-p)}\right)^{\frac{q-p}{pq}}$$
$$= \|\nabla f\|_q \left(\int f^q\right)^{\frac{q-p}{pq}} = \|\nabla f\|_q \|f\|_q^{\frac{q-p}{p}}.$$

This gives

$$\|f\|_q^{q/p} \leq \phi(|\Omega|) \frac{q}{p} \|\nabla f\|_q \|f\|_q^{\frac{q-p}{p}}$$

that is to say

$$\|f\|_q \leq \phi(|\Omega|) \frac{q}{p} \|\nabla f\|_q.$$

■

We now give some equivalent inequalities to S_ϕ^p in the following theorem.

Theorem 58.8. *We have*

$$S_\phi^\infty \iff \phi(|B(x,r)|) \geq r \iff |B(x,r)| \geq \phi^{-1}(r)$$
$$S_\phi^1 \iff |\Omega| \leq \phi(|\Omega|) |\partial\Omega| \;\; (i.e.\ \frac{|\partial\Omega|}{|\Omega|} \geq \frac{1}{\phi(|\Omega|)}) \text{ for all reasonable } \Omega$$
$$S_\phi^2 \iff \text{to the } \phi - \text{Nash inequality}$$

(up to constants) where the ϕ – Nash inequality is the inequality,

$$\|f\|_2 \le \phi\left(\frac{\|f\|_1^2}{\|f\|_2^2}\right)\|\nabla f\|_2 \text{ for all } f \in C_c^\infty(M). \qquad (\phi - Nash) \qquad (58.8)$$

The ϕ – Nash inequality is clearly equivalent to

$$\theta\left(\|f\|_2^2\right) := \frac{\|f\|_2^2}{\phi^2\left(\frac{1}{\|f\|_2^2}\right)} \le \|\nabla f\|_2^2 \text{ for all } f \in C_c^\infty(M) \ni \|f\|_1 = 1$$

where $\theta(x) := x/\phi^2(1/x)$.

Proof. $\left(S_\phi^\infty\right)$ Let $f(x) := (r - d(x_0, x))_+ = h_r(d(x_0, x))$ where $h_r(t) = \max((r-t), 0)$. Then

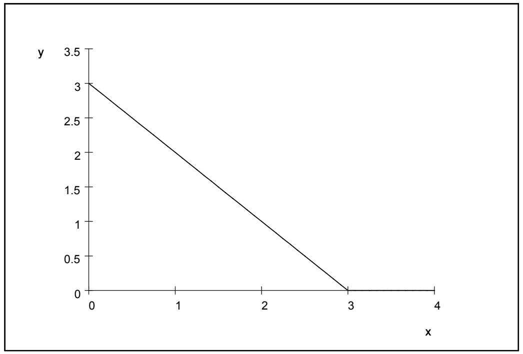


Fig. 58.1. Plot of h_r when $r = 3$.

$$|\nabla f(x)| = |h_r'(d(x_0, x))| \, |\nabla_x d(x_0, x)| = 1_{d(x_0,x)\le r}$$

and $\|f\|_\infty = r$. So putting this function into $\left(S_\phi^\infty\right)$ then implies

$$r = \|f\|_\infty \le \phi(|B(x_0, r)|) \|\nabla f\|_\infty = \phi(|B(x_0, r)|).$$

For the converse, if $\text{supp}(f) \subset \Omega$, then by the mean value theorem,

$$\|f\|_\infty \le in(\Omega) \|\nabla f\|_\infty, \qquad (58.9)$$

where $in(\Omega) :=$ (in radius of Ω) is the radius of the largest ball contained in Ω. To prove this last equation, let $x_0 \in \Omega$ and $y \in \partial\Omega$, then by the mean value theorem and the definition of $in(\Omega)$,

$$|f(x_0)| = |f(y) - f(x_0)| \le d(x_0, y) \|\nabla f\|_\infty \le in(\Omega) \|\nabla f\|_\infty.$$

This proves Eq. (58.9). Now suppose $\phi(|B(x,r)|) \ge r$ holds for all x and r, and let $B(x,r) \subset \Omega$. Then since ϕ is increasing,

$$r \le \phi(|B(x,r)|) \le \phi(|\Omega|)$$

and taking sups over all $B(x,r) \subset \Omega$ we learn that $in(\Omega) \le \phi(|\Omega|)$. Using this inequality in the estimate in Eq. (58.9) shows

$$\|f\|_\infty \le \phi(|\Omega|) \|\nabla f\|_\infty,$$

as desired.

$\left(S_\phi^1\right)$ Applying $\left(S_\phi^1\right)$ to the function 1_Ω shows

$$|\Omega| = \|1_\Omega\|_1 \le \phi(|\Omega|) \|\nabla 1_\Omega\|_1 = \phi(|\Omega|) |\partial\Omega|_1$$

as desired. For the converse, we have by the co-area formula, the above inequality and the fact that ϕ is increasing for any f with $\text{supp}(f) \subset \Omega$ and positive that

$$\begin{aligned}\|\nabla f\|_{L^1(\mu)} &= \int_0^\infty |\partial\{f > t\}| \, dt \ge \int_0^\infty \frac{|f > t|}{\phi(|f > t|)} dt \\ &\ge \frac{1}{\phi(|\Omega|)} \int_0^\infty \mu(f > t) \, dt = \frac{1}{\phi(|\Omega|)} \|f\|_1.\end{aligned}$$

$\left(S_\phi^2\right)$ Clearly $\left(S_\phi^2\right)$, i.e.

$$\|f\|_2 \le \phi(|\Omega|) \|\nabla f\|_2 \text{ for all } \text{supp}(f) \subset \Omega$$

is equivalent to

$$\lambda_1(\Omega) = \sup_{f:\text{supp}(f)\subset\Omega} \frac{(Af, f)}{\|f\|_2^2} = \sup_{f:\text{supp}(f)\subset\Omega} \frac{\|\nabla f\|_2^2}{\|f\|_2^2} \ge \frac{1}{\phi^2(|\Omega|)}.$$

Now suppose that ϕ – Nash of Eq. (58.8) holds. Since

$$\|f\|_1^2 = (f, 1_\Omega)^2 \le \|f\|_2^2 \|1_\Omega\|_2^2 = |\Omega| \, \|f\|_2^2$$

so that $|\Omega| \ge \frac{\|f\|_1^2}{\|f\|_2^2}$. Since ϕ is increasing $\phi(|\Omega|) \ge \phi\left(\frac{\|f\|_1^2}{\|f\|_2^2}\right)$, so that ϕ – Nash implies

$$\|f\|_2 \le \phi\left(\frac{\|f\|_1^2}{\|f\|_2^2}\right) \|\nabla f\|_2 \le \phi(|\Omega|) \|\nabla f\|_2$$

which is $\left(S_\phi^2\right)$.

Conversely suppose (S_ϕ^2) holds and let $f \in C_c^\infty(M, [0,\infty))$ and $t > 0$ (t to be chosen later). Then using $f < 2(f-t)$ on $f > 2t$ we find

$$\int f^2 = \int_{f>2t} f^2 + \int_{f\le 2t} f^2 \le 4\int_{f>2t} (f-t)^2 + 2t\int_{f\le 2t} f$$
$$\le 4\int (f-t)_+^2 + 2t\,\|f\|_1$$

Now applying $\left(S_\phi^2\right)$ to $(f-t)_+$ gives

$$\int (f-t)_+^2 \le \phi(|f>t|)\,\|\nabla (f-t)_+\|_2^2$$
$$\le \int (f-t)_+^2 \le \phi(|f>t|)\,\|\nabla f\|_2^2 \le \phi\left(t^{-1}\|f\|_1\right)\|\nabla f\|_2^2$$

and combining this with the last inequality implies

$$\int f^2 \le 4\phi\left(t^{-1}\|f\|_1\right)\|\nabla f\|_2^2 + 2t\,\|f\|_1\,.$$

Letting $\epsilon > 0$ and taking $t = \epsilon\,\|f\|_2^2 / \|f\|_1$ in this equation shows

$$\|f\|_2^2 \le 4\phi\left(\frac{\|f\|_1^2}{\epsilon\,\|f\|_2^2}\right)\|\nabla f\|_2^2 + 2\epsilon\,\|f\|_2^2$$

or equivalently that

$$\|f\|_2^2 \le \frac{4}{1-2\epsilon}\phi\left(\frac{\|f\|_1^2}{\epsilon\,\|f\|_2^2}\right)\|\nabla f\|_2^2\,.$$

Taking $\epsilon = 1/4$, for example, in this equation shows

$$\|f\|_2^2 \le 8\phi\left(4\frac{\|f\|_1^2}{\|f\|_2^2}\right)\|\nabla f\|_2^2$$

which is ϕ – Nash up to constants. ■

58.5 Semi-Group Theory

Definition 58.9. *A one parameter semi group T_t on a Banach space X is equicontinuos if $\|T_t\| \le M$ for all $t \ge 0$.*

Theorem 58.10. *Let (X,μ) measure space, T_t a semigroup of operators on $L^p(X,\mu)$ for $1 \le p \le \infty$ and $A := -\frac{d}{dt}|_0 T_t f$ so that $T_t = e^{-tA}$. Assume that $\|T_t\|_{1\to 1} \le M < \infty$ and $\|T_t\|_{\infty\to\infty} \le M < \infty$ for all t. Also that there exists $\theta : \mathbb{R}_+ \to \mathbb{R}_+$ such that $\int_\epsilon^\infty \frac{dx}{\theta(x)} < \infty$ and $\int_0^\epsilon \frac{dx}{\theta(x)} = \infty$ for all $\epsilon \in (0,\infty)$ such that*

$$\theta\left(\|f\|_2^2\right) \le \mathrm{Re}\,(Af,f) \text{ for all } f \in \mathcal{D}(A) \ni \|f\|_1 \le M. \tag{58.10}$$

Then T_t is ultracontractive, i.e. $\|T_t\|_{1\to\infty} < \infty$ for all $t > 0$, and moreover we have

$$\|T_t\|_{1\to\infty} \le m(t) \text{ for all } t > 0$$

where m satisfies

$$t = \int_{m(t)}^\infty \frac{dx}{\theta(x)}.$$

Remark 58.11. This type of result appears implicitly in Nash 1958. Also see Carlen, Kusuoka and Stroock 1986 and Tmisoki in 1990.

Proof. Let $f \in L^1$ with $\|f\|_1 = 1$ and so by assumption $\|T_t f\|_1 \le M$ for all $t \ge 0$. Letting $I(t) := \|T_t f\|_2^2$ we have using Eq. (58.10) to find

$$I'(t) = -2\,\mathrm{Re}\,(AT_t f, T_t f) \le -2\theta\left(\|T_t f\|_2^2\right) = -2\theta(I(t)).$$

Thus $-\frac{I'(t)}{\theta(I(t))} \ge 2$ and upon integration gives

$$\int_{I(T)}^\infty \frac{dx}{\theta(x))} \ge \int_{I(T)}^{I(0)} \frac{dx}{\theta(x))} = -\int_0^T \frac{I'(t)}{\theta(I(t))}dt$$
$$\ge \int_0^T 2dt = 2T = \int_{m(2T)}^\infty \frac{dx}{\theta(x)}$$

and therefore we have $I(T) \le m(2T)$ for all T. From this we conclude that

$$\|T_t f\|_2^2 = I(t) \le m(2t)\,\|f\|_1^2$$

showing $\|T_t\|_{1\to 2}^2 \le m(2t)$.

We will now apply this same result to T_t^* using the following comments:

1. $A^* = -\frac{d}{dt}|_0 T_t^*$ and $\mathrm{Re}\,(A^* f, f) = \mathrm{Re}\,(f, Af) = \mathrm{Re}\,(Af,f)$ we have
$$\theta\left(\|f\|_2^2\right) \le \mathrm{Re}\,(Af,f) = \mathrm{Re}\,(A^*f,f) \text{ for all } f \in \mathcal{D}(A^*) \ni \|f\|_1 \le M.$$
(Actually I am little worried about domain issues here but I do not pause to worry about them now.)
2. We also have
$$\|T_t^*\|_{1\to 1} = \sup_{\|f\|_1=1} \|T_t^* f\|_1 = \sup_{\|f\|_1=1}\sup_{\|g\|_\infty=1} |(T_t^* f, g)|$$
$$= \sup_{\|g\|_\infty=1}\sup_{\|f\|_1=1} |(f, T_t g)| = \sup_{\|g\|_\infty=1} \|T_t g\|_\infty \le M.$$

Using these comments we have $\|T_t^*\|_{1\to 2}^2 \leq m(2t)$ and hence by duality again,

$$\begin{aligned}\|T_t\|_{2\to\infty} &= \sup_{\|f\|_2=1} \|T_t f\|_\infty = \sup_{\|f\|_2=1} \sup_{\|g\|_1=1} |(T_t f, g)| \\ &= \sup_{\|g\|_1=1} \sup_{\|f\|_2=1} |(f, T_t^* g)| \\ &= \sup_{\|g\|_1=1} \|T_t^* g\|_2 = \|T_t^*\|_{1\to 2} \leq \sqrt{m(2t)}.\end{aligned}$$

Hence

$$\begin{aligned}\|T_t\|_{1\to\infty} &= \left\|T_{t/2} T_{t.2}\right\|_{1\to\infty} \leq \left\|T_{t/2}\right\|_{2\to\infty} \|T_t\|_{1\to 2} \\ &\leq \sqrt{m(t)}\sqrt{m(t)} = m(t)\end{aligned}$$

as desired. ■

Part XVII

Heat Kernels on Vector Bundles

These notes are on the construction and the asymptotic expansion for heat kernels on vector bundles over compact manifolds using Levi's method. The construction described here follows closely the presentation given in Berline, Getzler, and Vergne, "Heat Kernels and Dirac Operators."

59

Heat Equation on $\mathbb{R}^n$

Let $\Delta = \sum_{i=1}^n \partial^2/\partial x_i^2$ be the usual Laplacian on $\mathbb{R}^n$ and consider the Heat Equation

$$\left(\partial_t - \frac{1}{2}\Delta\right) u = 0 \text{ with } u(0,x) = f(x), \tag{59.1}$$

where f is a given function on $\mathbb{R}^n$. By Fourier transforming the equation in the x – variables one finds that (59.1) implies that

$$\left(\partial_t + \frac{1}{2}|\xi|^2\right) \hat{u}(t,\xi) = 0 \text{ with } \hat{u}(0,\xi) = \hat{f}(\xi). \tag{59.2}$$

and hence that $\hat{u}(t,\xi) = e^{-t|\xi|^2/2}\hat{f}(\xi)$. Inverting the Fourier transform then shows that

$$u(t,x) = \mathcal{F}^{-1}\left(e^{-t|\xi|^2/2}\hat{f}(\xi)\right)(x) = \left(\mathcal{F}^{-1}\left(e^{-t|\xi|^2/2}\right) * f\right)(x).$$

Now by well known Gaussian integral formulas one shows that

$$\mathcal{F}^{-1}\left(e^{-t|\xi|^2/2}\right)(x) = (2\pi)^{-n}\int_{\mathbb{R}^n} e^{-t|\xi|^2/2}e^{i\xi\cdot x}d\xi = (2\pi t)^{-n/2}e^{-|x|^2/2t}.$$

Let us summarize the above computations in the following Theorem.

Theorem 59.1. *Let*

$$p(t,x,y) := (2\pi t)^{-n/2} e^{-|x-y|^2/2t} \tag{59.3}$$

be the ***heat kernel*** *on* $\mathbb{R}^n$*. Then*

$$\left(\partial_t - \frac{1}{2}\Delta_x\right) p(t,x,y) = 0 \text{ and } \lim_{t\downarrow 0} p(t,x,y) = \delta_x(y), \tag{59.4}$$

where δ_x is the δ – function at x in $\mathbb{R}^n$. More precisely, if f is a continuous bounded (can be relaxed considerably) function on $\mathbb{R}^n$, then $u(t,x) = \int_{\mathbb{R}^n} p(t,x,y)f(y)dy$ is a solution to Eq. (59.1) where $u(0,x) := \lim_{t\downarrow 0} u(t,x)$.

Proof. Direct computations show that $\left(\partial_t - \frac{1}{2}\Delta_x\right)p(t,x,y) = 0$, see Proposition 63.1 and Remark 63.2 below. The main issue is to prove that $\lim_{t\downarrow 0} p(t,x,y) = \delta_x(y)$ or equivalently that $\lim_{t\downarrow 0}\int_{\mathbb{R}^n} p(t,x,y)f(y)dy = f(x)$. To show this let $p_t(v) := (2\pi t)^{-n/2}e^{-|v|^2/2t}$ and notice

$$\begin{aligned}\left|\int_{\mathbb{R}^n} p(t,x,y)f(y)dy - f(x)\right| &\le \int_{\mathbb{R}^n} p(t,x,y)\,|f(y)-f(x)|\,dy \\ &= \int_{\mathbb{R}^n} p_t(v)\,|f(x+v)-f(x)|\,dy. \end{aligned} \tag{59.5}$$

Now for a bounded function g on $\mathbb{R}^n$ we have that

$$\begin{aligned}\int_{\mathbb{R}^n}|g(v)|p_t(v)dv &= \int_{B(\delta)}|g(v)|p_t(v)dv + \int_{B(\delta)^c}|g(v)|p_t(v)dv \\ &\le \sup_{v\in B(\delta)}|g(v)| + \|g\|_0\int_{B(\delta)^c}p_t(v)dv \\ &\le \sup_{v\in B(\delta)}|g(v)| + C\,\|g\|_0\,e^{-\delta^2/4t}, \end{aligned} \tag{59.6}$$

where $\|g\|_0$ denotes the supremum norm of g. Applying this estimate to Eq. (59.5) implies,

$$\left|\int_{\mathbb{R}^n} p(t,x,y)f(y)dy - f(x)\right| \le \sup_{v\in B(\delta)}|f(x+v)-f(x)| + C\,\|f\|_0\,e^{-\delta^2/4t}.$$

Therefore if K is a compact subset of $\mathbb{R}^n$, then

$$\begin{aligned}&\overline{\lim}_{t\downarrow 0}\sup_{x\in K}\left|\int_{\mathbb{R}^n} p(t,x,y)f(y)dy - f(x)\right| \\ &\le \sup_{v\in B(\delta)}\sup_{x\in K}|f(x+v)-f(x)| \to 0 \text{ as } \delta\to 0\end{aligned}$$

by uniform continuity. This shows that $\lim_{t\downarrow 0} u(t,x) = f(x)$ uniformly on compact subsets of $\mathbb{R}^n$. ■

Notation 59.2 *We will write $\left(e^{t\Delta/2}f\right)(x)$ for $\int_{\mathbb{R}^n} p(t,x,y)f(y)dy$.*

60

An Abstract Version of E. Levi's Argument

The idea for the construction of the heat kernel for more general heat equations will be based on a method due to E. Levi. Let us illustrate the method with the following finite dimensional analogue. Suppose that L is a linear operator on a finite dimensional vector space V and let $P_t := e^{tL}$, i.e. P_t is the unique solution to the ordinary differential equation

$$\frac{d}{dt}P_t = LP_t \text{ with } P_0 = I. \tag{60.1}$$

In this finite dimensional setting it is very easy to solve Eq. (60.1), namely one may take

$$P_t = \sum_{k=0}^{\infty} \frac{t^k}{k!} L^k.$$

Such a series solution will in general not converge when L is an unbounded operator on an infinite dimensional space as are differential operators. On the other hand for the heat equation we can find quite good parametrix (approximate solution) to Eq. (60.1). Let us model this by a map $t \in \mathbb{R}_+ \to K_t \in End(V)$ such that $K_0 = I$ and

$$\frac{d}{dt}K_t - LK_t = -R_t, \tag{60.2}$$

where $\|R_t\| = O(t^\alpha)$ for some $\alpha > -1$. Using du Hamell's principle (or variation of parameters if you like) we see that K_t is given by

$$K_t = P_t - \int_0^t P_{t-s}R_s ds = P_t - (QP)_t\,, \tag{60.3}$$

where

$$(Qf)_t := \int_0^t f_{t-s}R_s ds = \int_0^t f_s R_{t-s} ds. \tag{60.4}$$

We may rewrite Eq. (60.3) as $K = (I-Q)\,P$ and hence we should have that

$$P = (I-Q)^{-1} K = \sum_{m=0}^{\infty} Q^m K. \tag{60.5}$$

Now a simple change of variables shows that $(Qf)_{t-s} = \int_s^t f_{t-r}R_{r-s}dr$ and by induction one shows that

$$(Q^m f)_t = \int_{t\Delta_m} f_{t-s_m} R_{s_m - s_{m-1}} \dots R_{s_2-s_1} R_{s_1} d\mathbf{s} \tag{60.6}$$

where

$$t\Delta_m := \{\mathbf{s} = (s_1, s_2, \dots, s_m) : 0 \le s_1 \le s_2 \dots \le s_m \le t\} \tag{60.7}$$

and $d\mathbf{s} = ds_1 ds_2 \dots ds_m$. Alternatively one also shows that

$$(Q^m K)_t = \int_{t\Delta_m} K_{s_1} R_{s_2-s_1} R_{s_3-s_2} \dots R_{s_m - s_{m-1}} R_{t-s_m} d\mathbf{s}. \tag{60.8}$$

Equation (60.6) implies that

$$(Q^m f)_t = \int_0^t f_{t-s} \left(Q^{m-1}R\right)_s ds. \tag{60.9}$$

Using this result, we may write Eq. (60.5) as

$$P_t = K_t + \int_0^t K_{t-s}V_s ds = K_t + \int_0^t K_s V_{t-s}\, ds \tag{60.10}$$

where

$$V_t = \sum_{m=0}^{\infty} (Q^m R)_t = R_t + \sum_{m=1}^{\infty} \int_{t\Delta_m} R_{t-s_m} R_{s_m-s_{m-1}} \dots R_{s_2-s_1} R_{s_1} d\mathbf{s}. \tag{60.11}$$

Let us summarize these results in the following proposition.

Proposition 60.1. *Let $\alpha \ge 0$, K, P, $R = LK - \dot{K}$, Q and V be as above. Then the series in Eq. (60.5) and Eq. (60.11) are convergent and Eq. (60.10) holds, where $P_t = e^{tL}$ is the unique solution to Eq. (60.1). Moreover,*

$$\|P - K\|_t \le \frac{C}{\alpha+1} e^{Ct^{\alpha+1}} \|K\|_t\, t^{\alpha+1} = O(t^{1+\alpha}), \tag{60.12}$$

where $\|f\|_t := \max_{0 \le s \le t} |f_s|$.

Remark 60.2. In the finite dimensional case or where L is a bounded operator, we may take $K = I$ in the previous proposition. Then $R_s = L$ is constant independent of s and

$$V_t = \sum_{m=0}^{\infty} \frac{t^m}{m!} L^{m+1}$$

which used in Eq. (60.10) gives the standard formula:

$$P_t = I + \int_0^t \sum_{m=0}^{\infty} \frac{(t-s)^m}{m!} L^{m+1}\, ds = I + \sum_{m=0}^{\infty} \frac{t^{m+1}}{(m+1)!} L^{m+1} = e^{tL}.$$

Proof. From Eq. (60.6),

$$\begin{aligned} \left|(Q^m R)_t\right| &= \left| \int_{t\Delta_m} R_{t-s_m} R_{s_m - s_{m-1}} \dots R_{s_2 - s_1} R_{s_1} d\mathbf{s} \right| \\ &\le (Ct)^{(m+1)\alpha} \int_{t\Delta_m} d\mathbf{s} = (Ct^\alpha)^{m+1} \frac{t^m}{m!} \end{aligned}$$

Therefore the series in Eq. (60.11) is absolutely convergent and

$$|V_t| \le Ct^\alpha \sum_{m=0}^{\infty} \frac{1}{m!} (Ct^{\alpha+1})^m = Ce^{Ct^{\alpha+1}} t^\alpha.$$

Using this bound on V and the uniform boundedness of K_t,

$$\int_0^t |K_s V_{t-s}|\, ds \le Ce^{Ct^{\alpha+1}} \|K\|_t \int_0^t (t-s)^\alpha\, ds = \frac{C}{\alpha+1} e^{Ct^{\alpha+1}} \|K\|_t\, t^{\alpha+1} \tag{60.13}$$

and hence P_t defined in Eq. (60.10) is well defined and is continuous in t. Moreover, (60.13) implies Eq. (60.12) once we shows that $P_t = e^{tL}$. This is checked as follows,

$$\begin{aligned} \frac{d}{dt} \int_0^t K_{t-s} V_s ds &= V_t + \int_0^t \dot{K}_{t-s} V_s ds = V_t + \int_0^t (LK_{t-s} - R_{t-s}) V_s ds \\ &= V_t + L \int_0^t K_{t-s} V_s ds - (QV)_t = L \int_0^t K_{t-s} V_s ds + R_t. \end{aligned}$$

Thus we have,

$$\begin{aligned} \frac{d}{dt} P_t &= \dot{K}_t + L \int_0^t K_{t-s} V_s ds + R_t \\ &= LK_t + L \int_0^t K_{t-s} V_s ds = LP_t. \end{aligned}$$

■

61

Statement of the Main Results

Let M be a compact Riemannian Manifold of dimension n and Δ denote the Laplacian on $C^\infty(M)$. We again wish to solve the heat equation (59.1). It is natural to define a kernel $\rho(t,x,y)$ in analogy with the formula for $p(t,x,y)$ in Eq. (59.3), namely let

$$\rho(t,x,y) := (2\pi t)^{-n/2}\, e^{-d^2(x,y)/2t}, \tag{61.1}$$

where $d(x,y)$ is the Riemannian distance between two point $x, y \in M$. We may then define the operator T_t on $C^\infty(M)$ by

$$T_t f(x) = \int_M \rho(t,x,y) f(y) d\lambda(y), \tag{61.2}$$

where λ is the volume measure on M. Although, $\lim_{t\downarrow 0} T_t f = f$, it is not the case that $u(t,x) = T_t f(x)$ is a solution to the heat equation on M. This is because ρ does not satisfy the Heat equation. Nevertheless, ρ is an approximate solution as will be seen Proposition 63.1 below. Moreover, ρ will play a crucial role in constructing the true heat kernel $p(t,x,y)$ for M. Let us now summarize the main theorems to be proved.

61.1 The General Setup: the Heat Eq. for a Vector Bundle

Let $\pi : E \to M$ be a Vector bundle with connection ∇^E. We will usually denote the covariant derivatives on TM and E all by ∇. For a section $S \in \Gamma(E)$, let $\square S := \operatorname{tr}\left(\nabla^{T^*M\otimes E}\nabla^E S\right)$ be the rough or Bochner Laplacian on E and let

$$L := \frac{1}{2}\square + \mathcal{R},$$

where $\mathcal{R}$ is a section of $\operatorname{End}(E)$. We are interested in solving

$$(\partial_t - L)\, u = 0 \text{ with } u(0,x) = f(x), \tag{61.3}$$

where $u(t,\cdot)$ and $f(\cdot)$ are section of $\Gamma(E)$.

61.2 The Jacobian (J – function)

Definition 61.1. *Let $D : TM \to \mathbb{R}$ be defined so that for each $y \in M$, $\exp_y^* \lambda = D\lambda_y$ where $\lambda = \lambda_M$ is the volume form on M and λ_y is the volume form on T_yM. More explicitly, if $\{e_i\}_{i=1}^n$ is an oriented orthonormal basis for T_yM, then and $v \in T_yM$, then*

$$\begin{aligned} D(v) &= \lambda(\exp_*(e_1)_v, \exp_*(e_2)_v \dots, \exp_*(e_n)_v) \\ &= \sqrt{\det\left\{\left(\exp_{y*}(e_i)_v, \exp_{y*}(e_j)_v\right)\right\}_{i,j=1}^n}, \end{aligned} \tag{61.4}$$

where $\exp_* w_v = \frac{d}{dt}|_0 \exp(v+tw)$. Further define $J(x,y) = D(\exp_y^{-1}(x))$.

Notice that $J(\cdot,y)$ satisfies

$$\exp_y^*\left(\frac{1}{J(\cdot,y)}\lambda\right) = \lambda_y.$$

Alternatively we have that $J(x,y) = \det\left(\exp_{y*}(\cdot)_v\right)$ where $v = \exp_y^{-1}(x)$. To be more explicit, let $\{e_i\}_{i=1}^n$ be an orthonormal basis for T_yM, then

$$J(x,y) = \sqrt{\det\left\{\left(\exp_{y*}(e_i)_v, \exp_{y*}(e_j)_v\right)\right\}_{i,j=1}^n}.$$

Remark 61.2 (Symmetry of J). It is interesting to notice that J is a symmetric function. We will not need this fact below so the proof may be skipped. We will also be able to deduce the symmetry of J using the asymptotic expansion of the heat kernel along with the symmetry of the heat kernel.

Proof. Let $w_y \in T_yM$, then

$$w_y d^2(x,\cdot) = 2\left(V(x,y), w_y\right) = -2\left(\exp_y^{-1}(x), w_y\right)$$

where $V(x,y) := \frac{d}{dt}|_{t=1}\exp(t\exp_x^{-1}(y)) = -\exp_y^{-1}(x)$. Thus if $u_x \in T_xM$, then

$$u_x w_y d^2(x,y) = -2\left(\left(\exp_y^{-1}\right)_* u_x, w_y\right).$$

Now

$$J(x,y) = \det\left[\left(\exp_y^{-1}\right)_*\right] = \det\left[\left\{-\frac{1}{2}u_i w_j d^2(x,y)\right\}_{i,j=1}^n\right]$$

where $\{u_i\}$ and $\{w_j\}$ is an orthonormal basis of T_xM and T_yM respectively. From this last formula it is clear from the fact that $d(x,y) = d(y,x)$ that $J(x,y) = J(y,x)$. ■

Lemma 61.3 (Expansion of D). *The function J is symmetric, $J(x,y) = J(y,x)$. Moreover*

$$D(v) = 1 - \frac{1}{6}(\operatorname{Ric}\ \ v, v) + O(v^3)$$

and hence

$$J(x,y) = \left(\operatorname{Ric}\ \ \exp_y^{-1}(x), \exp_y^{-1}(x)\right) + O(d^3(x,y)).$$

The proof of this result will be given in the Appendix below since the result is not really needed for our purposes.

61.3 The Approximate Heat Kernels

Theorem 61.4 (Approximate Heat Kernel). *For $x, y \in M$, let*

$$\gamma_{y,x}(t) := \exp_y(t \exp_y^{-1}(x))$$

so that $\gamma_{x,y}$ is the geodesic connecting y to x. Also let $//_t(\gamma_{x,y})$ denote parallel translation along $\gamma_{x,y}$. (Read $\gamma_{x,y}$ as $\gamma_{x\leftarrow y}$.) Define, for (x,y) near the diagonal $\Delta \subset M \times M$ and $k = 0, 1, 2 \ldots$, $u_k(x,y) : E_y \to E_x$ inductively by

$$u_{k+1}(x,y) = u_0(x,y) \int_0^1 s^k u_0(\gamma_{x,y}(s), y)^{-1} \left(L_x u_k\right)(\gamma_{x,y}(s), y) ds \text{ for } k = 0, 1, 2 \ldots \tag{61.5}$$

and

$$u_0(x,y) = \frac{1}{\sqrt{J(x,y)}} //_1(\gamma_{x,y}). \tag{61.6}$$

For $q = 0, 1, 2 \ldots$,let

$$\Sigma_q(t,x,y) := \rho(t,x,y) \sum_{k=0}^{q} t^k u_k(x,y), \tag{61.7}$$

where

$$\rho(t,x,y) := (2\pi t)^{-n/2} e^{-d^2(x,y)/2t}. \tag{61.8}$$

Then

$$\left(\partial_t - L_x\right) \Sigma_q(t,x,y) = -t^q \rho(t,x,y) L_x u_q(x,y). \tag{61.9}$$

Definition 61.5 (Cut off function). *Let $\epsilon > 0$ be less than the injectivity radius of M and choose $\Psi \in C_c^\infty(-\epsilon^2, \epsilon^2)$ such that $0 \le \Psi \le 1$ and Ψ is 1 in a neighborhood of 0. Set $\psi(x,y) = \Psi(d^2(x,y))$, a cutoff function which is one in a neighborhood of the diagonal and such that $\psi(x,y) = 0$ if $d(x,y) \ge \epsilon$.*

Corollary 61.6 (Approximate Heat Kernel). *Let $k_q(t,x,y) := \psi(x,y)\Sigma_q(t,x.$ Define*

$$r_q(t,x,y) := -\left(\partial_t - L_x\right) k_q(t,x,y),$$

then

$$\left\| \partial_t^k r_q(t,\cdot,\cdot) \right\|_l \le C t^{q-n/2-l/2-k}, \tag{61.10}$$

where $\|r\|_l$ denotes the supremum norm of r and all of its derivatives up to order l.

61.4 The Heat Kernel and its Asymptotic Expansion

Theorem 61.7 (Existence of Heat Kernels). *There exists a heat kernel $p(t,x,y) : E_y \to E_x$ for L, i.e. p is a C^1 –function in t and C^2 in (x,y) such that*

$$\left(\partial_t - L_x\right) p(t,x,y) = 0 \text{ with } \lim_{t\downarrow 0} p(t,x,y) = \delta_x(y) id. \tag{61.11}$$

Remark 61.8. The explicit formula for $p(t,x,y)$ is derived by a formal application of equations (60.10) and (60.11) above. The results are:

$$p(t,x,y) = k_q(t,x,y) + \int_0^t \int_M k_q(t-s,x,z) v_q(s,z,y) d\lambda(z) ds \tag{61.12}$$

where

$$v_q(s,x,y) = \sum_{m=1}^{\infty} r^m(s,x,y) \tag{61.13}$$

and the kernels r^m are defined inductively by where $r^1 = r$ and for $m \ge 2$

$$r^m(s,x,y) = \int_0^s \int_M r(s-r,x,z) r^{m-1}(r,z,y) d\lambda(z) dr. \tag{61.14}$$

Corollary 61.9 (Uniqueness of Heat Kernels). *The heat equation (61.3) has a unique solution. Moreover, there is exactly one solution to (61.11).*

Proof. Let $u(t,x) := \int_M p(t,x,y) f(y) d\lambda(y)$, then u solves the heat equation (61.3). We will prove uniqueness of u using the existence of the adjoint problem. In order to carry this out we will need to know that $L^t = \frac{1}{2}\square_{E^*} + \mathcal{R}^t : \Gamma(E^*) \to \Gamma(E^*)$ is the formal transpose of L in the sense that

$$\int_M \langle Lf, g \rangle d\lambda = \int_M \langle f, L^t g \rangle d\lambda \tag{61.15}$$

for all sections $f \in \Gamma(E)$ and $g \in \Gamma(E^*)$. Here $\square_{E^*}$ is the rough Laplacian on E^*. Indeed, let X be the vector field on M such that $(X, v) = \langle \nabla_v f, g \rangle$ for all $v \in T_m M$. Then

$$(\nabla_v X, \cdot) = \nabla_v \langle \nabla_{\cdot} f, g \rangle = \langle \nabla_v \left(\nabla_{\cdot} f\right), g \rangle + \langle \nabla_{\cdot} f, \nabla_v g \rangle,$$

so that in particular

$$(\nabla_v X, v) = \langle \nabla^2_{v \otimes v} f, g \rangle + \langle \nabla_v f, \nabla_v g \rangle.$$

Let $v = e_i$, where $\{e_i\}$ an orthonormal frame, and sum this equation on i to find that

$$\div(X) = \langle \square f, g \rangle + \langle \nabla f, \nabla g \rangle. \tag{61.16}$$

Using the Riemannian metric on M to identify one forms with vector fields, we may write this equality as:

$$\div\langle\nabla_{\cdot} f, g\rangle = \langle\Box f, g\rangle + \langle\nabla f, \nabla g\rangle.$$

Similarly one shows that

$$\div\langle f, \nabla_{\cdot} g\rangle = \langle f, \Box g\rangle + \langle\nabla f, \nabla g\rangle,$$

which subtracted from the previous equation gives "Green's identity,"

$$\langle f, \Box g\rangle - \langle\Box f, g\rangle = \div\left(\langle f, \nabla_{\cdot} g\rangle - \langle\nabla_{\cdot} f, g\rangle\right) \tag{61.17}$$

Integrating equations (61.16) and (61.17) over M, we find, using the divergence theorem, that

$$\int_M \langle\Box f, g\rangle d\lambda = \int_M \langle f, \Box g\rangle d\lambda = -\int_M \langle\nabla f, \nabla g\rangle d\lambda.$$

Thus

$$\begin{aligned}\int_M \langle Lf, g\rangle d\lambda &= \frac{1}{2}\int_M \langle\Box f, g\rangle d\lambda + \int_M \langle\mathcal{R} f, g\rangle d\lambda \\ &= \frac{1}{2}\int_M \langle f, \Box g\rangle d\lambda + \int_M \langle f, \mathcal{R}^t g\rangle d\lambda = \int_M \langle f, L^t g\rangle d\lambda,\end{aligned}$$

proving Eq. (61.15).

Suppose that u is a solution to Eq. (61.3) with $u(0,x) = 0$ for all x. By applying Theorem 61.7, we can construct a heat kernel q_t for L^t. Given $g \in \Gamma(E^*)$, let

$$v(t,x) := \int_M q_{T-t}(x,y) g(y) d\lambda(y)$$

for $t < T$. Now consider

$$\begin{aligned}&\frac{d}{dt}\int_M \langle u(t,x), v(t,x)\rangle d\lambda(x) \\ &\quad = \int_M \langle Lu(t,x), v(t,x)\rangle d\lambda(x) - \int_M \langle u(t,x), L^t v(t,x)\rangle d\lambda(x) = 0,\end{aligned}$$

and therefore, $\int_M \langle u(t,x), v(t,x)\rangle d\lambda(x)$ is constant in t. Considering this expression in the limit that t tends to 0 and T implies that

$$0 = \int_M \langle 0, v(0,x)\rangle d\lambda(x) = \int_M \langle u(T,x), g(x)\rangle d\lambda(x).$$

Since g is arbitrary, this implies that $u(T,x) = 0$ for all x. Hence the solution to equation (61.3) is unique. It is now easy to use this result to show that $p(t,x,y)$ must be unique as well. ■

Theorem 61.10 (Assymptotics of the Heat Kernel). *Let $p(t,x,y)$ be the heat kernel described by Eq. (61.11), then p is smooth in (t,x,y) for $t > 0$. Moreover if K^q is as in Corollary 61.6, then*

$$\left\|\partial_t^k \left(p(t,\cdot,\cdot) - K^q(t,\cdot,\cdot)\right)\right\|_l = O(t^{q-n/2-l/2-k}) \tag{61.18}$$

provided that $q > n/2 + l/2 + k$.

62

Proof of Theorems 61.7 and 61.10

In this section we will give the proof of Theorems 61.7 and 61.10 assuming Theorem 61.4 and Corollary 61.6.

62.1 Proof of Theorem 61.7

Let l and k be given and fix $q > n/2 + l/2 + k.$. Let $k(t, x, y) = k_q(t, x, y)$ and $r(t, x, y) = r_q(t, x, y)$ as in Corollary 61.6. Let

$$K_t f(x) := \int_M k(t, x, y) f(y) d\lambda(y) \text{ and } R_t f(x) := \int_M r(t, x, y) f(y) d\lambda(y).$$

Following the strategy described in Section 60, we will let $v(t, x, y)$ be the kernel of the operator $\sum_{m=0}^{\infty} Q^m R$, where Q is as in Eq. (60.4). That is

$$v(s, x, y) = \sum_{m=1}^{\infty} r^m(s, x, y) \tag{62.1}$$

where $r^1 = r$ and for $m \geq 2$

$$\begin{aligned} r^m&(s,x,y) \\ &= \int_{s\Delta_{m-1}} \int_{M^{m-1}} r(s - s_{m-1}, x, y_{m-1}) r(s_{m-1} - s_{m-2}, y_{m-1}, y_{m-2}) \dots r(s_1, \\ &= \int_0^s \int_M r(s - r, x, z) r^{m-1}(r, z, y) d\lambda(z) dr \end{aligned} \tag{62.2}$$

and $d\mathbf{y} = d\lambda(y_1) \dots d\lambda(y_m)$. The kernel r^m is easy to estimate using (61.10) to find that

$$\begin{aligned} \left\|\partial_s^k r^m(s, \cdot, \cdot)\right\|_l &\leq \left(C s^{q-n/2}\right)^m s^{-l/2-k} Vol(M)^{m-1} s^{m-1}/(m-1)! \\ &= C^m s^{m(q-n/2)-l/2-k} Vol(M)^{m-1} s^{m-1}/(m-1)!. \end{aligned} \tag{62.3}$$

and from this it follows that $\sum_{m=1}^{\infty} \left\|\partial_s^k r^m(s, \cdot, \cdot)\right\|_l = O(s^{(q-n/2)-l/2-k})$. Therefore, v is well defined with $\partial_s^k v(s, x, y)$ in Γ_l.

Proof. Let $p(t, x, y)$ be the kernel of the operator P_t defined in Eq. (60.10) i.e.

$$\begin{aligned} p(t, x, y) &= k(t, x, y) + \int_0^t \int_M k(t - s, x, z) v(s, z, y) d\lambda(z) ds & (62.4) \\ &= k(t, x, y) + \int_0^t \int_M k(s, x, z) v(t - s, z, y) d\lambda(z) ds. & (62.5) \end{aligned}$$

Using Eq. (62.5), we find (since $\partial_t v(0, z, y) = 0$) that

$$\begin{aligned} \partial_t p(t, x, y) &= \partial_t k(t, x, y) + \int_0^t \int_M k(s, x, z) \partial_t v(t - s, z, y) d\lambda(z) ds \\ &= L_x k(t, x, y) - r(t, x, y) \\ &\quad + \int_0^t \int_M k(s, x, z) \partial_t v(t - s, z, y) d\lambda(z) ds \end{aligned} \tag{62.6}$$

More generally,

$$\partial_t^i p(t, x, y) = \partial_t^i k(t, x, y) + \int_0^t \int_M k(s, x, z) \partial_t^i v(t - s, z, y) d\lambda(z) ds,$$

from which it follows that $\partial_t^i p$ is continuous in Γ_l for all $i \leq k$. Furthermore,

$$\begin{aligned} \left\|\partial_t^i p(t, \cdot, \cdot) - \partial_t^i k(t, \cdot, \cdot)\right\|_l &\leq C \int_0^t \left\|\partial_t^i v(t - s, \cdot, \cdot)\right\|_l ds \\ &\leq C \int_0^t (t - s)^{q-n/2-l/2-i} ds \\ &= O(t^{q-n/2-l/2-i+1}). \end{aligned} \tag{62.7}$$

To finish the proof of Theorem 61.7, we need only verify Eq. (61.11). The assertion that $\lim_{t \downarrow 0} p(t, x, y) = \delta_x(y) id$ follows from the previous estimate and the analogous property of $k(t, x, y)$. Fubini's theorem and integration by parts shows that

$$\begin{aligned}
\int_\epsilon^t \int_M & k(s,x,z)\partial_t v(t-s,z,y)d\lambda(z)ds \\
&= -\int_M k(s,x,z)v(t-s,z,y)d\lambda(z)\,|_{s=\epsilon}^{s=t} \\
&+ \int_\epsilon^t \int_M \partial_s k(s,x,z)v(t-s,z,y)d\lambda(z)ds \\
&= \int_M k(\epsilon,x,z)v(t-\epsilon,z,y)d\lambda(z) \\
&+ \int_\epsilon^t \int_M \left(L_x k(s,x,z) - r(s,x,z)\right) v(t-s,z,y)d\lambda(z)ds \\
&= \int_M k(\epsilon,x,z)v(t-\epsilon,z,y)d\lambda(z) \\
&+ L_x \int_\epsilon^t \int_M k(s,x,z)v(t-s,z,y)d\lambda(z)ds \\
&- \int_\epsilon^t \int_M r(s,x,z)v(t-s,z,y)d\lambda(z)ds.
\end{aligned}$$

Making use of the fact that K is uniformly bounded on Γ_l and that the strong–$\lim_{\epsilon\downarrow 0} K_\epsilon = I$, we may pass to the limit, $\epsilon \to 0$, in this last equality to find that

$$\begin{aligned}
\int_0^t \int_M & k(s,x,z)\partial_t v(t-s,z,y)d\lambda(z)ds \\
&= v(t,x,y) + L_x \int_0^t \int_M k(s,x,z)v(t-s,z,y)d\lambda(z)ds \\
&- \int_0^t \int_M r(s,x,z)v(t-s,z,y)d\lambda(z)ds \\
&= r(t,x,y) + L_x \int_0^t \int_M k(s,x,z)v(t-s,z,y)d\lambda(z)ds,
\end{aligned} \tag{62.8}$$

wherein the last equality we have made use of equations (62.1) and (62.2) to conclude that

$$v(t,x,y) - \int_0^t \int_M r(s,x,z)v(t-s,z,y)d\lambda(z)ds = r(t,x,y).$$

Combining (62.6) and (62.8) implies that $(\partial_t - L_x)\, p(t,x,y) = 0$. ■

62.2 Proof of Theorem 61.10

Because q in the above proof was arbitrary, we may construct a kernel $p(t,x,y)$ as in the previous section which is C^N for any N we desire. By the uniqueness

of p, Corollary 61.9, the kernel $p(t,x,y)$ constructed in the proof of Theorem 61.7 is independent of the parameter q. Therefore, by choosing q as large, we see that p is in fact infinitely differentiable in (t,x,y) with $t > 0$. Finally the estimate in Eq. (61.18) has already been proved in Eq. (62.7).

63

Properties of ρ

For the time being let $y \in M$ be a fixed point and $r(x) := d(x,y)$. Also let $v(x) := \exp_y^{-1}(x)$, $\gamma_x(t) := \exp(tv(x))$ and V be the "radial" vector field,

$$V(x) = V_y(x) := \frac{d}{dt}|_1 \exp(tv(x)) = //_1(\gamma_x)v(x), \tag{63.1}$$

where $//_t(\gamma_x)$ is used to denote parallel translation along γ_x up to time t. Notice that V is a smooth vector field on a neighborhood of y. To simplify notation we will write $\rho(t,x)$ for $\rho(t,x,y)$, i.e.

$$\rho(t,x) = (2\pi t)^{-n/2} e^{-r^2(x)/2t}. \tag{63.2}$$

The main proposition of this section is as follows.

Proposition 63.1. *Fix $y \in M$, let $J(x) = J(x,y)$ (see Definition 61.1 above) and $\rho(t,x)$ be as in Eq. (63.2), then*

$$\begin{aligned}\left(\partial_t - \frac{1}{2}\Delta\right)\rho &= \frac{1}{2t} r \partial \ln J/\partial r\, \rho = \frac{1}{2t}(V \ln J)\rho \\ &= \frac{1}{2t}(\nabla \cdot V - n)\rho. \end{aligned} \tag{63.3}$$

Remark 63.2. If $M = \mathbb{R}^n$ with the standard metric, then $V(x) = x, \nabla \cdot V = n$ (and $J \equiv 1$) so that

$$p(t,x) = (2\pi t)^{-n/2} e^{-x^2/2t}$$

is an exact solution to the heat equation as is seen from Eq. (63.3). Moreover, the constants have been chosen such that $\int_{\mathbb{R}^n} p(t,x)dx = 1$ for all $t > 0$. From this fact and the fact that $(2\pi t)^{-n/2} e^{-x^2/2t}$ has most of its mass within a radius of size order $\sqrt{t}$, it follows that $\lim_{t\downarrow 0} p(t,x) = \delta(x)$. Similar statements hold for $\rho(t,x)$ given in Eq. (63.2).

In order to prove the Proposition we will need to introduce some more notation which will allow us to compute the Laplacian on radial functions $f(r)$, see Lemma 65.3 below.

Notation 63.3 (Geodesic Polar Coordinates) *Let $y \in M$ be fixed, $r(x) := d(x,y)$ and $\theta(x) := \exp_y^{-1}(x)/r(x)$. So that $(r,\theta) : M \to \mathbb{R}_+ \times S_y$ where $S = S_y$ is the unit sphere in T_yM. We also write $\partial f/\partial r = g_r(r,\theta)$ when $f = g(r,\theta)$. Alternatively, where*

$$\partial f/\partial r = \frac{1}{r}Vf = \frac{1}{r}\frac{d}{dt}|_1 f(\exp_y(t\exp_y^{-1}(\cdot))) \tag{63.4}$$

$$= \frac{d}{dt}|_0 f(\exp_y((r+t)\,\theta) \tag{63.5}$$

and V is the vector field given in Eq. (63.2) above.

Notice that with this notation $\exp_y^{-1}(x) = r(x)\theta(x)$ and $J(x,y) = D(r(x)\theta(x),y)$.

63.0.1 Proof of Proposition 63.1

We begin with the logaritheoremic derivatives of ρ,

$$\partial_t \ln \rho(t,x) = -\frac{n}{2t} + \frac{r^2}{2t^2}$$

and

$$\nabla \ln \rho(t,x) = -\frac{\nabla r^2}{2t} = -\frac{1}{t}V.$$

Therefore,

$$\begin{aligned}\Delta\rho = \nabla \cdot (\rho \nabla \ln \rho(t,x)) &= -\frac{1}{t}\nabla \cdot (\rho V) \\ &= \frac{1}{t^2}|V|^2 \rho - \frac{1}{t}\rho \nabla \cdot V = \rho\left(\frac{r^2}{t^2} - \frac{1}{2t}\Delta r^2\right)\end{aligned}$$

and hence

$$\begin{aligned}\left(\partial_t - \frac{1}{2}\Delta\right)\rho &= \rho\left(-\frac{n}{2t} + \frac{r^2}{2t^2} - \frac{r^2}{2t^2} + \frac{1}{4t}\Delta r^2\right) \\ &= \frac{1}{2t}\left(\frac{1}{2}\Delta r^2 - n\right)\rho \\ &= \frac{1}{2t}(\nabla \cdot V - n)\rho = \frac{1}{2t}(r\partial \ln J/\partial r)\rho.\end{aligned}$$

63.0.2 On the Operator Associated to the Kernel ρ

We now modify the definition of T_t in Eq. (61.2) by inserting the cutoff function ψ as in Definition 61.5, that is let

$$T_t f(x) := \int_M \psi(x,y)\rho(t,x,y)f(y)d\lambda(y). \tag{63.6}$$

We will end this section with some basic properties of T_t.

Theorem 63.4. *Let $T_t f$ be as in Eq. (63.6). Then for $t>0$, $T_t : C(M) \to C^\infty(M)$, for each l, there is a constant C_l such that $\|T_t f\|_l \le C_l \|f\|_l$ for all $0 < t \le 1$ and $f \in C^l(M)$ and moreover $\lim_{t\downarrow 0} \|T_t f - f\|_l = 0$. Here, $\|f\|_l$ denotes the sup–norm of f and all of its derivatives up to order l.*

Proof. First off, since $\psi(x,y)\rho(t,x,y)$ is a smooth function in (x,y), it is clear that $T_t f(x)$ is smooth. To prove the remaining two assertions, let us make the change of variables, $y = \exp_x(v)$ in the definition of $T_t f$. This gives,

$$\begin{aligned} T_t f(x) &= \int_{B_x(\epsilon)} \psi(x,\exp_x(v))\rho(t,x,\exp_x(v))f(\exp_x(v))D(v)dv \\ &= \int_{T_x M} \Psi(|v|^2)\,(2\pi t)^{-n/2}\, e^{-|v|^2/2t} f(\exp_x(v))D(v)dv \end{aligned}$$

where $B_x(\epsilon)$ be the ball of radius ϵ centered at $0_x \in T_x M$. Now let $u(x)$ be a local orthonormal frame on M, so that $u(x) : \mathbb{R}^n \to T_x M$ is a smoothly varying orthogonal isomorphism for x in some neighborhood of M. We now make the change of variables $v \to u(x)v$ and $v \to \sqrt{t}u(x)v$ with $v \in \mathbb{R}^n$ in the above displayed equation to find,

$$\begin{aligned} T_t f(x) &= \int_{\mathbb{R}^n} \Psi(|v|^2)f(\exp_x(u(x)v))D(u(x)v)p_t(v)dv & (63.7)\\ &= \int_{\mathbb{R}^n} \Psi(t\,|v|^2)f(\exp_x(u(x)\sqrt{t}v))D(u(x)\sqrt{t}v)p_1(v)dv & (63.8) \end{aligned}$$

where $p_t(v) := (2\pi t)^{-n/2}\, e^{-|v|^2/2t}$.

Suppose that L is a l 'th order differential operator on M, then from Eq. (63.7) we find that

$$(LT_t f)(x) = \int_{\mathbb{R}^n} \Psi(|v|^2)L_x\left[f(\exp_x(u(x)v))D(u(x)v)\right]p_t(v)dv$$

from which we see that

$$|(LT_t f)(x)| \le C_l(L)\,\|f\|_l \int_{\mathbb{R}^n} p_t(v)dv = C_l(L)\,\|f\|_l\,.$$

This shows that $\|T_t f\|_l \le C_l \|f\|_l\,.$

Using the product and the chain rule,

$$L_x\left[f(\exp_x(u(x)v))D(u(x)v,x)\right] = \sum_k a_k(x,v)\,(\mathcal{L}_k f)\,(\exp_x(u(x)v)),$$

where $a_k(x,v)$ are smooth functions of (x,v) with $|v| \le \epsilon$ and $\mathcal{L}_k$ are differential operators of degree at most l. Noting that

$$Lf(x) = \sum_k a_k(x,0)\,(\mathcal{L}_k f)\,(x),$$

we find that

$$\begin{aligned} &|\,(LT_t f)\,(x) - Lf(x)| \\ &\quad\le \sum_k \int_{\mathbb{R}^n} \left| \begin{array}{c} a_k(x,v)\,(\mathcal{L}_k f)\,(\exp_x(u(x)v)) \\ -a_k(x,0)\,(\mathcal{L}_k f)\,(x) \end{array} \right| \Psi(|v|^2)p_t(v)dv. \end{aligned}$$

Applying the estimate in Eq. (59.6) to the previous equation implies that

$$\begin{aligned} &||LT_t f - Lf||_0 \\ &\quad\le C\,\|f\|_l\, e^{-\delta^2/4t} + \sum_k \sup_{v\in B(\delta)} \sup_x \left| \begin{array}{c} a_k(x,v)\,(\mathcal{L}_k f)\,(\exp_x(u(x)v)) \\ -a_k(x,0)\,(\mathcal{L}_k f)\,(x) \end{array} \right| \end{aligned}$$

and therefore

$$\overline{\lim}_{t\downarrow 0}||LT_t f - Lf||_0 \le \sum_k \sup_{v\in B(\delta)} \sup_x \left| \begin{array}{c} a_k(x,v)\,(\mathcal{L}_k f)\,(\exp_x(u(x)v)) \\ -a_k(x,0)\,(\mathcal{L}_k f)\,(x) \end{array} \right|$$

which tends to zero as $\delta \to 0$ by uniform continuity. From this we conclude that $\lim_{t\downarrow 0} ||T_t f - f||_l = 0$. ■

To conclude this section we wish to consider $\lim_{t\downarrow 0}(\partial_t - \frac{1}{2}\Delta)T_t$.

Theorem 63.5. *Let T_t be as above and S be the scalar curvature on M. Then $\partial_t T_t = (\frac{1}{2}\Delta - \frac{1}{6}S)T_t + O(\sqrt{t})$. So if we used T_t for K in the construction in Proposition 60.1, we would construct $e^{t(\Delta/2 - S/6)}$ rather than $e^{t\Delta/2}$.*

Proof. We will start by computing,

$$\begin{aligned} &(\partial_t - \frac{1}{2}\Delta)T_t f(x) \\ &\quad= \int_M (\partial_t - \frac{1}{2}\Delta_x)\psi(x,y)\rho(t,x,y)f(y)d\lambda(y) \\ &\quad= \int_M \psi(x,y)(\partial_t - \frac{1}{2}\Delta_x)\rho(t,x,y)f(y)d\lambda(y) + O(t^\infty), \\ &\quad= \int_M \psi(x,y)\frac{1}{2t}\,(V_y(x)\ln J(\cdot,y))\rho_t(x,y)\, f(y)d\lambda(y) + O(t^\infty). \end{aligned} \tag{63.9}$$

Using Lemma 61.3, we find when $x = \exp_y(v)$, that

$$\begin{aligned}(V_y(x)\ln J(\cdot,y)) &= \partial_v \ln D(v) = \partial_v \ln(1 - \frac{1}{6}(\mathrm{Ric}\ v, v) + O(v^3)) \\ &= \partial_v(-\frac{1}{6}(\mathrm{Ric}\ v, v) + O(v^3)) \\ &= -\frac{1}{3}\left(\mathrm{Ric}\ \exp_y^{-1}(x), \exp_y^{-1}(x)\right) + O(d^3(x,y)).\end{aligned}$$

Using the symmetry of J or by direct means one may conclude that

$$\left(\mathrm{Ric}\ \exp_y^{-1}(x), \exp_y^{-1}(x)\right) = \left(\mathrm{Ric}\ \exp_x^{-1}(y), \exp_x^{-1}(y)\right) + O(d^3(x,y)) \tag{63.10}$$

so that

$$(V_y(x)\ln J(\cdot,y)) = -\frac{1}{3}\left(\mathrm{Ric}\ \exp_x^{-1}(y), \exp_x^{-1}(y)\right) + O(d^3(x,y)).$$

To check Eq. (63.10) directly, let $\gamma_v(t) = \exp(tv)$ and notice that

$$\frac{d}{dt}(\mathrm{Ric}\ \dot\gamma(t), \dot\gamma(t)) = \left((\nabla_{\dot\gamma(t)}\mathrm{Ric}\ \dot\gamma(t)), \dot\gamma(t)\right) = O(v^3).$$

Integrating this expression implies that $(\mathrm{Ric}\ \dot\gamma(1), \dot\gamma(1)) = (\mathrm{Ric}\ v, v) + O(v^3)$. Taking $v = \exp_x^{-1}(y)$ implies Eq. (63.10).

Using this result in (63.9) and making the change of variables $y = \exp_x v$ as above we find that

$$\begin{aligned}&(\partial_t - \frac{1}{2}\Delta)T_t f(x) \\ &= -\frac{1}{3}\frac{1}{2t}\int_{T_xM}\Psi(|v|^2)\frac{e^{-|v|^2/2t}}{(2\pi t)^{n/2}}\left\{(\mathrm{Ric}\ v, v) + O(v^3)\right\} f(\exp_x(v))D(v)dv \\ &\quad + O(t^\infty) \\ &= -\frac{1}{6t}\left\{S(x)f(x)t + O(t^{3/2})\right\} = -\frac{1}{6}S(x)f(x) + O(t^{1/2}).\end{aligned}$$

Therefore, $\partial_t T_t = (\frac{1}{2}\Delta - \frac{1}{6}S)T_t + O(\sqrt{t})$. ■

64

Proof of Theorem 61.4 and Corollary 61.6

64.1 Proof of Corollary 61.6

We will begin with a Proof of Corollary 61.6 assuming Theorem 61.4. Using Eq. (61.9) and the product rule,

$$\begin{aligned}
r_q(t,x,y) &= (\partial_t - L_x)\, K_q(t,x,y) \\
&= \psi(x,y)\,(\partial_t - L_x)\, \Sigma_q(t,x,y) \\
&\quad - \frac{1}{2}\Delta_x \psi(x,y) \Sigma_q(t,x,y) - \nabla_{\nabla_x \psi(x,y)} \Sigma_q(t,x,y) \\
&= -t^q \psi(x,y) \rho(t,x,y) L_x u_q(x,y) \\
&\quad - \frac{1}{2}\Delta_x \psi(x,y)\, \Sigma_q(t,x,y) - \nabla_{\nabla_x \psi(x,y)} \Sigma_q(t,x,y).
\end{aligned}$$

Let $\epsilon > 0$ be chosen such that $\psi(x,y) = 1$ if $d(x,y) \le \epsilon$. It is easy to see for any l that

$$\left\| \partial_t^k \left(\frac{1}{2}\Delta_x \psi(x,y)\, \Sigma_q(t,x,y) + \nabla_{\nabla_x \psi(x,y)} \Sigma_q(t,x,y) \right) \right\|_l = O(e^{-\epsilon/3t}),$$

where $\|f\|_l$ denotes the supremum norm of f along with all of its derivatives in (x,y) up to order l. We also have that

$$\|t^q \psi(x,y) \rho(t,x,y) L_x u_q(x,y)\| \le C t^{q-n/2}.$$

Furthermore, if W is a vector field on M, then by Lemma 65.4,

$$\begin{aligned}
\left| W_x e^{-d^2/2t} \right| &= \left| -\frac{(V,W)}{t} e^{-d^2/2t} \right| \le |W| \frac{d}{t} e^{-d^2/2t} \\
&= \frac{1}{\sqrt{t}} |W| \frac{d}{\sqrt{t}} e^{-d^2/2t} \le \frac{e^{-1/2}}{\sqrt{t}} |W|.
\end{aligned}$$

Similarly we have the same estimate for $\left| W_y e^{-d^2/2t} \right|$. Let us now consider higher order spatial derivatives, for example

$$U_y W_x e^{-d^2/2t} = -U_y \left(\frac{(V,W)}{t} e^{-d^2/2t} \right) = -\left(\frac{U_y(V,W)}{t} e^{-d^2/2t} + \frac{(V,W)}{t} U_y e^{-d^2/}\right.$$

from which we find that

$$\left| U_y W_x e^{-d^2/2t} \right| \le \frac{C}{t} e^{-d^2/2t} + C \frac{d^2}{t^2} e^{-d^2/2t} \le \frac{C}{t}.$$

Continuing in this way we learn that

$$\left\| e^{-d^2/2t} \right\|_l \le C t^{-l/2}.$$

Let us consider the t–derivatives of $\rho(t,x,y)$,

$$|\partial_t \rho| = | - \rho(\frac{n}{2t} + \frac{d^2}{2t^2})| \le C t^{-n/2} t^{-1}.$$

Similarly,

$$|\partial_t^2 \rho| = |\rho(\frac{n}{2t^2} + \frac{d^2}{t^3}) + \rho(\frac{n}{2t} + \frac{d^2}{2t^2})^2| \le C t^{-n/2} t^{-2}.$$

Continuing this way, one learns that $|\partial_t^k \rho| \le CCt^{-(n/2+k)}$. Putting this all together gives Eq. (61.10).

64.2 Proof of Theorem 61.4

Proposition 64.1. *Let $y \in M$ be fixed, $r(x) = d(x,y)$, $J(x) = J(x,y)$, V and $\rho(t,x)$ be as above. Suppose that $g(t,x) : E_y \to E_x$ is a time dependent section of* $\hom(E_y \to E)$ *and $u(t,x) = \rho(t,x) g(t,x)$. Then*

$$\begin{aligned}
(\partial_t - L)\, u &= \rho \left(\partial_t - L + \frac{1}{t} \left(\nabla_V + \frac{1}{2} r \partial \ln J/\partial r \right) \right) g \\
&= \rho \left(\partial_t - L + \frac{1}{t} S \right) g, \qquad (64.1)
\end{aligned}$$

where

$$S = \nabla_V + \frac{1}{2} r \partial \ln J/\partial r. \qquad (64.2)$$

Proof. First let us recall that

$$\begin{aligned}
\square(\rho g) &= \mathrm{tr} \nabla^2 (\rho g) = \mathrm{tr}\left(\nabla^2 \rho\, g + 2 \nabla \rho \otimes \nabla g + \rho \nabla^2 g \right) \\
&= \Delta \rho\, g + 2 \nabla_{\nabla \rho} g + \rho \square g \\
&= \Delta \rho\, g + 2 \rho\, \nabla_{\nabla \ln \rho} g + \rho \square g
\end{aligned}$$

and that

$$\nabla \ln \rho = \nabla \left(-\frac{1}{2t}r^2\right) = -\frac{V}{t}.$$

Hence

$$\begin{aligned}(\partial_t - L)\, u &= \left(\partial_t - \frac{1}{2}\Box\right)(\rho g) - \rho \mathcal{R} g \\ &= \left(\partial_t - \frac{1}{2}\Delta\right)\rho\, g - \rho \nabla_{\nabla \ln \rho} g + \rho\,(\partial_t - L)\, g \\ &= \frac{1}{2t}\,(r\partial \ln J/\partial r)\,\rho\, g + \frac{1}{t}\rho \nabla_V g + \rho\,(\partial_t - L)\, g \\ &= \rho \left(\partial_t - L + \frac{1}{t}\left(\nabla_V + \frac{1}{2} r \partial \ln J/\partial r\right)\right) g.\end{aligned}$$

■

Now let

$$g_q(t,x) = \sum_{k=0}^{q} t^k u_k(x) \text{ and } \Sigma_q(t,x) = \rho(t,x) g_q(t,x) \tag{64.3}$$

where $u_k(x) : E_y \to E_x$ are to be determined. Then

$$\begin{aligned}\left(\partial_t + \frac{1}{t}S - L\right) g &= \sum_{k=0}^{q} \left\{t^{k-1}\,(k u_k + S u_k) - t^k L u_k\right\} \\ &= \frac{1}{t} S u_0 + \sum_{k=0}^{q-1} t^k\left((k+1) u_{k+1} + S u_{k+1} - L u_k\right) - t^q L u_q.\end{aligned}$$

Thus if we choose u_0 such that

$$S u_0(x) = \left(\nabla_V + \frac{1}{2} V \ln J\right) u_0(x) = 0. \tag{64.4}$$

and u_k such that

$$(S + k + 1) u_{k+1} - L u_k = 0 \tag{64.5}$$

then $\left(\partial_t + \frac{1}{t}S - L\right) g = -t^q L u_q$ or equivalently by Eq. (64.1),

$$(\partial_t - L)\, k_q = (\partial_t - L)\,(\rho g) = -t^q \rho L u_q. \tag{64.6}$$

Let us begin by solving (64.4) for u_0. For $x, y \in M$, let $\gamma(t) = \gamma_{x,y}(t) := \exp_y(t \exp_y^{-1}(x))$ so that $\gamma_{x,y}$ is the geodesic connecting y to x. Notice that $V(\gamma(t)) = t\dot\gamma(t)$ and therefore $\nabla_V = t\nabla_{\dot\gamma(t)}$. Therefore, the equation $S u_0 = 0$ is implies that

$$t\frac{\nabla}{dt} u_0(\gamma_{x,y}(t)) + \frac{1}{2} t \left[\frac{\partial}{\partial t} \ln J(\gamma_{x,y}(t))\right] u_0(\gamma_{x,y}(t)) = 0$$

or equivalently that

$$\left(\frac{d}{dt} + \frac{1}{2}\frac{d}{dt} \ln J(\gamma_{x,y}(t))\right)\left\{//_t(\gamma_{x,y})^{-1} u_0(\gamma_{x,y}(t))\right\} = 0. \tag{64.7}$$

We may solve this last equation to find that

$$//_t(\gamma_{x,y})^{-1} u_0(\gamma_{x,y}(t)) = \frac{1}{\sqrt{J(\gamma_{x,y}(t))}} u_0(y)$$

and hence that

$$u_0(x) = //_1(\gamma_{x,y}) \frac{1}{\sqrt{J(x,y)}} u_0(y).$$

Since we are going to want u to be a fundamental solution, it is natural to require the $u_0(y) = Id_{E_y}$. This gives a first order parametrix,

$$k_0(t,x,y) := \frac{1}{\sqrt{J(x,y)}} \rho(t,x,y) \tau(x,y) : E_y \to E_x,$$

where

$$\tau(x,y) := //_1(\gamma_{x,y}) \text{ and } \rho(t,x,y) := \rho(t,x) = (2\pi t)^{-n/2}\, e^{-d^2(x,y)/2t}.$$

This kernel satisfies,

$$(\partial_t - L_x)\, k_0(t,x,y) = -\rho(t,x,y) L_x \left(\frac{1}{\sqrt{J(x,y)}} \tau(x,y)\right).$$

Proposition 64.2. *Let $y \in M$ be fixed and set, for x near y,*

$$u_0(x,y) = //_1(\gamma_{x,y}) \frac{1}{\sqrt{J(x,y)}}. \tag{64.8}$$

Then $u_0(x,y)$ is smooth for (x,y) near the diagonal in $M \times M$ and $S_x u_0(x,y) = 0$.

Proof. Because of smooth dependence of differential equations on initial conditions and parameters, it follows that $u_0(x,y)$ is smooth for (x,y) near the diagonal in $M \times M$. To simplify notation, let $u_0(x) := u_0(x,y)$, $\tau(x) = \tau(x,y)$, and $J(x) = J(x,y)$. We must verify that $S u_0 = 0$. This is seen as follows:

$$\begin{aligned}\nabla_V u_0(x) &= \nabla_V \left(\frac{1}{\sqrt{J(x)}} \tau(x)\right) \\ &= -\frac{1}{2}\frac{1}{\sqrt{J(x)}}\,(V \ln J)\,(x)\tau(x) + \frac{1}{\sqrt{J(x)}} \nabla_V \tau(x) \\ &= \frac{1}{\sqrt{J(x)}}\left\{\nabla_V \tau(x) - \frac{1}{2}\,(V \ln J)\,(x)\tau(x)\right\} \\ &= -\frac{1}{2}\,(V \ln J)\,(x) u_0(x),\end{aligned}$$

since

$$\nabla_V \tau(x) = r(x)\frac{\nabla}{dt}|_1 \tau(\gamma_{x,y}(t)) = r(x)\frac{\nabla}{dt}|_1 //_t(\gamma_{x,y}) = 0.$$

Hence

$$Su_0(x) = \nabla_V u_0(x) + \frac{1}{2}(V \ln J)(x) u_0(x) = 0.$$

■

We now consider solving Eq. (64.5). Fixing x and y and letting $\gamma(t) := \gamma_{x,y}(t)$, Eq. (64.5) may be written as

$$\left(t\frac{\nabla}{dt} + \frac{1}{2}t\frac{d}{dt}\ln J(\gamma(t)) + k + 1\right) u_{k+1}(\gamma(t)) - Lu_k(\gamma(t)) = 0$$

or equivalently that

$$\left(\frac{d}{dt} + \frac{1}{2}\frac{d}{dt}\ln J(\gamma(t)) + \frac{k+1}{t}\right) //_t(\gamma)^{-1} u_{k+1}(\gamma(t)) - \frac{1}{t} //_t(\gamma)^{-1} Lu_k(\gamma(t)) = 0. \tag{64.9}$$

Letting f be a solution to

$$\left(\frac{d}{dt} - \frac{1}{2}\frac{d}{dt}\ln J(\gamma(t)) - \frac{k+1}{t}\right) f(t) = 0$$

it follows that Eq. (64.9) may be written as

$$\frac{d}{dt}\left[f(t) //_t(\gamma)^{-1} u_{k+1}(\gamma(t))\right] - \frac{f(t)}{t} //_t(\gamma)^{-1} Lu_k(\gamma(t)) = 0. \tag{64.10}$$

We now let f be given by

$$\begin{aligned} f(t) &= \exp\left(\int \left[\frac{1}{2}\frac{d}{dt}\ln J(\gamma(t)) + \frac{k+1}{t}\right] dt\right) \\ &= \exp\left(\frac{1}{2}\ln J(\gamma(t)) + (k+1)\ln t\right) = \sqrt{J(\gamma(t))} t^{(k+1)}. \end{aligned}$$

Integrating (64.10) over $[0, t]$ implies that

$$f(t) //_t(\gamma)^{-1} u_{k+1}(\gamma(t)) = \int_0^t \frac{f(\tau)}{\tau} //_\tau(\gamma)^{-1} Lu_k(\gamma(\tau)) d\tau.$$

Evaluating this equation at $t = 1$ and solving for $u_{k+1}(x)$ gives:

$$\begin{aligned} u_{k+1}(x) &= \frac{1}{\sqrt{J(x,y)}} \tau(x,y) \int_0^1 \frac{\sqrt{J(\gamma(t))} s^{(k+1)}}{s} //_s(\gamma)^{-1} Lu_k(\gamma(s)) ds \\ &= \frac{1}{\sqrt{J(x,y)}} \tau(x,y) \int_0^1 s^k \sqrt{J(\gamma_{x,y}(s), y)} //_s(\gamma_{x,y})^{-1} Lu_k(\gamma_{x,y}(s)) ds. \end{aligned} \tag{64.11}$$

Theorem 64.3. *Let $u_0(x, y)$ be given as in Equation (64.8), and define the smooth sections $u_k(x, y)$ inductively by*

$$u_{k+1}(x,y) = u_0(x,y) \int_0^1 s^k u_0(\gamma_{x,y}(s), y)^{-1} L_x u_k(\gamma_{x,y}(s), y) ds \tag{64.12}$$

for $k = 0, 1, 2, \ldots$. Then u_k solves Eq. (64.5).

Proof. Let us begin by noting that Eq. (64.11) and (64.12) are the same equation because

$$u_0(\gamma_{x,y}(s), y) = \frac{1}{\sqrt{J(\gamma_{x,y}(s), y)}} //_s(\gamma_{x,y}).$$

Let $x, y \in M$ be fixed and set $\gamma(t) = \gamma_{x,y}(t)$. Since $\gamma_{y,\gamma(t)}(s) = \gamma(ts)$ and $//_s(\gamma_{y,\gamma(t)}) = //_{ts}(\gamma)$,it follows that

$$\begin{aligned} u_{k+1}(\gamma(t), y) &= u_0(\gamma(t), y) \int_0^1 s^k 1\sqrt{J(\gamma(ts), y)} //_{ts}(\gamma)^{-1} L_x u_k(\gamma(ts), y) ds \\ &= t^{-(k+1)} u_0(\gamma(t), y) \int_0^t r^k \sqrt{J(\gamma(r), y)} //_r(\gamma)^{-1} L_x u_k(\gamma(r), y) dr. \end{aligned}$$

From this equation we learn that

$$\begin{aligned} &(\nabla_V u_0(x) + \frac{1}{2}(V \ln J)(x) + k + 1) u_{k+1}(x,y) \\ &= \left(t\frac{\nabla}{dt} + \frac{1}{2}t\frac{d}{dt}\ln J(\gamma(t)) + k + 1\right)|_{t=1} u_{k+1}(\gamma(t)) \\ &= u_0(x,y)\frac{d}{dt}|_1 \int_0^t r^k \sqrt{J(\gamma(r), y)} //_r(\gamma)^{-1} L_x u_k(\gamma(r), y) dr \\ &= u_0(x,y)\sqrt{J(\gamma(r), y)} //_1(\gamma)^{-1} L_x u_k(x,y) = L_x u_k(x,y), \end{aligned}$$

wherein the second equality we have used the product rule and the fact that

$$\left(t\frac{\nabla}{dt} + \frac{1}{2}t\frac{d}{dt}\ln J(\gamma(t)) + k + 1\right) t^{-(k+1)} u_0(\gamma(t), y) = 0$$

which is verified using Eq. (64.7). ■

65

Appendix: Gauss' Lemma & Polar Coordinates

Lemma 65.1 (Gauss' Lemma). *Let $y \in M, v, w \in T_yM$, then*

$$\left(\exp_{*y} v_v, \exp_{*y} w_v\right)_{\exp(v)} = (v, w)_y$$

Proof. Let $\Sigma(t,s) := \exp(t(v+sw))$, then

$$\begin{aligned}\frac{d}{dt}(\dot{\Sigma}(t,0), \Sigma'(t,0)) &= (\frac{\nabla}{dt}\dot{\Sigma}(t,0), \Sigma'(t,)) + (\dot{\Sigma}(t,0), \frac{\nabla}{dt}\Sigma'(t,0)) \\ &= (\dot{\Sigma}(t,0), \frac{\nabla}{ds}|_0\dot{\Sigma}(t,s)) = \frac{1}{2}\frac{d}{ds}|_0(\dot{\Sigma}(t,s), \dot{\Sigma}(t,s)) \\ &= \frac{1}{2}\frac{d}{ds}|_0 \left|v+sw\right|^2 = (v,w).\end{aligned}$$

Combining this equation with the observation that $(\dot{\Sigma}(t,0), \Sigma'(t,0))|_{t=0} = 0$ implies that

$$(\dot{\Sigma}(1,0), \Sigma'(1,0)) = (v,w)_0.$$

■

Corollary 65.2. *Suppose that $y \in M$ and choose $\delta > 0$ such that $\exp_y$ is a diffeomorphism on $B(0_y, \delta)$. Then $d(x,y) = \left|\exp_y^{-1}(x)\right|$ for all $x \in V :=$* $\exp_y(B(0_y,\delta))$.

Proof. Let $\sigma(t)$ be a curve in M such that $\sigma(0) = y$ and $\sigma(1) = x$. Suppose for the moment that $\sigma(t)$ is contained in V and write $\sigma(t) = \exp_y(c(t))$. Set $u = c(1)/\left|c(1)\right|$ and decompose $c(t) = (c(t),u)u + d(t)$ where $(d(t),u) = 0$.Then

$$\begin{aligned}\left|\dot{\sigma}(t)\right|^2 &= \left|\exp_{y*}\dot{c}(t)_{c(t)}\right|^2 = \left|\exp_{y*}\left((\dot{c}(t),u)u_{c(t)} + \dot{d}(t)_{c(t)}\right)\right|^2 \\ &= \left|\exp_{y*}(\dot{c}(t),u)u_{c(t)}\right|^2 + \left|\exp_{y*}\left(\dot{d}(t)_{c(t)}\right)\right|^2 \\ &= \left|(\dot{c}(t),u)\right|^2 + \left|\exp_{y*}\left(\dot{d}(t)_{c(t)}\right)\right|^2 \geq \left|(\dot{c}(t),u)\right|^2.\end{aligned}$$

From this we learn that

$$\text{Length}(\sigma) = \int_0^1 \left|\dot{\sigma}(t)\right| dt \geq \int_0^1 \left|(\dot{c}(t),u)\right| dt \geq \int_0^1 (\dot{c}(t),u)dt = \left|c(1)\right|.$$

That is $\text{Length}(\sigma) \geq \left|\exp_y^{-1}(x)\right|$. It is easily to use the same argument to show that if σ leaves the open set V then $\text{Length}(\sigma) \geq \delta > \left|\exp_y^{-1}(x)\right|$ and hence $\text{Length}(\sigma) \geq \left|\exp_y^{-1}(x)\right|$ for all path σ such that $\sigma(0) = y$ and $\sigma(1) = x$. Moreover we have equality if $\sigma(t)$ is the geodesic joining y to x. This shows that

$$d(x,y) = \inf_\sigma \text{Length}(\sigma) = \left|\exp_y^{-1}(x)\right|.$$

■

For more on geodesic coordinates, see Appendix 67.

65.1 The Laplacian of Radial Functions

Lemma 65.3. *Let $r(x) := d(x,y)$ and $J(x) = J(x,y)$ as in Definition 61.1. Then*

$$\Delta f(r) = \frac{\partial\left(Jr^{n-1}f'(r)\right)/\partial r}{Jr^{n-1}} = f''(r) + \left(\frac{n-1}{r} + \frac{\partial \ln J}{\partial r}\right)f'(r). \tag{65.1}$$

We also have that

$$\Delta f(r) = f''(r) + \frac{\nabla\cdot V - 1}{r}f'(r) \tag{65.2}$$

and that

$$\Delta r^2 = 2\nabla\cdot V = 2\left(n + r\frac{\partial \ln J}{\partial r}\right).$$

Proof. We will give two proofs of this result. For the first proof recall that if $\left\{z^i\right\}$ is a chart on M, then

$$\Delta F = \frac{1}{\sqrt{g}}\frac{\partial}{\partial z^i}\left(\sqrt{g}g^{ij}\frac{\partial}{\partial z^j}F\right), \tag{65.3}$$

where $ds^2 = g_{ij}dz^idz^j$, g^{ij} is the inverse of (g_{ij}) and $\sqrt{g} = \sqrt{\det(g_{ij})}$ We now choose the coordinate system z to be $z^n := r$, and $z^i := \alpha^i \circ \theta$, where $\left\{\alpha^i\right\}_{i=1}^{n-1}$ is a chart on $S = S_y \subset T_yM$. We now need to compute g_{ij} in this case. Let us begin by noting that $\partial/\partial z^i = \exp_{y*}\left(r\frac{\partial}{\partial\alpha^i}|_\theta\right)$ for $i = 1, 2, \ldots, n-1$ and $\partial/\partial z^n = \partial/\partial r = \exp_{y*}\left(\frac{d}{dt}|_0(r+t)\theta\right)$. By Gauss' lemma, it follows that $ds^2 = dr^2 + h_{ij}(r,\theta)d\theta^i d\theta^j$, i.e. $g = \begin{bmatrix} h & 0 \\ 0 & 1\end{bmatrix}$. Therefore $g^{-1} = \begin{bmatrix} h^{-1} & 0 \\ 0 & 1\end{bmatrix}$ and $\sqrt{g} = \sqrt{h}$. So if $F = f(r) = f(z^n)$, it follows from Eq. (65.3) that

$$\begin{aligned}\Delta f(r) &= \frac{1}{\sqrt{g}}\frac{\partial}{\partial z^i}\left(\sqrt{g}g^{in}\frac{\partial}{\partial z^n}F\right) = \frac{1}{\sqrt{h}}\frac{\partial}{\partial z^n}\left(\sqrt{h}\frac{\partial}{\partial z^n}F\right) \\ &= f''(r) + \frac{\partial \ln(\sqrt{h})}{\partial r} f'(r). \end{aligned} \tag{65.4}$$

So to finish the proof we need to describe $\sqrt{g} = \sqrt{h}$ in terms of J. In order to do this, let us notice that D in Eq. (61.4) may also be expressed as

$$\begin{aligned} D(v) &= \frac{\lambda(\exp_*(w_1)_v, \exp_*(w_2)_v \dots, \exp_*(w_n)_v)}{\lambda_y(w_1, w_2, \dots, w_n)} \\ &= \sqrt{\frac{\det\left\{(\exp_{y*}(w_i)_v, \exp_{y*}(w_j)_v)\right\}_{i,j=1}^n}{\det\left\{(w_i, w_j)\right\}_{i,j=1}^n}}, \end{aligned} \tag{65.5}$$

where now $\{w_i\}$ is any oriented basis for T_yM. From this expression it follows that

$$\sqrt{g} = D(r\theta)\sqrt{\det\left\{\left(r\frac{\partial}{\partial \alpha^i}|_\theta, r\frac{\partial}{\partial \alpha^j}|_\theta\right)\right\}_{i,j=1}^{n-1}} = J(\cdot, y)r^{n-1}K(\theta)$$

where

$$K(\theta) := \sqrt{\det\left\{\left(\frac{\partial}{\partial \alpha^i}|_\theta, \frac{\partial}{\partial \alpha^j}|_\theta\right)\right\}_{i,j=1}^{n-1}}.$$

Using these expression in Eq. (65.4) along with the observation that $\partial K(\theta)/\partial r = 0$ proves Eq. (65.1).

(Second more direct Proof.) Let (ρ, ω) denote a generic point in $\mathbb{R}_+ \times S$ and $d\omega$ denote the volume form on $S = S_y \subset T_yM$. Then

$$\begin{aligned}\int_M f(r,\theta) d\lambda &= \int_{T_yM} \tilde{f} D \, d\lambda_y \\ &= \int_{\mathbb{R}_+ \times S} f(\rho, \omega) D(\rho\omega)\rho^{n-1} d\rho d\omega, \end{aligned}$$

where $\tilde{f}(\rho\omega) := f(\rho, \omega)$. Therefore,

$$\begin{aligned}\int_M \Delta f(r) g(r,\theta) d\lambda &= -\int_M (\nabla f(r), \nabla g(r,\theta))\, d\lambda \\ &= -\int_M f'(r)\partial g(r,\theta)/\partial r \, d\lambda \\ &= -\int_{\mathbb{R}_+ \times S} f'(\rho)\partial g(\rho,\omega)/\partial\rho \, D(\rho\omega)\rho^{n-1} d\rho d\omega \\ &= \int_{\mathbb{R}_+ \times S} \frac{\partial\left(f'(\rho)D(\rho\omega)\rho^{n-1}\right)/\partial\rho}{D(\rho\omega)\rho^{n-1}} g(\rho,\omega)\, D(\rho\omega)\rho^{n-1} d\rho d\omega \\ &= \int_M \frac{\partial\left(f'(r)Jr^{n-1}\right)/\partial r}{Jr^{n-1}} g(r,\theta)\, d\lambda, \end{aligned}$$

which proves Eq. (65.1).

To prove Eq. (65.2), we compute more directly:

$$\begin{aligned}\Delta f(r) &= \nabla\cdot\nabla f(r) = \nabla\cdot(f'(r)\partial/\partial r) = \nabla\cdot\left(\frac{f'(r)}{r}V\right) \\ &= \left(\frac{\partial}{\partial r}\frac{f'(r)}{r}\right)\left(\frac{\partial}{\partial r}, V\right) + \frac{f'(r)}{r}\nabla\cdot V \\ &= f''(r) - \frac{f'(r)}{r} + \frac{f'(r)}{r}\nabla\cdot V \end{aligned}$$

which proves Eq. (65.2).

In particular we have that $\Delta r^2 = 2\nabla\cdot V$ and

$$\Delta r^2 = 2 + \left(\frac{n-1}{r} + \frac{\partial \ln J}{\partial r}\right)2r = 2\left(n + r\frac{\partial \ln J}{\partial r}\right)$$

■

Lemma 65.4. *Let $r(x) := d(x,y)$ as above, then*

$$Vr^2(x) = 2r^2(x),$$

$$\nabla r^2 = 2V$$

and

$$\nabla\cdot V(y) = n.$$

Proof. We first claim that $|V| = r$. Moreover,

$$Vr^2(x) = \frac{d}{dt}|_1 r^2(\gamma_x(t)) = \frac{d}{dt}|_1 t^2 r^2(x) = 2r^2(x).$$

That is to say $Vr^2 = 2r^2$. Moreover, if $w \in T_xM$ is perpendicular to $V(x)$, then $wr^2 = 0$ so that ∇r^2 is proportional to V. One way to argue this is that $V(x)$ points in the direction of maximum increase of r^2 by the triangle inequality. Hence

$$\nabla r^2 = \frac{(\nabla r^2, V)}{(V,V)}V = \frac{2r^2}{r^2}V = 2V.$$

If we do not like this explanation, then use Gauss's lemma I guess. Come back to this point.

Now we wish to compute $\Delta r^2 = 2\nabla\cdot V$. We would like to at least do this at $x = y$. To this end, let us work out $\nabla_w V$ for $w \in T_yM$. Setting $\sigma(t) = \exp(tw)$, we find that

$$\nabla_w V = \frac{d}{dt}|_0 //_t^{-1}(\sigma)V(\sigma(t)) = \frac{d}{dt}|_0 //_t^{-1}(\sigma)//_t(\sigma)tw = w.$$

Therefore $\nabla\cdot V(y) = \sum_{i=1}^n (\nabla_{e_i}V, e_i) = n$ and hence $\Delta r^2(y) = 2n$. ■

66

The Dirac Equation a la Roe's Book

In this section, we consider the Dirac equation:

$$\partial_t S = iDS \text{ with } S_{t=0} = S_0 \text{ given.} \tag{66.1}$$

Here $D = \gamma_{e_i} \nabla_{e_i}$ is the Dirac operator on some spinor bundle over M. The most interesting statement made by Roe about the Dirac equation is it's finite speed of propagation property. Given a compact region $\Omega \subset M$ with smooth boundary and a solution to Eq. (66.1), let

$$E_\Omega(t) := \int_\Omega |S_t(x)|^2 d\lambda(x)$$

be the energy of S_t in the region Ω. Let us begin by computing the derivative of E_Ω.

Lemma 66.1. *Let $\Omega \subset M$ be a compact region with smooth boundary, and S be a solution to Eq. (66.1). Then*

$$\frac{d}{dt} E_\Omega(t) = i \int_\Omega \nabla \cdot X_t \, d\lambda = i \int_{\partial\Omega} (X_t, N) \, d\sigma, \tag{66.2}$$

where X_t is the smooth vector field on M such that $(X_t, Y) = (\gamma_Y S_t, S_t)$ for all vector fields Y on M and N is the outward pointing normal to Ω and σ is surface measure on Ω_t.

Proof. Differentiating under the integral sign implies that

$$\begin{aligned}
\frac{d}{dt} E_\Omega(t) &= \int_\Omega \left\{ \left(\dot{S}_t, S_t\right) + \left(S_t, \dot{S}_t\right) \right\} d\lambda \\
&= \int_\Omega \{(iDS, S) + (S, iDS)\} \, d\lambda \\
&= i \int_\Omega \{(DS, S) - (S, DS)\} \, d\lambda \qquad (66.3) \\
&= - \int_\Omega \operatorname{Im}(DS, S) \, d\lambda \qquad (66.4)
\end{aligned}$$

We also have,

$$\begin{aligned}
(\nabla_* X, \cdot) &= \nabla_*(X, \cdot) = \nabla_*(\gamma.S, S) \\
&= ((\nabla_* \gamma).S, S) + (\gamma.\nabla_* S, S) + (\gamma.S, \nabla_* S) \\
&= (\gamma.\nabla_* S, S) - (S, \gamma.\nabla_* S),
\end{aligned}$$

where we have made use of the fact that $\nabla \gamma = 0$ and $\gamma_X^* = -\gamma_X$. Taking $* = e_i$ and $\cdot = e_i$ and summing on i in the above equation implies that

$$\nabla \cdot X = (DS, S) - (S, DS). \tag{66.5}$$

Combining Eq. (66.3) and (66.5) along with the divergence (Stoke's) theorem proves Eq. (66.2). ■

Corollary 66.2. *The total energy $E(t) = E_M(t)$ remains constant and solutions to Eq. (66.1) are unique if they exist.*

We now want to examine how $E_\Omega(t)$ depends on Ω.

Lemma 66.3. *Suppose that $\Omega \subset M$ is as above and $\phi_t : \Omega \to M$ is a one parameter family of smooth injective local diffeomorphisms depending smoothly on t and let $\Omega_t := \phi_t(\Omega)$. Also define a vector field Y_t on Ω_t by $\dot{\phi}_t = Y_t \circ \phi_t$, i.e. $Y_t := \dot{\phi}_t \circ \phi_t^{-1}$. If $f : M \to \mathbb{R}$ is a smooth function, then*

$$\frac{d}{dt} \int_{\Omega_t} f d\lambda = \int_{\Omega_t} \nabla \cdot (f Y_t) \, d\lambda = \int_{\partial\Omega_t} f(Y_t, N) \, d\sigma, \tag{66.6}$$

where again N is the outward pointing normal to $\partial\Omega_t$ and σ is surface measure on Ω_t.

Proof. Since

$$\begin{aligned}
\int_{\Omega_t} f d\lambda &= \int_{\phi_t(\Omega)} f\lambda = \int_\Omega \phi_t^* (f\lambda) = \int_\Omega f \circ \phi_t \, \phi_t^* \lambda, \\
&= \int_\Omega f \circ \phi_t \, \phi_t^* \lambda
\end{aligned}$$

$\frac{d}{dt} f \circ \phi_t = Y_t f \circ \phi_t$ and

$$\frac{d}{dt} \phi_t^* \lambda = \phi_t^* \left((d i_{Y_t} + i_{Y_t} d) \lambda \right) = \phi_t^* \left(\nabla \cdot Y_t \lambda \right),$$

it follows that

$$\begin{aligned}
\frac{d}{dt} \int_{\Omega_t} f d\lambda &= \int_\Omega Y_t f \circ \phi_t \, \phi_t^* \lambda + \int_\Omega f \circ \phi_t \, \phi_t^* \left(\nabla \cdot Y_t \lambda \right) \\
&= \int_{\Omega_t} Y_t f \, \lambda + \int_{\Omega_t} f \, \nabla \cdot Y_t \, \lambda \\
&= \int_{\Omega_t} \nabla \cdot (f Y_t) \, \lambda
\end{aligned}$$

from which Eq. (66.6) follows. ■

Fix a point $m \in M$ and let $B(m,R)$ be the geodesic ball centered at $m \in M$ with radius R. If we believe that the speed of propagation of the Dirac equation is 1, then we should have

$$e(t) := \int_{B(m,R-t)} |S_t|^2 \, d\lambda \tag{66.7}$$

is non-increasing as t increases to R. The reason is that, we are shrinking the ball at a rate equal to the speed of propagation, so no energy which was in the wave at time T outside the ball $B(m,R-T)$ can enter the region $B(m,R-t)$ for $t \le T$. We will now verify that $e(t)$ is non-increasing.

Proposition 66.4. *For R smaller than the injectivity radius of M, the function $e(t)$ in Eq. (66.7) is non-increasing as t increases to R.*

Proof. Let $\phi_t : T_mM \to M$ be given by $\phi_t(v) = \exp(tv)$ and Y_t be the locally defined vector field on M such that $\dot{\phi}_t(v) = Y_t \circ \phi_t(v)$ for all v small. Since $\phi_t(D) = B(m,t)$, where D is the unit disc in T_mM, we have that $(Y_t, N) = 1$, by Gauss's lemma. So By Lemmas 66.1 and 66.3,

$$\begin{aligned}\frac{d}{dt}e(t) &= i\int_{\partial B(m,R-t)} (X_t, N)\, d\sigma - \int_{\partial B(m,R-t)} |S_t|^2 (Y_{R-t}, N)\, d\sigma \\ &= i\int_{\partial B(m,R-t)} (X_t, N)\, d\sigma - \int_{\partial B(m,R-t)} |S_t|^2 \, d\sigma.\end{aligned}$$

Now

$$|(X_t, N)| = |(\gamma_N S, S)| \le |\gamma_N S|\,|S| \le |S|^2$$

since N is a unit vector and γ_N is an isometry. (Recall that γ_N is skew adjoint and $\gamma_N^2 = -I$.). This shows that

$$\begin{aligned}\frac{d}{dt}e(t) &= -\operatorname{Im}\int_{\partial B(m,R-t)} (X_t, N)\, d\sigma - \int_{\partial B(m,R-t)} |S_t|^2 \, d\sigma \\ &\le \int_{\partial B(m,R-t)} |S|^2 \, d\sigma - \int_{\partial B(m,R-t)} |S_t|^2 \, d\sigma = 0.\end{aligned}$$

■

Corollary 66.5. *Suppose that the support of S_0 is contained in Ω. Then the support of S_t is contained in*

$$\Omega_t := \{x \in M : d(x,m) < t \text{ for all } m \in \Omega\}.$$

Proof. By repeating the argument and using the semi-group property of e^{itD}, we may and do assume that t is positive and less than the injectivity radius of M. Let $x \notin \bar{\Omega}_t$, so that there exists $R > t$ such that $B(x,R) \cap \Omega = \emptyset$. By the previous proposition, $e(\tau) := \int_{B(x,R-\tau)} |S_\tau|^2 \, d\lambda$ is decreasing and hence

$$\int_{B(x,R-t)} |S_t|^2 \, d\lambda = e(t) \le e(0) = \int_{B(x,R)} |S_0|^2 \, d\lambda = 0.$$

This shows that $S_t \equiv 0$ on $B(x,R-t)$ and in particular at x. ■

66.1 Kernel Construction

Lemma 66.6. *Suppose that $U \subset \mathbb{R}^N$ is a open set and $A : L^2(E) \to C^{r+1}(U)$ is bounded linear map. For each $x \in U$, let $T_x \in L^2(E)$ such that $(T_x, S) = AS(x)$ for all $S \in L^2(E)$. Then the map $x \in U \to T_x \in L^2(E)$ is C^r – smooth. Moreover, we have estimates of the derivatives of $x \to T_x$ in terms of the operator norm $\|A\|_{op}$ of A as an operator from $A : L^2(E) \to C^{r+1}(U)$.*

Remark 66.7. The above Lemma may has well been formulated with $L^2(E)$ replaced by an abstract Hilbert space H. The proof given below would still go through without any change.

Proof. First notice that

$$\begin{aligned}|(T_x - T_y, S)| &= |AS(x) - AS(y)| \le |\nabla AS|_U |x-y| \\ &\le |AS|_{C^1(U)}|x-y| \le C|S|_{L^2(E)}|x-y|,\end{aligned}$$

which shows that $|T_x - T_y|_{L^2(E)} \le C|x-y|$, so the $T.$ is continuous. Let us now consider the directional derivatives of T_x. For $x \in U$ and $v \in \mathbb{R}^N$, let $BS(x,v) := \partial_v AS(x)$. As above there exists $T_{x,v} \in L^2(E)$ such $(T_{x,v}, S) = \partial_v AS(x) = BS(x,v)$ for all $S \in L^2(E)$ and moreover $(x,v) \to T_{x,v}$ is locally Lipschitz continuous and linear in v. Indeed,

$$\begin{aligned}|(T_{x,v} - T_{y,v}, S)| &= |\partial_v AS(x) - \partial_v AS(y)| \le |AS|_{C^2(U)} |x-y||v| \\ &\le \|A\|_{op} |S|_{L^2(E)}|x-y||v|.\end{aligned}$$

That is to say, $x \in U \to T_{x,\cdot} \in B(\mathbb{R}^N, L^2(E))$ is a Lipschitz continuous map.

Now let $x \in U$ and $v \in \mathbb{R}^N$, then

$$\begin{aligned}(T_{x+v} - T_x, S) = AS(x+v) - AS(x) &= \int_0^1 \partial_v AS(x+tv)dt \\ &= \int_0^1 (T_{x+tv,v}, S)dt,\end{aligned}$$

which shows that

$$\begin{aligned}|T_{x+v} - T_x - T_{x,v}| &\le \int_0^1 |T_{x+tv,v} - T_{x,v}|\, dt \le C|v|\int_0^1 |tv|\, dt \\ &= \|A\|_{op} |v|^2/2.\end{aligned}$$

This shows that T_x is differentiable and that $T_x' v = T_{x,v}$. We have already seen that T_x' is continuous. This shows that T_x is C^1. We may continue this way inductively to finish the proof of the lemma. ■

Proposition 66.8. *Suppose that $A : L^2(E) \to C^{r+1}(E)$ is a bounded operator, then AA^* has an integral kernel which is C^r – smooth. Moreover the C^r – norm of the kernel is bounded by the square of the operator norm for $A : L^2(E) \to C^{r+1}(E)$.*

Proof. Let $x \in M$ the map $S \in L^2(E) \to AS(x) \in E_x$ is a bounded linear map and hence there is a unique element $T(x,\cdot) \in L^2(End(E, E_x))$ such that

$$AS(x) = \int_M T(x,y)S(y)dy.$$

Notice that if $\xi \in \Gamma(E)$, then

$$\begin{aligned}(\xi(x), AS(x)) &= \int_M (\xi(x), T(x,y)S(y))\, dy \\ &= \int_M (T^*(x,y)\xi(x), S(y))\, dy,\end{aligned}$$

which by the previous lemma shows that $x \to T^*(x,\cdot)\xi(x) \in L^2(E)$ is a C^r – map with bounds determined by $\|A\|_{op}$.

(Surely one can show that there is a version of $T(x,y)$ such that $(x,y) \to T(x,y)$ is jointly measurable. We will avoid this issue here however.) Ignoring measurability issues, we know that

$$A^*S(x) = \int_M T^*(y,x)S(y)dy$$

so the

$$\begin{aligned}AA^*S(x) &= \int_{M\times M} T(x,y)T^*(z,y)S(z)dydz \\ &= \int_M k(x,z)S(z)dz,\end{aligned}$$

where

$$k(x,z) := \int_M T(x,y)T^*(z,y)dy.$$

Even though the derivation of k above was suspect because of measurability questions, the formula make perfect sense. Indeed suppose that ξ and η are in $\Gamma(E)$, then

$$\begin{aligned}(\eta(x), k(x,z)\xi(z))_{E_x} &= \int_M (\eta(x), T(x,y)T^*(z,y)\xi(z))_{E_x}\, dy \\ &= \int_M (T^*(x,y)\eta(x), T^*(z,y)\xi(z))_{E_y}\, dy \\ &= (T^*(x,\cdot)\eta(x), T^*(z,\cdot)\xi(z))_{L^2(E)}.\end{aligned}$$

Furthermore this shows that $(x,z) \to (\eta(x), k(x,z)\xi(z))_{E_x}$ is C^r in (x,y) with a C^r norm which is controlled by the C^r – norms of η, ξ, and $\|A\|_{op}^2$. Since η and ξ are arbitrary, we find that $k(x,z)$ is C^r as well and the C^r – norm is bounded by a constant times $\|A\|_{op}^2$.

So the only thing left to check is that

$$AA^*S(x) = \int_M k(x,z)S(z)d\lambda(z).$$

Letting $\xi, \eta \in \Gamma(E)$ as before, then

$$\begin{aligned}(\eta(x), k(x,z)\xi(z))_{E_x} &= \int_M (\eta(x), T(x,y)T^*(z,y)\xi(z))_{E_x}\, dy \\ &= (\eta(x), (AT^*(z,\cdot)\xi(z))\,(x))_{E_x}\end{aligned}$$

so that

$$\begin{aligned}\int_M (\eta(x), k(x,z)\xi(z))_{E_x}\, dx &= \int_M (\eta(x), (AT^*(z,\cdot)\xi(z))\,(x))_{E_x}\, dx \\ &= \int_M (A^*\eta(x), T^*(z,x)\xi(z))_{E_x}\, dx \\ &= \int_M (T(z,x)A^*\eta(x), \xi(z))_{E_x}\, dx.\end{aligned}$$

Integrating this last expression over z shows that

$$\int_{M\times M} (\eta(x), k(x,z)\xi(z))_{E_x}\, dxdz = (AA^*\eta, \xi)_{L^2(E)} = (\eta, AA^*\xi)_{L^2(E)}.$$

Since η is arbitrary we conclude that

$$AA^*\xi(x) = \int_M k(x,z)\xi(z)dz.$$

■

Using the above results, one can show that $f(D)$ has a smooth kernel for any function $f : \mathbb{R} \to \mathbb{C}$ which has rapid decrease. To see this, by writing f in its real and imaginary parts, we may assume that f is real valued. Furthermore, by decomposing f into its positive and negative parts we may assume that $f \geq 0$. Let $g = f^{1/2}$, a function with rapid decrease still, we see that $g(D)$ is a self-adjoint smoothing operator. Therefore $f(D) = g^2(D)$ has a smooth integral kernel. In this way we find that $e^{-tD^2/2}$ has a smooth integral kernel. Let $k_t(x,y) = e^{-tD^2/2}(x,y)$ denote the smooth kernel.

Proposition 66.9. *The function $k_t(x,y) \to 0$ in C^∞ as $t \to 0$ off of any neighborhood of the diagonal $x = y$ in $M \times M$.*

Proof. Let $\delta > 0$ be given, and let ϕ and ψ be smooth functions on $\mathbb{R}$ such that $\phi + \psi = 1$, the support of ϕ is contained in $(-\delta, \delta)$ and ψ is supported in $\{x : |x| \geq \delta/2\}$. Also let $p_t(\lambda) = (2\pi t)^{-1/2} e^{-\lambda^2/2t}$. Then

$$e^{-tD^2/2} = \int_{\mathbb{R}} p_t(\lambda) e^{i\lambda D} d\lambda = \int_{\mathbb{R}} p_t(\lambda)\phi(\lambda) e^{i\lambda D} d\lambda + \int_{\mathbb{R}} p_t(\lambda)\psi(\lambda) e^{i\lambda D} d\lambda.$$

This can be written as $e^{-tD^2/2} = h_t(D) + g_t(D)$, where $h_t(\xi) := \int_{\mathbb{R}} p_t(\lambda)\phi(\lambda) e^{i\lambda\xi} d\lambda$ and $g_t(\xi) = \int_{\mathbb{R}} p_t(\lambda)\psi(\lambda) e^{i\lambda\xi} d\lambda$. Now we notice that

$$\begin{aligned} |\xi^n g_t(\xi)| &= \left| i^{-n} \int_{\mathbb{R}} p_t(\lambda)\psi(\lambda) \partial_\lambda^n e^{i\lambda\xi} d\lambda \right| \\ &= \left| i^n \int_{\mathbb{R}} e^{i\lambda\xi} \partial_\lambda^n \left(p_t(\lambda)\psi(\lambda)\right) d\lambda \right| \\ &\leq K_n(t) \end{aligned}$$

where $\lim_{t\downarrow 0} K_n(t) = 0$ for each n. From this it follows for any n that $D^n g_t(D) : L^2(E) \to L^2(E)$ tends to zero in the operator norm as $t \to 0$. This fact, elliptic regularity, and the Sobolev embedding theorems implies that $g_t(D) : L^2(E) \to C^r(E)$ tends to zero in operator norm for any $r \geq 0$. Using the previous proposition, this shows that the integral kernel of $g_t(D)$ goes to 0 in C^∞. (Note, Roe proves some of this by appealing to the closed graph theorem for Frechet spaces.) Finally, $h_t(D) = \int_{\mathbb{R}} p_t(\lambda)\phi(\lambda) e^{i\lambda D} d\lambda$ is an operator which does not increase the support of a section by more than size δ. This implies that the support of the integral kernel of $h_t(D)$ is contained $\{(x, y) : d(x, y) < 3\delta\}$. Since δ is arbitrary, we are done. ■

66.2 Asymptotics by Sobolev Theory

Let me end this section by explaining how Roe shows that the formal asymptotic expansions of the heat kernel are close to the heat kernel.

Let $\xi \in E_m$ and let $w_t(x) := e^{tL}(x, m)\xi$, then

$$\left(\partial_t + D^2\right) w_t = 0 \text{ and } \lim_{t\downarrow 0} w_t = \xi\delta_m. \tag{66.8}$$

Conversely, if w_t solves Eq. (66.8), then for all smooth sections S of E, $\partial_t (e^{-(T-t)D^2} S, w_t) = 0$. Therefore

$$\begin{aligned} (S, w_T) &= \lim_{t\downarrow 0} (e^{-(T-t)D^2} S, w_t) = \left(\left(e^{-TD^2} S\right)(m), \xi\right) \\ &= \int \left(S(x), \left[e^{-TD^2}(m, x)\right]^* \xi\right) dx = \int \left(S(x), e^{-TD^2}(x, m)\xi\right) dx, \end{aligned}$$

which shows that $w_t(x) := e^{-tD^2}(x, m)\xi$. since S is arbitrary.

Now suppose that S_t is an approximate fundamental solution at $\xi_m \in E_m$, so that

$$\left(\partial_t + D^2\right) S_t = t^k r_t \text{ and } \lim_{t\downarrow 0} S_t = \xi\delta_m$$

where r_t is a smooth section of $\Gamma(E)$. Let α_t denote the solution to

$$\left(\partial_t + D^2\right) \alpha_t = t^k r_t \text{ with } \alpha_0 = 0,$$

which can be written by du Hamel's principle as $\alpha_t = \int_0^t e^{-(t-\tau)D^2} \tau^k r_\tau d\tau$. Since $w_t := S_t - \alpha_t$ satisfies

$$\left(\partial_t + D^2\right) w_t = 0 \text{ with } \lim_{t\downarrow 0} w_t = \xi\delta_m,$$

we find the $w_t = e^{-tD^2}(m, \cdot)$. Therefore, for any $k > n/2$,

$$\begin{aligned} |e^{-tD^2}(m, \cdot) - S_t| &= |\alpha_t| = \left| \int_0^t e^{-(t-\tau)D^2} \tau^k r_\tau d\tau \right| \\ &\leq C_k \left\| \int_0^t e^{-(t-\tau)D^2} \tau^k r_\tau d\tau \right\|_{L^2_k(E)} \\ &\leq C_k \left\| D^k \int_0^t e^{-(t-\tau)D^2} \tau^k r_\tau d\tau \right\|_{L^2(E)} + C_k \left\| \int_0^t e^{-(t-\tau)D^2} \tau^k r_\tau d \right. \end{aligned}$$

wherein the second to last inequality we have used the Sobolev embedding theorem and in the last we use elliptic regularity. Since e^{-tD^2} is a bounded operator, we find that

$$\begin{aligned} |e^{-tD^2}(m, \cdot) - S_t| &\leq C_k \left\| \int_0^t e^{-(t-\tau)D^2} \tau^k D^k r_\tau d\tau \right\|_{L^2(E)} + C_k \left\| \int_0^t e^{-(t-\tau)D^2} \tau^k r_\tau d \right. \\ &\leq C_k t^{k+1} \sup_{0\leq\tau\leq t} \left\{ \left\| D^k r_\tau \right\|_{L^2(E)} + \|r_\tau\|_{L^2(E)} \right\} = K t^{k+1}. \end{aligned}$$

67

Appendix: VanVleck Determinant Properties

67.1 Proof of Lemma 61.3

The first step is to get a more explicit expression for D. To this end, fix $v \in T_yM$ and for any $w \in T_yM$, let

$$Y_w(t) := \frac{d}{ds}|_0 \exp(t(v+sw)) = \exp_*(tw)_{tv} \tag{67.1}$$

and $\gamma(t) = \exp(tv)$. Then Y_w solves Jacobi's equation:

$$\begin{aligned}\frac{\nabla^2}{dt^2}Y_w(t) &= \frac{\nabla^2}{dt^2}\frac{d}{ds}|_0 \exp(t(v+sw)) \\ &= \frac{\nabla}{dt}\frac{\nabla}{ds}|_0 \frac{d}{dt}\exp(t(v+sw)) \quad \text{(No Torsion)} \\ &= \left[\frac{\nabla}{dt}, \frac{\nabla}{ds}|_0\right]\frac{d}{dt}\exp(t(v+sw)) \quad \left(\frac{\nabla}{dt}\frac{d}{dt}\exp(t(v+sw)) = 0\right) \\ &= R(\dot\gamma(t), Y_w(t))\dot\gamma(t). \quad \text{(Definition of } R\text{)}\end{aligned}$$

Y_w obeys the initial conditions

$$Y_w(0) = \frac{d}{ds}|_0 \exp(0(v+sw)) = 0$$

and

$$\begin{aligned}\frac{\nabla}{dt}|_0 Y_w(0) &= \frac{\nabla}{ds}|_0 \frac{d}{dt}|_0 \exp(t(v+sw)) \\ &= \frac{\nabla}{ds}|_0 (v+sw) = w.\end{aligned}$$

Notice that $Y_w(t) \in T_{\gamma(t)}M$ for each t, so by using parallel translation $u(t) := //_t(\gamma)$ along γ we may pull this back to T_yM. Set

$$Z_w(t) = u^{-1}(t)Y_w(t).$$

Using the fact that $\frac{\nabla}{dt}u(t) = 0$, the previous equations imply that Z_w satisfies:

$$\ddot Z_w(t) = A_\gamma(t)Z_w(t) \text{ with } Z_w(0) = 0 \text{ and } \dot Z_w(0) = w, \tag{67.2}$$

where

$$\begin{aligned}A_\gamma(t)w &= u^{-1}(t)R(\dot\gamma(t), u(t)w)\dot\gamma(t) \\ &= u^{-1}(t)R(u(t)v, u(t)w)u(t)v.\end{aligned} \tag{67.3}$$

Since $u(t)$ is orthogonal for all t, we may now compute $D(v)$ as

$$D(v) = \lambda_y(Z_{w_1}(1), \dots, Z_{w_n}(1))/\lambda_y(w_1, \dots, w_n) = \det Z_\gamma(1), \tag{67.4}$$

where Z_γ is the matrix solution to the differential equation

$$\ddot Z_\gamma(t) = A_\gamma(t)Z_\gamma(t) \text{ with } Z_\gamma(0) = 0 \text{ and } \dot Z_\gamma(0) = I. \tag{67.5}$$

By Taylor's theorem,

$$Z_\gamma(1) = I + \frac{1}{2}\ddot Z_\gamma(0) + \frac{1}{6}Z_\gamma^{(3)}(0) + \int_0^1 Z_\gamma^{(4)}(t)d\mu(t), \tag{67.6}$$

where μ is a positive measure such that $\mu([0,1]) = 1/4!$. Now from the differential equation $\ddot Z_\gamma(0) = 0$,

$$\begin{aligned}Z_\gamma^{(3)}(t) &= \dot A_\gamma(t)Z_\gamma(t) + A_\gamma(t)\dot Z_\gamma(t) \text{ and} \\ Z_\gamma^{(4)}(t) &= \ddot A_\gamma(t)Z_\gamma(t) + 2\dot A_\gamma(t)\dot Z_\gamma(t) + A_\gamma(t)\ddot Z_\gamma(t).\end{aligned}$$

In particular $Z_\gamma^{(3)}(0) = A_\gamma(0) = R(v,\cdot)v$, and

$$Z_\gamma^{(4)}(t) = \ddot A_\gamma(t)Z_\gamma(t) + 2\dot A_\gamma(t)\dot Z_\gamma(t) + A_\gamma(t)A_\gamma(t)Z_w(t).$$

Now $A_\gamma(t) = O(v^2)$,

$$\dot A_\gamma(t) = u^{-1}(t)\left(\nabla_{u(t)v}R\right)(u(t)v, u(t)w)u(t)v = O(v^3),$$

$$\ddot A_\gamma(t) = u^{-1}(t)\left(\nabla^2_{u(t)v\otimes u(t)v}R\right)(u(t)v, u(t)w)u(t)v = O(v^4)$$

and hence $Z_\gamma^{(4)}(t) = O(v^3)$. Using these estimates in Eq. (67.6) shows that

$$Z_\gamma(1) = I + \frac{1}{6}R(v,\cdot)v + O\left(v^3\right). \tag{67.7}$$

Taking the determinant of this equation shows that

$$\begin{aligned}D(v) &= 1 + \frac{1}{6}\mathrm{tr}\,(w \to R(v,w)v) + O(v^3) \\ &= 1 - \frac{1}{6}(\mathrm{Ric}\ v, v) + O(v^3).\end{aligned}$$

Before finishing this section, let us write out Eq. (67.7) in detail.

Lemma 67.1. *Let $v, w \in T_yM$, then*

$$//_1^{-1}(\gamma_v)\exp_*(w_v) = w + \frac{1}{6}R(v,w)v + O\left(v^3\right)w. \tag{67.8}$$

In particular we have for $v, w, u \in T_yM$ that

$$\begin{aligned}
(\exp_*(w_v), \exp_*(w_v)) \\
&= (w + \frac{1}{6}R(v,w)v + O\left(v^3\right)w, u + \frac{1}{6}R(v,u)v + O\left(v^3\right)u) \\
&= (w,u) + \frac{1}{6}(w, R(v,u)v) + \frac{1}{6}(R(v,w)v,u) + O\left(v^3\right)(w,u) \\
&= (w,u) - \frac{1}{3}(R(w,v)v,u) + O\left(v^3\right)(w,u).
\end{aligned} \tag{67.9}$$

67.2 Another Proof of Remark 61.2: The Symmetry of $J(x,y)$.

Recall that λ denotes the Riemannian volume form on M and

$$J(x,y) = \left(\exp_x^* \lambda\right)_v / \lambda_x$$

where $\exp_x(v) = y$. Also recall that $D(v) := \left(\exp_x^* \lambda\right)_v / \lambda_x$ where $x = \pi(v)$ and $\pi : TM \to M$ is the canonical projection map. The precise meaning of this equation is, given any basis $\{w_i\}_{i=1}^n$ for T_xM, then

$$D(v) = \lambda(\exp_*(w_1)_v, \dots, \exp_*(w_n)_v)/\lambda_x(w_1, \dots, w_n) = \det(Z_\gamma(1)),$$

where $w_v = \frac{d}{ds}|_0 (v + sw) \in T_vT_xM$ and Z_γ is defined in Eq. (67.2) above. In particular,

$$\exp_* w_v := \frac{d}{ds}|_0 \exp(v + sw).$$

Notice that

$$J(x,y) = D(\exp_x^{-1}(y)).$$

Let $i : TM \to TM$ denote the involution given by $i(v) = -\dot\gamma(1)$, where $\gamma(t) = \exp(tv)$ is the geodesic determined by v. Alternatively we may describe $i(v) = -//_1(\gamma)v$. Now if $v = \exp_x^{-1}(y)$, i.e. $y = \exp_x(v) = \exp(v)$, then $\exp(i(v)) = x$. That is to say, $i(v) = \exp_y^{-1}(x)$. Hence to show $J(x,y) = J(y,x)$ if and only if $D(v) = D(i(v))$. This is what is proved in A. L. Bess, "Manifolds all of whose Geodesics are Closed," see Lemma 6.12 on p. 156.

Now let us work out $D(i(v))$. Let $\sigma(t) = \gamma(1-t) = \exp(ti(v))$. Since $//_t(\sigma) = //_{1-t}(\gamma)//_1(\gamma)^{-1}$, it follows after a short calculation that $A_\sigma(t) = u(1)A_\gamma(1-t)u(1)^{-1}$. Let $W(t) := u(1)^{-1}Z_\sigma(1-t)u(1)$, then $W(1) = 0$, $\dot W(1) = -I$ and

$$\ddot W(t) = u(1)^{-1}A_\sigma(1-t)Z_\sigma(1-t)u(1) = W(t) = A_\gamma(t)W(t).$$

Notice that

$$D(i(v)) = \det Z_\sigma(1) = \det\left[u(1)^{-1}Z_\sigma(1)u(1)\right] = \det W(0).$$

So to finish the proof, we must show that $\det W(0) = \det Z(1)$. For this observe that $A_\gamma(t)$ is a symmetric operator (by symmetry properties of the curvature tensor) and hence

$$\begin{aligned}
\frac{d}{dt}\left\{\dot Z^*(t)W(t) - Z^*(t)\dot W(t)\right\} &= \ddot Z^*(t)W(t) - Z^*(t)\ddot W(t) \\
&= Z^*(t)A_\gamma(t)W(t) - Z^*(t)A_\gamma(t)W(t) = 0
\end{aligned}$$

and hence

$$\left\{\dot Z^*(t)W(t) - Z^*(t)\dot W(t)\right\}|_0^1 = 0.$$

This implies that

$$W(0) = \dot Z^*(0)W(0) - Z^*(0)\dot W(0) = \dot Z^*(1)W(1) - Z^*(1)\dot W(1) = Z^*(1).$$

Therefore $\det W(0) = \det Z^*(1) = \det Z(1)$ as desired.

67.3 Normal Coordinates

Notation 67.2 *Suppose that $o \in M$ is given and let $x(m) := \exp_o^{-1}(m) \in T_oM$ for m in a neighborhood of o. The chart x is called a geodesic normal coordinate system near o.*

In geodesic coordinates, $t \to tx$ is a geodesic, therefore if Γ is the Christofel symbols in the this coordinate system, we have

$$0 = \frac{\nabla}{dt}\frac{d}{dt}(tx) = \left(\frac{d}{dt} + \Gamma(tx)\langle x\rangle\right)\frac{d}{dt}(tx) = \Gamma(tx)\langle x\rangle x$$

for all x near 0. Since ∇ has zero Torsion we also have that

$$\Gamma(z)\langle x\rangle y = \Gamma(z)\langle y\rangle x$$

for all x, y, z. From the previous two equations it follows that

$$0 = \Gamma(0)\langle x\rangle y \text{ for all } x, y,$$

i.e. that $\Gamma(0) = 0$ and that

$$\partial_x \Gamma(0)\langle x\rangle x = 0.$$

Let $B(x,y,z) := \partial_x\Gamma(0)\langle y\rangle z$, then we have shown that

$$B(x,y,z) = B(x,z,y) \text{ and } B(x,x,x) = 0 \text{ for all } x, y, z. \tag{67.10}$$

Thus

$$0 = \frac{d}{dt}|_0 B(x+ty, x+ty, x+ty) = B(y,x,x) + 2B(x,x,y)$$

and therefore,

$$\begin{aligned} 0 &= \frac{d}{dt}|_0 \{B(y, x+tz, x+tz) + 2B(x+tz, x+tz, y)\} \\ &= 2B(y,x,z) + 2B(z,x,y) + 2B(x,z,y) \\ &= 2\left(B(y,z,x) + B(z,x,y) + B(x,y,z)\right) \end{aligned}$$

wherin the last equality we use Eq. (67.10). Hence we have shown that

$$\partial_x \Gamma(0)\langle y\rangle z + \text{cyclic} = 0. \tag{67.11}$$

So at $x = 0$ the curvature tensor is given by $R = d\Gamma$ and hence

$$\begin{aligned} R(x,y)x &= \partial_x \Gamma(0)\langle y\rangle x - \partial_y \Gamma(0)\langle x\rangle x \\ &= \partial_x \Gamma(0)\langle y\rangle x + \partial_x \Gamma(0)\langle y\rangle x + \partial_x \Gamma(0)\langle x\rangle y \\ &= 3\partial_x \Gamma(0)\langle y\rangle x = 3\partial_x \Gamma(0)\langle x\rangle y \end{aligned}$$

and hence

$$\begin{aligned} \partial_x \Gamma(0)\langle x\rangle y &= \frac{1}{3} R(x,y)x \text{ or} \\ \Gamma(x)\langle x\rangle y &= \frac{1}{3} R(x,y)x + O(x^2)\langle x,y\rangle. \end{aligned}$$

Therefore, if

$$\begin{aligned} \partial_z^2(x,y)|_{z=0} &= \partial_z(\Gamma\langle z\rangle x, y)|_{z=0} + \partial_z(x, \Gamma\langle z\rangle y)|_{z=0} \\ &= \frac{1}{3}(R(z,x)z, y) + \frac{1}{3}(x, R(z,y)z) \\ &= \frac{2}{3}(R(z,x)z, y) \end{aligned}$$

and therefore by Taylor's theorem we learn that

$$\begin{aligned} (x_z, y_z) &= (x,y) + \frac{1}{2!}\frac{2}{3}(R(z,x)z, y) + O(z^3) \\ &= (x,y) - \frac{1}{3}(R(x,z)z, y) + O(z^3) \end{aligned}$$

and hence we have reproved Lemma 67.1.

We now change notation a bit. Let $E \to M$ be a vector bundle with connection ∇.

Notation 67.3 *Let x be a chart on M such that $x(o) = 0$ and let $\gamma_v(t) = x^{-1}(tv)$ for all $v \in \mathbb{R}^{\dim(M)} =: V$. Let u $u : \mathcal{D}(x) \to Gl(E)$ be the local orthonormal frame given by*

$$u(m) = //_1(\gamma_{x(m)}), \text{ i.e. } u(x^{-1}(tv)) = //_t(\gamma_v) : E_o \to E_m \tag{67.12}$$

for all $v \in V$ sufficiently small. Also let $\Gamma = \Gamma^u = u^{-1}\nabla u$ be the associated connection one form.

From Eq. (67.12) it follows that

$$\frac{\nabla}{dt} u(x^{-1}(tv)) = \frac{\nabla}{dt} //_t(\gamma_v) = 0$$

and in particular at $t = 0$ this shows that

$$0 = v^i \nabla_{\partial_i} u|_{\gamma_v(t)} = v^i u\left(\gamma_v(t)\right) \Gamma(\partial_i|_{\gamma_v(t)}) = u\left(\gamma_v(t)\right) \Gamma(v^i \partial_i|_{\gamma_v(t)}).$$

That is to say

$$\Gamma(v^i \partial_i|_{\gamma_v(t)}) = 0 \text{ for all } v \in V.$$

In particular at $t = 0$ we learn that $\Gamma(\partial_i|_o) = 0$ and

$$0 = \frac{d}{dt}|_0 \Gamma(v^i \partial_i|_{\gamma_v(t)}) = v^i v^j \partial_j \Gamma(\partial_i)|_o \text{ for all } v \in V.$$

This shows that $\partial_j \Gamma(\partial_i)|_o = -\partial_i \Gamma(\partial_j)|_o$. Since

$$\begin{aligned} R^E(\partial_i, \partial_j)|_o &= u^{-1} \nabla^2_{\partial_i \wedge \partial_j} u|_o \\ &= u^{-1}(o) \left\{\partial_i \Gamma(\partial_j) - \partial_j \Gamma(\partial_i) + [\Gamma(\partial_i), \Gamma(\partial_j)]\right\}|_o \\ &= -2\partial_j \Gamma(\partial_i)|_o \end{aligned}$$

from which it follows that

$$\partial_j \Gamma(\partial_i)|_o = -\frac{1}{2} R^E(\partial_i, \partial_j)|_o.$$

From Taylor's theorem we find

$$\Gamma(\partial_i)|_m = -\frac{1}{2} x^j(m) R^E(\partial_i, \partial_j)|_o + o(|x(m)|^2) = -\frac{1}{2} R^E(\partial_i, x^j(m)\partial_j)|_o + o(|x(m)|^2)$$

This result is summarized as follows.

Proposition 67.4. *Keeping the notation as above and let $w \in \mathbb{R}^{\dim(M)}$, then*

$$\Gamma(w^i \partial_i) = -\frac{1}{2} R^E(w^i \partial_i|_o, x^j \partial_j|_o) + o(x^2)(w) \tag{67.13}$$

near $o \in M$. In particular if $x(m) = \exp_o^{-1}(m)$ are normal coordinates on M, then

$$w^i\partial_i|_m = \frac{d}{dt}|_0 x^{-1}(x(m)+tw) = \frac{d}{dt}|_0 \exp(x(m)+tw)$$
$$= \exp_*(w_{x(m)}).$$

Therefore, Eq. (67.13) may be written as

$$\Gamma(\exp_*(w_{x(m)})) = -\frac{1}{2}R^E(w,x(m)) + O(x^2(m))(w).$$

Proof. The quick proof of these results is as follows. We work in the local frame u. Write $//_t(\sigma) = u(\sigma(t))P_t(\sigma)$ and recall the formula

$$\frac{d}{ds}|_0 P_t(\Sigma_s) = P_t(\Sigma_s)\int_0^t R^E_{//_\tau(\Sigma_s)}(\dot{\Sigma}_s(\tau),\Sigma_s'(\tau))d\tau - \Gamma(\Sigma_s'(t))P_t(\Sigma_s).$$

Apply this to $\Sigma_s(t) = x^{-1}(t(x(m)+sv))$ and use that fact that in the frame defined by u, $P_t(\Sigma_s) = id$ so that

$$0 = \int_0^t R^E_{//_\tau(\Sigma_0)}(\dot{\Sigma}_0(\tau),\Sigma_0'(\tau))d\tau - \Gamma(\Sigma_0'(t)).$$

Therefore, at $t=1$,

$$\begin{aligned}
\Gamma(v^i\partial_i|_m) &= \int_0^1 R^E_{//_\tau(\Sigma_0)}(\dot{\Sigma}_0(\tau),\Sigma_0'(\tau))d\tau \\
&= -\int_0^1 R^E_{//_\tau(\Sigma_0)}(\tau v^i\partial_i|_{x^{-1}(\tau x(m))}, x^j(m)\partial_j|_{x^{-1}(\tau x(m))})d\tau \\
&= -\int_0^1 R^E_{//_\tau(\Sigma_0)}(v^i\partial_i|_{x^{-1}(\tau x(m))}, x^j(m)\partial_j|_{x^{-1}(\tau x(m))})\tau d\tau \\
&= -\frac{1}{2}R^E(v^i\partial_i|_o, x^j(m)\partial_j|_o) + O(x^2(m)) \\
&= -\frac{1}{2}R^E(v^i\partial_i|_o, x(m)) + O(x^2(m))(v).
\end{aligned}$$

■

68

Miscellaneous

68.1 Jazzed up version of Proposition 68.1

Proposition 68.1. *Let $\alpha > -1$, R, K, P, Q and V be as above. Then the series in Eq. (60.5) and Eq. (60.11) is convergent and is equal to $P_t = e^{tL}$ – the unique solution to Eq. (60.1). Moreover,*

$$\|P-K\|_t \le \kappa(t)\,\|K\|_t\, t^{\alpha+1} = O(t^{1+\alpha}), \tag{68.1}$$

where $\|f\|_t := \max_{0\le s\le t}|f_s|$ and $\kappa(t)$ is an increasing function of t, see Eq. (68.4).

Proof. By making the change of variables $r = s + u(t-s)$ we find that

$$\int_s^t (t-r)^\alpha (r-s)^\beta dr = C(\alpha,\beta)(t-s)^{\alpha+\beta+1} \tag{68.2}$$

where

$$C(\alpha,\beta) := \int_0^1 u^\alpha (1-u)^\beta du = B(\alpha+1,\beta+1),$$

and B is the beta function. From Eq. (1.5.5) of Lebedev, "Special Functions and Their Applications", p.13,

$$C(\alpha,\beta) = B(\alpha+1,\beta+1) = \frac{\Gamma(\alpha+1)\Gamma(\beta+1)}{\Gamma(\alpha+\beta+2)}. \tag{68.3}$$

(See below for a proof of Eq. (68.3).)

By repeated use of Eq. (68.2),

$$\begin{aligned}
|(Q^m R)_t| &= \left|\int_{t\Delta_m} R_{t-s_m} R_{s_m - s_{m-1}} \dots R_{s_2-s_1} R_{s_1} d\mathbf{s}\right| \\
&\le C^m \int_{t\Delta_m} (t-s_m)^\alpha (s_m - s_{m-1})^\alpha \dots (s_2-s_1)^\alpha s_1^a d\mathbf{s} \\
&= C^m C(\alpha,\alpha) \int_{t\Delta_{m-1}} \begin{bmatrix} (t-s_{m-1})^{2\alpha+1}(s_{m-1}-s_{m-2})^\alpha \times \\ \dots (s_2-s_1)^\alpha s_1^a \end{bmatrix} d\mathbf{s} \\
&= C^m C(\alpha,\alpha) C(\alpha, 2\alpha+1) C(\alpha, 3\alpha+2) \dots \\
&\quad \dots C(\alpha,(m-1)\alpha+m-2) \int_0^t (t-s_1)^{m\alpha+m-1} ds_1 \\
&= C^m C(\alpha,\alpha) C(\alpha, 2\alpha+1) C(\alpha, 3\alpha+2) \dots \\
&\quad \dots C(\alpha,(m-1)\alpha+m-2) C(\alpha, m\alpha+m-1) t^{(m+1)\alpha+m}.
\end{aligned}$$

Now from Eq. (68.3) we find that

$$\begin{aligned}
&C(\alpha,\alpha)C(\alpha,2\alpha+1)C(\alpha,3\alpha+2)\dots C(\alpha,m\alpha+m-1) \\
&\quad = \frac{\Gamma(\alpha+1)\Gamma(\alpha+1)}{\Gamma(2\alpha+2)} \frac{\Gamma(\alpha+1)\Gamma(2\alpha+2)}{\Gamma(3\alpha+3)} \times \\
&\quad\quad \frac{\Gamma(\alpha+1)\Gamma(3\alpha+3)}{\Gamma(4\alpha+4)} \dots \frac{\Gamma(\alpha+1)\Gamma(m\alpha+1)}{\Gamma((m+1)\alpha+m+1)} \\
&\quad = \frac{\Gamma(\alpha+1)^m}{\Gamma((m+1)\alpha+m+1)}.
\end{aligned}$$

Therefore,

$$|(Q^m R)_t| \le C^m \frac{\Gamma(\alpha+1)^m}{\Gamma((m+1)\alpha+m+1)} t^\alpha t^{m(\alpha+1)}$$

and thus the series in Eq. (60.11) is absolutely convergent and $|V_t| \le (a+1)\kappa(t)t^\alpha$ where

$$\kappa(t) = (\alpha+1)^{-1}\sum_{m=0}^\infty C^m \frac{\Gamma(\alpha+1)^m}{\Gamma((m+1)\alpha+m+1)} t^{m(\alpha+1)} \tag{68.4}$$

which is seen to be finite by Stirlings formula,

$$\Gamma(m(\alpha+1)+1) \backsim (2\pi)^{1/2} e^{-(m(\alpha+1)+1)} (m(\alpha+1)+1)^{(m(\alpha+1)+1/2)},$$

see Eq. (1.4.12) of Lebedev. ■

Using the bound on V and the uniform boundedness of K_t,

$$\int_0^t |K_s V_{t-s}|\, ds \le \kappa(t)\,\|K\|_t (\alpha+1) \int_0^t (t-s)^\alpha ds = \kappa(t)\,\|K\|_t\, t^{\alpha+1} \tag{68.5}$$

and hence P_t defined in Eq. (60.10) is well defined and is continuous in t. Moreover, (68.5) implies Eq. (68.1) once we shows that $P_t = e^{tL}$. This is checked as follows,

$$\frac{d}{dt}\int_0^t K_{t-s}V_s ds = V_t + \int_0^t \dot{K}_{t-s}V_s ds = V_t + \int_0^t (LK_{t-s} - R_{t-s}) V_s ds$$
$$= V_t + L\int_0^t K_{t-s}V_s ds - (QV)_t = L\int_0^t K_{t-s}V_s ds + R_t.$$

Thus we have,

$$\frac{d}{dt}P_t = \dot{K}_t + L\int_0^t K_{t-s}V_s ds + R_t$$
$$= LK_t + L\int_0^t K_{t-s}V_s ds = LP_t.$$

68.1.1 Proof of Eq. (68.3)

Let us recall that $\Gamma(x) := \int_0^\infty t^x e^{-t}dt/t$ and hence let $z = x + y$ and then $x = uz$ we derive,

$$\begin{aligned}\Gamma(\alpha+1)\Gamma(\beta+1) &= \int_{[0,\infty)^2} x^\alpha y^\beta e^{-(x+y)}dxdy \\ &= \int_0^\infty dx \int_x^\infty dz\, x^\alpha (z-x)^\beta e^{-z} \\ &= \int dxdz 1_{0<x<z<\infty}\, x^\alpha (z-x)^\beta e^{-z} \\ &= \int_0^\infty dz \int_0^1 du\, u^\alpha (1-u)^\beta z^{\alpha+\beta+1} e^{-z} \\ &= C(\alpha,\beta)\Gamma(\alpha+\beta+2),\end{aligned}$$

i.e.

$$C(\alpha,\beta) = \frac{\Gamma(\alpha+1)\Gamma(\beta+1)}{\Gamma(\alpha+\beta+2)}.$$

68.1.2 Old proof of Proposition 60.1

Proof. Taking norms of Eq. (60.8) shows that

$$\|Q^m K\|_t = \|K\|_t C^m \int_{t\Delta_m} (s_2 - s_1)^\alpha (s_3 - s_2)^\alpha \dots (s_m - s_{m-1})^\alpha (t - s_m)^\alpha d\mathbf{s},$$

where $d\mathbf{s} = ds_1 ds_2 \dots ds_m$. To evaluate this last integral, we will make use Eq. (68.2) to find of Repeated use of Eq. (68.2) gives,

$$\begin{aligned}&\int_{t\Delta_m} (s_2 - s_1)^\alpha (s_3 - s_2)^\alpha \dots (s_m - s_{m-1})^\alpha (t - s_m)^\alpha d\mathbf{s} \\ &= C(\alpha,\alpha)\int_{t\Delta_{m-1}} \left[\begin{array}{c}(s_2 - s_1)^\alpha (s_3 - s_2)^\alpha \times \\ \dots (s_{m-1} - s_{m-2})^\alpha (t - s_{m-1})^{2\alpha+1}\end{array}\right] d\mathbf{s} \\ &= C(\alpha,\alpha)C(\alpha,2\alpha+1)\int_{t\Delta_{m-2}} \left[\begin{array}{c}(s_2 - s_1)^\alpha (s_3 - s_2)^\alpha \times \\ \dots (s_{m-2} - s_{m-3})^\alpha (t - s_{m-2})^{3\alpha+2}\end{array}\right] d\mathbf{s} \\ &= C(\alpha,\alpha)C(\alpha,2\alpha+1)C(\alpha,3\alpha+2)\times \\ &\quad \dots C(\alpha,(m-1)\alpha+m-2)\int_0^t (t-s_1)^{m\alpha+m-1} ds_1 \\ &= C(\alpha,\alpha)C(\alpha,2\alpha+1)C(\alpha,3\alpha+2)\times \\ &\quad \dots C(\alpha,(m-1)\alpha+m-2)\frac{t^{m\alpha+m}}{m\alpha+m}.\end{aligned}$$

Now from Eq. (68.3) we find that

$$\begin{aligned}&C(\alpha,\alpha)C(\alpha,2\alpha+1)C(\alpha,3\alpha+2)\dots C(\alpha,(m-1)\alpha+m-2) \\ &= \frac{\Gamma(\alpha+1)^2}{\Gamma(2\alpha+2)}\frac{\Gamma(\alpha+1)\Gamma(2\alpha+2)}{\Gamma(3\alpha+3)}\times \\ &\quad \dots \frac{\Gamma(\alpha+1)\Gamma((m-1)\alpha+m-1)}{\Gamma(m\alpha+m)} \\ &= \frac{\Gamma(\alpha+1)^m}{\Gamma(m\alpha+m)}.\end{aligned}$$

Combining these results gives the estimate,

$$\begin{aligned}\|Q^m K\|_t &\le \|K\|_t \frac{(C\Gamma(\alpha+1))^m}{\Gamma(m\alpha+m)}\frac{t^{m\alpha+m}}{m\alpha+m} \\ &= \|K\|_t \frac{\left(C\Gamma(\alpha+1)t^{\alpha+1}\right)^m}{\Gamma(m(\alpha+1)+1)} \\ &\backsim \|K\|_t \left(C\Gamma(\alpha+1)t^{\alpha+1}\right)^m \frac{e^{m(\alpha+1)+1}}{(2\pi)^{1/2}(m(\alpha+1)+1)^{(m(\alpha+1)+1/2)}} \\ &= \|K\|_t \frac{\left(C\Gamma(\alpha+1)(et)^{\alpha+1}\right)^m}{(2\pi)^{1/2}e\,(m(\alpha+1)+1)^{(m(\alpha+1)+1/2)}},\end{aligned}$$

where the second to last expression is a result of Stirlings formula, Eq. (1.4.12) of Lebdev.

From this estimate we learn that $\sum_{m=0}^\infty Q^m K$ is uniformly convergent on compact subsets of $[0,\infty)$ and that

$$\left\|\sum_{m=0}^{\infty}Q^mK-K\right\|_t\le\sum_{m=1}^{\infty}\|Q^mK\|_t$$
$$\le\|K\|_t\sum_{m=1}^{\infty}\frac{\left(C\Gamma(\alpha+1)t^{\alpha+1}\right)^m}{\Gamma(m(\alpha+1)+1)}=O(t^{1+\alpha}).$$

So it only remains to prove that $P:=\sum_{m=0}^{\infty}Q^mK$ solves (60.1).

Now by the chain rule and the fundamental theorem of calculus, $\frac{d}{dt}(Qf)_t=\int_0^t\dot f_{t-s}R_sds+f_0R_t$ or equivalently

$$\frac{d}{dt}(Qf)=Q\dot f+f_0R.$$

Applying this formula inductively using the fact that $(Q^mK)_0=0$ if $m\ge1$ implies that

$$\begin{aligned}\frac{d}{dt}(Q^mK)&=Q^{m-1}\frac{d}{dt}(QK)=Q^m\dot K+Q^{m-1}(K_0R)\\&=Q^m(LK-R)+Q^{m-1}R\\&=LQ^mK-Q^mR+Q^{m-1}R,\end{aligned}$$

wherein the last equality we have used the fact that L commutes with Q. Setting $P_t^N:=\sum_{m=0}^{N}Q^mK$, we find using the previous equation that

$$\frac{d}{dt}P_t^N=LP_t^N-\left(Q^NR\right)_t$$

or equivalently that

$$P_t^N=I+\int_0^tLP_s^Nds-\int_0^t\left(Q^NR\right)_sds. \tag{68.6}$$

Since,

$$\left\|Q^NR\right\|_t\le\|R\|_t\frac{\left(C\Gamma(\alpha+1)t^{\alpha+1}\right)^N}{\Gamma(N(\alpha+1)+1)}\to0\text{ as }N\to\infty,$$

we may pass to the limit, $N\to\infty$, in Eq. (68.6) to conclude that

$$P_t=I+\int_0^tLP_s\,ds.$$

This completes the proof.

For later purposes, let us rework the above derivative aspects of the proof. Let

$$\begin{aligned}R_m(s)&:=\int_{0\le s_1\le s_2\le\dots\le s_{m-1}\le s}R_{s-s_{m-1}}R_{s_{m-1}-s_{m-2}}\dots R_{s_2-s_1}R_{s_1}d\mathbf{s}\\&=\left(Q^{m-1}R\right)_s,\end{aligned}$$

then by Eq. (60.6)

$$(Q^mK)_t=\int_0^tK_{t-s}R_m(s)ds.$$

Hence

$$\begin{aligned}\frac{d}{dt}(Q^mK)_t&=R_m(t)+\int_0^t\frac{d}{dt}K_{t-s}R_m(s)ds\\&=R_m(t)+\int_0^t(LK_{t-s}-R_{t-s})R_m(s)ds\\&=R_m(t)+L(Q^mK)_t-R_{m+1}(t)\\&=L(Q^mK)_t+\left(Q^{m-1}R\right)_t-(Q^mR)_t.\end{aligned}$$

as before. ■

Taking norms of this equation implies that

$$\begin{aligned}|(Q^mf)_t|&\le C^m\int_{t\Delta_m}|f_{s_1}|(s_2-s_1)^\alpha(s_3-s_2)^\alpha\dots(s_m-s_{m-1})^\alpha(t-s_m)^\alpha\,d\mathbf{s}\\&\le C^mt^{\alpha m}\int_{t\Delta_m}|f_{s_1}|\,d\mathbf{s}=C^mt^{\alpha m}\int_0^t|f_{s_1}|\frac{(t-s_1)^{m-1}}{(m-1)!}ds_1\\&\le\frac{\left(Ct^{1+\alpha}\right)^m}{m!}\int_0^t|f_s|\,d\mu_m(s)\le\frac{\left(Ct^{1+\alpha}\right)^m}{m!}\max_{0\le s\le t}|f_s|,\end{aligned} \tag{68.7}$$

where $d\mu_m(s):=m(t-s)^{m-1}t^{-m}ds$, a probability measure on $[0,t]$. This shows that $\|Q^mf\|_t\le\left(Ct^{1+\alpha}\right)^m\|f\|_t/m!$. From this estimate we learn that $\sum_{m=0}^{\infty}Q^mK$ is uniformly convergent on compact subsets of $[0,\infty)$ and that

$$\left\|\sum_{m=0}^{\infty}Q^mK-K\right\|_t\le\sum_{m=0}^{\infty}\|Q^mK\|_t\le e^{Ct^{1+\alpha}}-1=O(t^{1+\alpha}).$$

$$\begin{aligned}\frac{d}{dt}\int_0^tK_{t-s}V_sds&=\lim_{\delta\to0}\frac{1}{\delta}\left(\int_0^{t+\delta}K_{t+\delta-s}V_sds-\int_0^tK_{t-s}V_sds\right)\\&=\lim_{\delta\to0}\frac{1}{\delta}\left(\int_t^{t+\delta}K_{t+\delta-s}V_sds+\int_0^t(K_{t+\delta-s}-K_{t-s})V_sds\right)\\&=\lim_{\delta\to0}\left(\frac{1}{\delta}\int_t^{t+\delta}K_{t-s}V_sds+\frac{1}{\delta}\int_t^{t+\delta}(K_{t+\delta-s}-K_{t-s})V_sds\right)\\&\quad+\frac{d}{d\delta}\int_0^tK_{t+\delta-s}V_sds\\&=V_t+\int_0^t\dot K_{t-s}V_sds=V_t+\int_0^t(LK_{t-s}-R_{t-s})V_sds\\&=V_t+L\int_0^tK_{t-s}V_sds-(QV)_t=L\int_0^tK_{t-s}V_sds+R_t\end{aligned}$$

68.1.3 Old Stuff related to Theorem 61.7

Proof. Let

$$F(t,x,y) = \int_0^t \int_M k(t-s,x,z)v(s,z,y)d\lambda(z)ds$$

so that $p(t,x,y) = k(t,x,y) + F(t,x,y)$. For $\delta > 0$ set

$$F_\delta(t,x,y) = \int_0^t \int_M k(t+\delta-s,x,z)v(s,z,y)d\lambda(z)ds,$$

then

$$\begin{aligned}
\partial_t F_\delta(t,x,y) &= \int_M k(\delta,x,z)v(t,z,y)d\lambda(z) \\
&\quad + \int_0^t \int_M \partial_t k(t+\delta-s,x,z)v(s,z,y)d\lambda(z)ds \\
&= \int_M k(\delta,x,z)v(t,z,y)d\lambda(z) \\
&\quad + \int_0^t \int_M L_x k(t+\delta-s,x,z)v(s,z,y)d\lambda(z)ds \\
&\quad - \int_0^t \int_M r(t+\delta-s,x,z)v(s,z,y)d\lambda(z)ds \\
&= \int_M k(\delta,x,z)v(t,z,y)d\lambda(z) + L_x F_\delta(t,x,y) \\
&\quad - \int_0^t \int_M r(t+\delta-s,x,z)v(s,z,y)d\lambda(z)ds.
\end{aligned}$$

We may let $\delta \to 0$ in this last expression using the fact that K_t is uniformly bounded on Γ_l to find that

$$\lim_{\delta \downarrow 0} F_\delta(t,x,y) = F(t,x,y)$$

and

$$\begin{aligned}
\lim_{\delta \downarrow 0} \partial_t F_\delta(t,x,y) &= v(t,x,y) + L_x F(t,x,y) \\
&\quad - \int_0^t \int_M r(t-s,x,z)v(s,z,y)d\lambda(z)ds
\end{aligned}$$

with the limits being uniform in t. Also by equation (62.1) and (62.2),

$$v(t,x,y) - \int_0^t \int_M r(t-s,x,z)v(s,z,y)d\lambda(z)ds = r(t,x,y),$$

and therefore $\partial_t F$ exists and

$$\partial_t F(t,x,y) = L_x F(t,x,y) + r(t,x,y).$$

Hence

$$\begin{aligned}
\partial_t p(t,x,y) &= \partial_t k(t,x,y) + \partial_t F(t,x,y) \\
&= L_x k(t,x,y) - r(t,x,y) + L_x F(t,x,y) + r(t,x,y) \\
&= L_x \left(k(t,x,y) + F(t,x,y)\right) = L_x p(t,x,y).
\end{aligned}$$

For $s < t$, let

$$b(t,s,x,y) := \int_M k(t-s,x,z)v(s,z,y)d\lambda(z).$$

It is clear the b is smooth in t and s and is C^l in (x,y), moreover

$$\lim_{s \uparrow 0} \|b(t,s,x,y) - v(t,x,y)\|_{C^l(x,y)} = 0.$$

Hence we set $b(t,t,x,y) := v(t,x,y)$ so that $b(t,s,\cdot,\cdot)$ is continuous for $s \in [0,t]$ in the space of C^l sections Γ_l. Similarly, for $s < t$,

$$\begin{aligned}
\partial_t b(t,s,x,y) &= \int_M \dot{k}(t-s,x,z)v(s,z,y)d\lambda(z) \\
&= \int_M \left(L_x k(t-s,x,z) - r(t-s,x,z)\right)v(s,z,y)d\lambda(z) \\
&= L_x b(t,s,x,y) - \int_M r(t-s,x,z)v(s,z,y)d\lambda(z).
\end{aligned}$$

From this last expression and our previous comments, $\lim_{s\uparrow 0} L_x b(t,s,x,y) = L_x v(t,x,y)$ in Γ_{l-2} and hence

$$\lim_{s \uparrow 0} \partial_t b(t,s,x,y) = L_x v(t,x,y) \text{ in } C^{l-2}.$$

More precisely, we will construct $p(t,x,y)$ as

$$\begin{aligned}
p(t,x,y) = \sum_{m=0}^{\infty} \int_{t\Delta_m} \int_{M^m} & k(t-s_m,x,y_m)r(s_m-s_{m-1},y_m,y_{m-1}) \\
& r(s_{m-1}-s_{m-2},y_{m-1},y_{m-2}) \dots r(s_1,y_1,y)d\mathbf{s}d\mathbf{y},
\end{aligned} \tag{68.8}$$

Consider,

$$\partial_t p(t,x,y) = \partial_t k(t,x,y) + \int_0^t \int_M k(t-s,x,z)v(s,y,z)dsd\lambda(y).$$

is a bounded operator and its derivatives in s up to order k is a convergent sum in converge. ■

69

Remarks on Covariant Derivatives on Vector Bundles

Let $\pi : E \to M$ be a vector bundle with fiber W. A local frame u on E is a local section of the bundle $Aut(V,E) \to M$, i.e. for $m \in \mathcal{D}(u)$ (the domain of u) $u(m) : W \to E_m = \pi^{-1}(m)$ is a linear isomorphism of vector spaces. Notice that any local section S of E may be written as $S(m) = u(m)s(m)$, where $s \in C^\infty(M,W)$. Suppose that ∇ is a covariant derivative on E, define for $v \in T_mM$, a linear transformation $\nabla_v u : W \to E_m$ by

$$(\nabla_v u)\, w := \nabla_v(u(\cdot)w) \text{ for each } w \in W.$$

With this notation and the basic properties of ∇, given $s \in C^\infty(M,W)$, we have that

$$\nabla_v(us) = (\nabla_v u)\, s(m) + u(m)\partial_v s,$$

where $\partial_v s := \frac{d}{dt}|_0 s(\sigma(t))$ provided that $\dot\sigma(0) = v$. In particular this shows that

$$u(m)^{-1}\nabla_v(us) = \partial_v s + A(v)s(m),$$

where $A(v) = A^u(v) := u(m)^{-1}(\nabla_v u)$. So the local representation of ∇ is $\nabla = d + A$, where A is a one form with values in $End(W)$.

Given a path $S(t) \in E$, let $\sigma(t) = \pi(S(t))$ and $s(t) = u(\sigma(t))^{-1}S(t)$. Then define

$$\nabla S(t)/dt := u(\sigma(t))\left(\dot s(t) + A(\dot\sigma(t))s(t)\right), \tag{69.1}$$

i.e. the local version of $\frac{\nabla}{dt} = \frac{d}{dt} + A(\dot\sigma(t))$. Notice that if $S = us$ is a local section of E, then

$$\begin{aligned}\nabla S(\sigma(t)))/dt &= u(\sigma(t))\left(\frac{d}{dt}s(\sigma(t)) + A(\dot\sigma(t))s(\sigma(t))\right)\\ &= u(\sigma(t))\left(\partial_{\dot\sigma(t)}s + A(\dot\sigma(t))s(\sigma(t))\right)\\ &= \nabla_{\dot\sigma(t)}S.\end{aligned}$$

This explains why ∇/dt is independent of the local frame u used in Eq. (69.1), a property which follows by direct computation as well.

We say that a path $S(t) \in E$ is parallel provided that $\nabla S(t)/dt = 0$ for all t. Given a curve $\sigma(t)$ in M and a point $S_0 \in E_{\sigma(0)}$, there is a unique path $S(t) \in E$ such that $\pi(S(t)) = \sigma(t)$ and $\nabla S(t)/dt = 0$. This path is constructed by solving (locally) the linear equation

$$\dot s(t) + A(\dot\sigma(t))s(t) = 0 \text{ with } s(0) = u(\sigma(0))^{-1}S_0$$

and then setting $S(t) = u(\sigma(t))s(t)$. It is easy to check that the map $S_0 \in E_{\sigma(0)} \to S(t) \in E_{\sigma(t)}$ is linear. In fact $S(t) = //_t(\sigma)S_0$, where $//_t(\sigma) = u(\sigma(t))g(t)u(\sigma(0))^{-1}$ and $g(t) \in End(W)$ is the unique solution to the linear differential equation,

$$\dot g(t) + A(\dot\sigma(t))g(t) = 0 \text{ with } g(0) = id \in End(W).$$

We will call $//_t(\sigma)$ parallel translation along σ. It is uniquely characterized as the solution to the differential equation

$$\nabla //_t(\sigma)/dt = 0 \text{ with } //_0(\sigma) = id \in End(E_{\sigma(0)}),$$

(If $U(t) \in End(E_{\sigma(0)}, E_{\sigma(t)})$ for each t, then $\nabla U(t)/dt$ is by definition the linear transformation from $E_{\sigma(0)}$ to $E_{\sigma(t)}$ determined by $(\nabla U(t)/dt)\,\xi := \nabla\left(U(t)\xi\right)/dt$ for all $\xi \in E_0$.) We have the following properties of parallel translation which follow from the uniqueness theorem for ordinary differential equations and the chain rule for covariant derivatives. Namely if $S(t)$ is a smooth path in E and $t = \tau(s)$, then

$$\nabla S(\tau(s))/ds = \tau'(s)\, \nabla S(t)/dt|_{t=\tau(s)}.$$

This property is easily verified from Eq. (69.1).

Proposition 69.1. *Let $\sigma(t) \in M$ be a smooth curve for $t \in [0,T]$ and let $\tau : [0,S] \to [0,T]$ be a smooth function such that $\tau(0) = 0$. Then $//_s(\sigma\circ\tau) = //_{\tau(s)}(\sigma)$, i.e. parallel translation does not depend on how the underlying curve is parametrized. Secondly, let $\tilde\sigma(t) := \sigma(T-t)$, then $//_t(\tilde\sigma) = //_{T-t}(\sigma)//_T(\sigma)^{-1}$. In particular $//_T(\sigma)^{-1} = //_T(\tilde\sigma)$.*

Proof. We have that

$$\nabla //_s(\sigma\circ\tau)/ds = 0 \text{ with } //_0(\sigma\circ\tau) = id \in End(E_{\sigma(0)})$$

and

$$\begin{aligned}&\nabla //_{\tau(s)}(\sigma)/ds = \nabla //_t(\sigma)/dt|_{t=\tau(s)}\tau'(s) = 0 \text{ with}\\ &//_{\tau(s)}(\sigma)|_{s=0} = id \in End(E_{\sigma(0)})\end{aligned}$$

and hence by uniqueness of solutions to O.D.E.'s we must have that $//_s(\sigma\circ\tau) = //_{\tau(s)}(\sigma)$. Similarly,

$$\nabla //_t(\tilde{\sigma})/dt = 0 \text{ with } //_0(\tilde{\sigma}) = id \in E_{\sigma(T)}$$

and

$$\nabla //_{T-t}(\sigma)//_T(\sigma)^{-1}/dt = -\nabla //_s(\sigma)//_T(\sigma)^{-1}/ds|_{s=T-t} = 0 \text{ with}$$
$$//_{T-t}(\sigma)//_T(\sigma)^{-1}|_{t=0} = id \in E_{\sigma(T)}.$$

Hence again by uniqueness of solutions to O.D.E.'s we must have that $//_t(\tilde{\sigma}) = //_{T-t}(\sigma)//_T(\sigma)^{-1}$. ■

70

Spin Bundle Stuff

Let M^n be a Riemannian manifold, $V = \mathbb{R}^n$, $Cl(V)$ be the Clifford algebra over V such that $v^2 = -(v,v)1$. Let $\text{Spin}(n) \subset Cl(V)$ be the spin group, ρ :Spin$(n) \to SO(n)$ be the spin representation and W be a left $Cl(V)$ module. The following compatibility condition is need below in the construction of Spinor bundles S over M such that $Cl(TM)$ acts on S.

Assumption 7 *For $h \in$Spin(n), $v \in V$ and $w \in W$, $h(vw) = (\rho(h)v)(hw)$.*

Now for the construction of spinor bundles. Let $\{U_\alpha\}_{\alpha \in A}$ be an open cover of M such that there exists $u_a : U_\alpha \to Hom(\mathbb{R}^n, TU_\alpha)$ which are isometries. For $m \in U_\alpha \cap U_\beta$ let $g_{\alpha\beta}(m) := u_\alpha(m)^{-1}u_\beta(m) \in O(n)$. Notice that the for $m \in U_\alpha \cap U_\beta \cap U_\delta$,

$$g_{\alpha\beta}(m)g_{\beta\delta}(m) = u_\alpha(m)^{-1}u_\beta(m)u_\beta(m)^{-1}u_\delta(m) = u_\alpha(m)^{-1}u_\delta(m) = g_{\alpha\delta}(m).$$

We now assume that M is orientable which means that we may choose u_α such that $g_{\alpha\beta}(m) \in SO(n)$. Now if M is spin as well, we may choose $\tilde{g}_{\alpha\beta}(m) \in$Spin$(n)$ such that

1. $g_{\alpha\beta}(m) = \rho(\tilde{g}_{\alpha\beta}(m))$ for all $m \in U_\alpha \cap U_\beta$ and all α and β.
2. $\tilde{g}_{\alpha\beta}(m)\tilde{g}_{\beta\delta}(m) = \tilde{g}_{\alpha\delta}(m)$.

Given this data, it is now possible to build a spin bundle over M as follows. For $m \in M$, let

$$S_m := \{(m, \alpha, w) : w \in W \text{ and } \alpha \in A \text{ s.t. } m \in U_\alpha\} / \backsim$$

where $(m, \alpha, w) \backsim (m, \alpha', w')$ if and only if $w' = \tilde{g}_{\alpha'\alpha}(m)w$. Let $S := \cup_{m \in M} S_m$ and $\pi : S \to M$ be the projection map which takes S_m to m for all $m \in M$. Given $\alpha \in A$, let $\tilde{u}_\alpha : U_\alpha \to S_{U_\alpha} := \pi^{-1}(U_\alpha) = \cup_{m \in U_\alpha} S_m$ be given by $\tilde{u}_\alpha(m)w$ is the equivalence class containing (m, α, w). Notice that $\tilde{u}_\alpha(m) : W \to S_m$ is a bijective map and that $\tilde{u}_\alpha(m)^{-1}\tilde{u}_\beta(m)w = \tilde{g}_{\alpha\beta}(m)w$. One may now easily check that we may make S_m in a well defined way into a linear

space by defining $\tilde{u}_\alpha(m)w + c\tilde{u}_\alpha(m)w' := \tilde{u}_\alpha(m)(w + cw')$, i.e. by requiring each $\tilde{u}_\alpha(m) : W \to S_m$ to be linear.

Let us now show that we can make S_m into a $Cl(T_mM)$ module. For $\eta \in T_mM$ and $\xi \in S_m$ choose $\alpha \in A$ such that $m \in U_\alpha$ and choose $v \in \mathbb{R}^n$ and $w \in W$ such that $\eta = u_\alpha(m)v$ and $\xi = \tilde{u}_\alpha(m)w$. We then define $\eta\xi := \tilde{u}_\alpha(m)(vw)$. To see this is well defined choose $\alpha' \in A$ such that $m \in U_{\alpha'}$ and choose $v' \in \mathbb{R}^n$ and $w' \in W$ such that $\eta' = u_{\alpha'}(m)v'$ and $\xi' = \tilde{u}_{\alpha'}(m)w'$. Then $w = \tilde{g}_{\alpha\alpha'}(m)w'$ and $v = g_{\alpha\alpha'}(m)v' = \rho(\tilde{g}_{\alpha\alpha'}(m))v'$, and hence

$$vw = (g_{\alpha\alpha'}(m)v')(\tilde{g}_{\alpha\alpha'}(m)w') = (\rho(\tilde{g}_{\alpha\alpha'}(m))v')(\tilde{g}_{\alpha\alpha'}(m)w') = \tilde{g}_{\alpha\alpha'}(v'w').$$

From this it follows that $\tilde{u}_\alpha(m)(vw) = \tilde{u}_{\alpha'}(m)(v'w')$ so that $\eta\xi$ is well defined independent of the choice of $\alpha \in A$. Since $\xi \in T_mM \xrightarrow{\phi} L_\xi \in End(S_m)$ (L_ξ denotes left multiplication by ξ) satisfies $\phi(\xi)^2 = -(\xi,\xi)_m I$, it follows that the action of T_mM on S_m extends uniquely to an action of $Cl(T_mM)$ on S_m.

Hence if M is a spin manifold, we have produced a vector bundle $S \to M$ such that each fiber of S_m of S is a $Cl(T_mM)$ Clifford module.

71

The Case where $M = \mathbb{R}^n$

71.1 Formula involving p

Let $L := \frac{1}{2}\Delta + B + c$, $L_0 := \frac{1}{2}\Delta + B$ and $\beta(x) := b(x) \cdot x$ where and $B = b \cdot \nabla = \sum_{i=1}^n b_i \partial_i$ and with $b_i(x)$ and $c(x)$ in $\mathbb{R}^{N\times N}$. Let g be an $\mathbb{R}^{N\times N}$ – valued function of (t,x) with $t > 0$ and $x \in \mathbb{R}^n$ and set

$$u(t,x) = p(t,x)g(t,x).$$

Then

$$\begin{aligned}(\partial_t - L)\, u &= (\partial_t - L)\, pg \\ &= (\partial_t - L_0)\, p \cdot g + p\,(\partial_t - L)\, g - \nabla p \cdot \nabla g \\ &= -Bp \cdot g + p\,(\partial_t - L)\, g - \nabla p \cdot \nabla g \\ &= p\,\{-B \ln p + \partial_t - L - \nabla \ln p \cdot \nabla\}\, g.\end{aligned}$$

Now

$$\nabla \ln p = \nabla\left(-x^2/2t\right) = -\frac{x}{t}$$

so the above equation may be written as:

$$\begin{aligned}(\partial_t - L)\, u &= p\left(\frac{1}{t} b \cdot x + \partial_t - L + \frac{1}{t} x \cdot \nabla\right) g \\ &= p\left(\partial_t - L + \frac{1}{t} S\right) g\end{aligned}$$

where

$$\begin{aligned}S &= \partial_x + b \cdot x \\ &= \partial_x + \beta\end{aligned}$$

and

$$\beta(x) = b(x) \cdot x.$$

71.2 Asymptotics of a perturbed Heat Eq. on $\mathbb{R}^n$

Let $L := \frac{1}{2}\Delta + B + c$, $L_0 := \frac{1}{2}\Delta + B$ and $\beta(x) := b(x) \cdot x$ where and $B = b \cdot \nabla = \sum_{i=1}^n b_i \partial_i$ and with $b_i(x)$ and $c(x)$ in $\mathbb{R}^{N\times N}$. As above let

$$p(t,x) = (2\pi t)^{-n/2} e^{-x^2/2t}$$

be the heat kernel with pole at 0 for $\mathbb{R}^n$.

Lemma 71.1. *Let $u_0 \in \mathbb{R}^{N\times N}$ be given, then there is a unique solution to the O.D.E*

$$\dot{U}(t,x) = -\frac{1}{t}\beta(tx)U(t,x) \text{ with } U(0,x) = u_0 \in \mathbb{R}^{N\times N}. \tag{71.1}$$

Moreover, $U(t,x)$ is smooth in (t,x) and

$$U(t,sx) = U(ts,x) \tag{71.2}$$

for all $s,t \in \mathbb{R}$ and $x \in \mathbb{R}^n$.

Proof. Since $\beta(0) = 0$,

$$\frac{1}{t}\beta(tx) = \frac{1}{t}\int_0^t \partial_x \beta(\tau x) d\tau = \int_0^1 \partial_x \beta(utx) du$$

and hence the matrix function $(t,x) \to \frac{1}{t}\beta(tx)$ is smooth even for t near 0. By basic O.D.E. theory this shows that U exists and is smooth. Since

$$\begin{aligned}\frac{d}{dt}U(ts,x) &= s\dot{U}(ts,x) = -s\frac{1}{st}\beta(tsx)U(ts,x) \\ &= -\frac{1}{t}\beta(tsx)U(ts,x),\end{aligned}$$

it follows that $U(ts,x)$ satisfies the same O.D.E. as $U(t,sx)$. Hence by uniqueness of solutions to O.D.E.'s we find that Eq. (71.2) holds. ■

Remark 71.2. If $[\beta(x), \beta(y)] = 0$ for all x and y, then the solution to Eq. (71.1) is given by

$$U(1,x) = \exp\left(-\int_0^1 \frac{1}{t}\beta(tx)\, dt.\right). \tag{71.3}$$

The rest of this section is devoted to the proof of the following Theorem.

Theorem 71.3. *Let U be defined as in Lemma 71.1 and defined*

$$u_0(x) = U(1,x), \tag{71.4}$$

and u_k inductively by

$$u_{k+1}(x) = u_0(x) \int_0^1 \tau^k u_0(\tau x)^{-1} L u_k(\tau x) d\tau. \tag{71.5}$$

for all $k = 0, 1, 2 \ldots.$ *If*

$$\Sigma_q(t,x) := p(t,x) \sum_{k=0}^{q} t^k u_k(x), \tag{71.6}$$

then

$$(\partial_t - L)\, \Sigma_q(t,x) = -t^q p(t,x) L u_q(t,x). \tag{71.7}$$

The proof of this theorem could be given by direct computation. However, we will take a longer route however and derive the formulas in the Theorem.

Let g be an $\mathbb{R}^{N\times N}$ – valued function of (t,x) with $t > 0$ and $x \in \mathbb{R}^n$ and set

$$u(t,x) = p(t,x) g(t,x).$$

Then

$$\begin{aligned}
(\partial_t - L)\, u &= (\partial_t - L)\, pg \\
&= (\partial_t - L_0)\, p \cdot g + p\,(\partial_t - L)\, g - \nabla p \cdot \nabla g \\
&= \left(p \frac{r}{2t} \partial_r \ln J\ - Bp\right) \cdot g + p\,(\partial_t - L)\, g - \nabla p \cdot \nabla g \\
&= p \left\{ \frac{r}{2t} \partial_r \ln J - B \ln p + \partial_t - L - \nabla \ln p \cdot \nabla \right\} g.
\end{aligned}$$

Now

$$\nabla \ln p = \nabla \left(-x^2/2t\right) = -\frac{x}{t}$$

so the above equation may be written as:

$$\begin{aligned}
(\partial_t - L)\, u &= p \left(b \cdot \frac{x}{t} + \partial_t - L + \frac{x}{t} \cdot \nabla \right) g \\
&= p \left(\partial_t - L + \frac{1}{t} S \right) g
\end{aligned} \tag{71.8}$$

where

$$S := x \cdot \nabla + b(x) \cdot x = \partial_x + \beta(x), \tag{71.9}$$

and $\beta(x) := b(x) \cdot x$ as above.

Now let

$$g(t,x) = g_q(t,x) = \sum_{k=0}^{q} t^k u_k(x) \tag{71.10}$$

and consider

$$\begin{aligned}
\left(\partial_t + \frac{1}{t} S - L\right) g &= \sum_{k=0}^{q} \left\{ t^{k-1} \left(k u_k + S u_k\right) - t^k L u_k \right\} \\
&= \frac{1}{t} S u_0 + \sum_{k=0}^{q-1} t^k \left((k+1) u_{k+1} + S u_{k+1} - L u_k\right) - t^q L u_q.
\end{aligned}$$

Thus if we choose u_0 such that

$$S u_0(x) = (x \cdot \nabla. + \beta(x))\, u_0(x) = 0. \tag{71.11}$$

and u_k such that

$$(k+1) u_{k+1} + S u_{k+1} - L u_k = 0 \tag{71.12}$$

then $\left(\partial_t + \frac{1}{t} S - L\right) g = -t^q L u_q$ or equivalently by Eq. (71.8),

$$(\partial_t - L)\, \Sigma_q = (\partial_t - L)\,(pg) = -t^q L u_q$$

which then proves Theorem 71.3 show that u_k defined in the theorem solve Equations (71.11) and (71.12).

Suppose that u_0 is a solution to Eq. (71.11). If $U(t,x) := u_0(tx)$, then

$$\begin{aligned}
\dot{U}(t,x) &= x \cdot \nabla u_0(tx) = \frac{1}{t} tx \cdot \nabla u_0(tx) \\
&= \frac{1}{t} \partial_{tx} u_0(tx) = -\frac{1}{t} \beta(tx)\, u_0(tx) \\
&= -\frac{1}{t} \beta(tx) U(t,x)
\end{aligned}$$

and hence $u_0(x)$ must be given by $U(1,x)$ as in Eq. (71.4). Conversely if u_0 is defined by Eq. (71.4) then from Lemma 71.1, $u_0(sx) = U(s,x)$ and hence

$$\begin{aligned}
\partial_x u_0(x) &= \frac{d}{ds}|_1 u_0(sx) = \frac{d}{ds}|_1 U(s,x) \\
&= -\beta(x) \dot{U}(1,x) = -\beta(x) u_0(x).
\end{aligned}$$

Thus we have shown that u_0 solves Eq. (71.11).

We now turn our attention to solving Eq. (71.12). Assuming u_{k+1} is a solution to Eq. (71.12), then $V_{k+1}(t,x) := u_{k+1}(tx)$ satisfies

$$\begin{aligned}
t\dot{V}_{k+1}(t,x)(t) &= tx \cdot \nabla u_{k+1}(tx) \\
&= -\left(\frac{1}{t} \beta(tx) + k + 1\right) V_{k+1}(t,x) + L u_k(tx)
\end{aligned}$$

or equivalently

$$\dot{V}_{k+1}(t,x) = -\left(\frac{1}{t} \beta(tx) + \frac{k+1}{t}\right) V_{k+1}(t,x) + \frac{1}{t} L u_k(tx). \tag{71.13}$$

This equation may be solved by introducing an integrating factor, i.e. let

$$U_{k+1}(t,x) := t^{k+1} u_0(tx)^{-1} V_{k+1}(t,x).$$

Then U_{k+1} solves

$$\begin{aligned}\dot{U}_{k+1}(t,x) &= t^{k+1}u_0(tx)^{-1}\dot{V}_{k+1}(t,x) + (k+1)\,t^k u_0(tx)^{-1}V_{k+1}(t,x)\\ &\quad + t^{k+1}u_0(tx)^{-1}\frac{\beta(tx)}{t}V_{k+1}(t,x)\\ &= t^k u_0(tx)^{-1}Lu_k(tx).\end{aligned}$$

Hence

$$\begin{aligned}U_{k+1}(t,x) &= U_{k+1}(0,x) + \int_0^t \tau^k u_0(\tau x)^{-1}Lu_k(\tau x)d\tau\\ &= \int_0^t \tau^k u_0(\tau x)^{-1}Lu_k(\tau x)d\tau.\end{aligned}$$

Therefore if u_{k+1} exists it must be given by Eq. (71.5).

Conversely if u_{k+1} is defined by Eq. (71.5) then

$$\begin{aligned}u_{k+1}(sx) &= u_0(sx)\int_0^1 \tau^k u_0(\tau sx)^{-1}Lu_k(\tau sx)d\tau\\ &= u_0(sx)\int_0^s s^{-(k+1)}t^k u_0(tx)^{-1}Lu_k(tx)dt\end{aligned}$$

and hence

$$\begin{aligned}\partial_x u_{k+1}(x) &= \frac{d}{ds}|_1 u_{k+1}(sx)\\ &= -\beta(x)u_{k+1}(x) + u_0(x)\frac{d}{ds}|_1\int_0^s s^{-(k+1)}t^k u_0(tx)^{-1}Lu_k(tx)dt\\ &= -\beta(x)u_{k+1}(x) + Lu_k(x) - (k+1)u_0(x)\int_0^1 t^k u_0(tx)^{-1}Lu_k(tx)dt\\ &= -\left(\beta(x)+k+1\right)u_{k+1}(x) + Lu_k(x)\end{aligned}$$

which is shows that u_{k+1} solves Eq. (71.12). This finishes the proof of Theorem 71.3.

Part XVIII

PDE Extras

72

Higher Order Elliptic Equations

Definition 72.1. $H^s(\mathbb{R}^d, \mathbb{C}^N) := \{u \in \mathcal{S}'(\mathbb{R}^d, \mathbb{C}^N) : |\hat{u}| \in L^2((1+|\xi|^2)^s ds)\}$.

Note For $s = 0, 1, 2, \ldots$ this agrees with our previous Notations of Sobolev spaces.

Lemma 72.2. *(A) $C_C^\infty(\mathbb{R}^d, \mathbb{C}^N)$ is dense in $H^s(\mathbb{R}^d, \mathbb{C}^N)$.*
(B) If $s > 0$ then $\|u\|_{-s} = \|u\|_{H^{-s}} = \sup_{\phi \in C_c^\infty} \frac{|<u,\phi>|}{\|\phi\|_{-s}}$.

Proof. (A) We first rate that $\mathcal{S}(\mathbb{R}^d, \mathbb{C}^N)$ is dense in $H^s(\mathbb{R}^d, \mathbb{C}^N)$ because $\mathcal{F}\mathcal{S} = \mathcal{S}$ and $\mathcal{S}$ is dense in $L^2(1+|\xi|^2)^s d\xi)$ for all $s \in \mathbb{R}$. It is easily seen that C_c^∞ is dense in $\mathcal{S}$ for all $s = 0, 1, 2, \ldots$ relative to H^s-norm. But this enough to prove (A) since for all $s \in \mathbb{R} \| \cdot \|_S \leq \| \cdot \|_k$ for some $k \in \mathbb{N}$.
(B)

$$\begin{aligned}\|u\|_{-s}^2 = \|\hat{u}\|_{L^2((1+|\xi|^2)^{-2}ds)} &= \sup_{\phi \in \mathcal{S}} \frac{|(\hat{u}, \phi)_{-s}|}{\|\phi\|_{-s}} \\ &= \sup_{\phi \in \mathcal{S}} \frac{|\int \hat{u}(\xi)\phi(\xi)(1+|\xi|^2)^{-s} d\xi|}{\sqrt{S|\phi(\xi)|^2(1+|\xi|^2)^{-s} d\xi}}.\end{aligned}$$

Let $\hat{\psi}(\xi) = \hat{\varpi}(\xi)(a+|\xi|^2)^{-s}$, the arbitrary element of $\mathcal{S}$ still to find

$$\begin{aligned}\|u\|_{-s} &= \sup_{\psi \in \mathcal{S}} \frac{|\int \hat{u}(\xi)\hat{\psi}(\xi) d\xi|}{\sqrt{\int |\hat{\psi}(\xi)|^2(1+|\xi|^2)^s d\xi}} \\ &= \sup_{\psi \in \mathcal{S}} \frac{|<u,\psi>|}{\|\psi\|_s} = \sup_{\psi \in C_c^\infty} \frac{|<u,\psi>|}{\|\psi\|_s}\end{aligned}$$

Since C_c^∞ is dense as well. ∎

Definition 72.3. $H^s(\mathbb{R}^d, \mathbb{C}^N) := \{u \in \mathcal{S}'(\mathbb{R}^d, \mathbb{C}^N) : |\hat{u}| \in L^2((1+|\xi|^2)^s d\xi)\}$.

Proposition 72.4. *Suppose $L = L_0$ has constant coefficients then $\|u\|_s \leq C(\|Lu\|_{s-k} + \|u\|_{s-1})$ for all $s \in \mathbb{R}$.*

Proof.

$$\begin{aligned}\widehat{Lu}(s)\text{``}=\text{''}L\int \widehat{u(x)e^{ix-\xi}}dx &= \int \sigma(L_0)\widehat{(x,\xi)u}(x)e^{ix-\xi}dx \\ &= \int \sigma(\xi)u(x)e^{ix\cdot\xi}dx = \sigma(\xi)\hat{u}(\xi).\end{aligned}$$

Therefore

$$\begin{aligned}\|\hat{u}(\xi)\| = \|\sigma(\xi)^{-1}\sigma(\xi)\hat{u}(\xi)\| &\leq \|\sigma(\xi)^{-1}\| \, \|\sigma(\xi)\hat{u}(\xi)\| \\ &\leq \frac{\epsilon}{|\xi|^k}\|\sigma(\xi)\hat{u}(\xi)\|.\end{aligned}$$

Therefore

$$\|\sigma(\xi)\hat{u}(\xi)\| \geq \frac{|\xi|^k}{C_1}\|\hat{u}(\xi)\|$$

Notice there exist $C > 0$ such that

$$1 \leq C\left(|\xi|^{2k}\left(\frac{1}{1+|x\imath|^2}\right)^k + \frac{1}{1+|\xi|^2}\right).$$

Then

$$\begin{aligned}|\hat{u}|^2(1+|\xi|^2)^2 &\leq C(|\xi|^{2k}(1+|\xi|^2)^{s-k}|\hat{u}|^2 + (1+|\xi|^2)^{s-1}|\hat{u}|^2) \\ &\leq \widetilde{C}\left((1+|\xi|^2)^{s-k}|\sigma(\xi)\hat{u}|^2 + (1+|\xi|^2)^{s-1}(\hat{u}(\xi)^2\right) \\ &\leq \widetilde{C}\left((1+|\xi|^2)^{s-k}|\widehat{Lu}(\xi)|^2 + (1+|\xi|^2)^{s-1}(\hat{u}(\xi)^2\right).\end{aligned}$$

Integrate this on ξ to get the desired inequality. ∎

Notation 72.5 *Suppose $\Omega \subset \mathbb{R}^d$ is an open set and $a_\alpha \in C^\infty(\overline{\Omega}, \mathbb{C}^N)$ for some N and $|\alpha| \leq k$. Set $L : C^\infty(\Omega, \mathbb{C}^N) \to C^\infty(\Omega, \mathbb{C}^N)$ to be the operator $L = \sum_{|\alpha| \leq k} a_\alpha \partial^\alpha$ and $L_0 := \sum_{|\alpha|=k} a_\alpha \partial^\alpha$*

$$\sigma(L_0)(\xi) = \sum_{|\alpha|=k} a_\alpha (i\xi)^\alpha = i^{|\alpha|} \sum_{|\alpha|=k} a_\alpha \xi^\alpha$$

for $\xi \in \mathbb{R}^d$. Notice that $L_0 e^{i\xi\cdot x} = \sigma(L_0)(x,\xi)e^{i\xi\cdot x}$.

Definition 72.6. *L is elliptic at $x \in \Omega$ if $\sigma(L_0)(x,\xi)^{-1}$ exists for all $\xi \neq 0$ and L is elliptic on $\overline{\Omega}$ if $\sigma(L_0)(x,\xi)^{-1}$ exist for all $\xi \neq 0$, $x \in \overline{\Omega}$.*

Remark 72.7. If L is elliptic on $\overline{\Omega}$ then there exist $\in> 0$ and $C > 0$ such that $\|\sigma(L_0)(x,\xi)\| \geq \epsilon|\xi|^k$ for all $x \in \overline{\Omega}, \xi \in \mathbb{R}^d$. Also $\|\sigma(L_0)(x,\xi)^{-1}\| \leq C|\xi|^k$ for all $x \in \overline{\Omega}, \xi \in \mathbb{R}^d \setminus \{0\}$.

Indeed for $|\xi| = 1, (x,\xi) \to \|\sigma(L_0)(x,\xi)^{-1}\|$ is a continuous function of (x,ξ) on $\overline{\Omega} \times S^{d-1}$. Therefore it has a global maximum say C. Then for $|\xi| = t \neq 0$.

$\|\sigma(L_0)\left(x,\frac{\xi}{t}\right)^{-1}\| \le C$ implies $\left\|\left[\left(\frac{1}{t}\right)^k \sigma(L_0,x,\xi)\right]^{-1}\right\| \le C$ or

$$\left\|\sigma(L_0)(x,\xi)^{-1}\right\| \le Ct^{-k} = C|\xi|^{-k}.$$

Given Ω as above let

$$H_0^s(\Omega,\mathbb{C}^N) := \overline{C_C^\infty(\Omega,\mathbb{C}^N)} H^s(\mathbb{R}^d,\mathbb{C}^N).$$

Theorem 72.8 (A priori Estimates). *For all $s \in \mathbb{R}$ there exist $C > 0$ such that $\forall\ u \in H_s^0(\Omega)$*

$$\|u\|_s \le C(\|Lu\|_{s-k} + \|u\|_{s-1}). \tag{72.1}$$

Reference *Theorem (6.28) of Folland p. 210 chapter 6.*

Proof. I will only prove the inequality (72.1) for $s = 0$. (However, negative s are needed to prove desired elliptic regularity results. This could also be done using the Theory of pseudo differential operators.) With out loss of generality we may assume $u \in C_c^\infty(\Omega,\mathbb{C}^N)$.

Step (1) The inequality holds if $L = L_0$ with constant coefficients by above proposition.

Step (2) Suppose now $L = L_0$ but does not have constant coefficients. Define $L_{x_0} = \sum_{|\alpha|=k} a_\alpha(x_0)\partial^\alpha$ for all $x_0 \in \Omega$. Then there exist C_0 independent of x_0 such that

$$\|u\|_0 \le C_0(\|L_{x_0}u\|_{-k} + \|u\|_{-1})$$

for all $x_0 \in \Omega$. Suppose supp $(u) \subset B(x_0,\delta)$ with δ small. Consider

$$\|(L - L_{x_0})u\|_{-k} = \sup_{\phi \in C_C^\infty} \frac{|((L - L_{x_0})u,\phi)|}{\|\phi\|_s}.$$

Now

$$\begin{aligned}((L - L_{x_0})u,\phi) &= (-1)^k \sum (u, \partial^\alpha(a_\alpha^+ - a_\alpha(x_0))\phi) \\ &= (-1)^k (u, \sum (a_\alpha^+ - a_\alpha(x_0)\partial^\alpha\phi + s\phi).\end{aligned}$$

Where $s\phi = \sum_{|\alpha|=k} [\partial^\alpha, a_\alpha^+]\phi$ is a $k-1$ order differential operator. Therefore

$$\begin{aligned}|((L = L_{x_0})u,\phi)| &\le |(s^+u,\phi)| + \sum_\alpha |(u,(a_\alpha^+ - a_\alpha(x_0))\partial^\alpha\phi)| \\ &\le \|S^+u\|_{-k}\|\phi\|_k \\ &\quad + \underbrace{\sum \sup_{\alpha|x-x_0|\le\delta} |a_\alpha(x) - a_\alpha(x_0)|}_{C(\delta)} \|u\|_0\|\phi\|_k \\ &\le k\|u\|_{-1}\|\phi\|_k + C(\delta)\|u\|_0\|\phi\|_k.\end{aligned}$$

Therefore

$$\|(L - L_{x_0})u\|_{-k} \le C(\delta)\|u\|_0 + K\|u\|_{-1}.$$

Choose δ small such that $C_0C(\delta) \le \frac{1}{2}$ by unit continuous. Therefore

$$\|(L - L_{x_0})u\|_{-k} \le \frac{1}{2}\|u\|_0 + K\|u\|_{-1}.$$

Hence

$$\begin{aligned}\|u\|_0 &\le C_0(\|L_{x_0}u\|_{-k} + \|u\|_{-1}) \\ &\le C_0(\|L - L_{x_0}u\|_{-k} + \|Lu\|_{-k} + \|u\|_{-1}) \\ &\le C_0(C(\delta)\|u\|_0 + \|Lu\|_{-k} + (1+K)\|u\|_{-1}) \\ &\le \frac{1}{2}\|u\|_0 + C_0(1+K)(\|Lu\|_{-k} + \|u\|_{-1}).\end{aligned}$$

Therefore

$$\|u\|_0 \le 2C_0(1+K)(\|Lu\|_{-k} + \|u\|_{-1})$$

provided supp $u \subset B(x_0,\delta)$ for some $x_0 \in R$. Now cover $\overline{\Omega}$ by finite collection of balls with radius δ and choose a partition of unity subordinate to this cover. $\{x_i\}$. Therefore if $u \in C_c^\infty(\Omega)$, then $\sum \psi_i u = u$ and

$$\begin{aligned}\|u\|_0 &\le C(\|Lu\|_{-k} + \|su\|_{-k} + \|u\|_{-1}) \\ &\le \widetilde{C}(\|Lu\|_{-k} + \|u\|_{-1}).\end{aligned}$$

■

73

Abstract Evolution Equations

73.1 Basic Definitions and Examples

Let $(X, \|\cdot\|)$ be a normed vector space. A linear operator L on X consists of a subspace $\mathcal{D}(L)$ of X and a linear map $L : D(L) \to X$.

Notation 73.1 *Given a function $v : [0,\infty) \to X$, we write $v(t) = e^{tL}v(0)$, provided that $v \in C([0,\infty) \to X) \cap C^1((0,\infty) \to X)$, $v(t) \in \mathcal{D}(L)$ for all $t > 0$, and $\dot{v}(t) = Lv(t)$.*

Example 73.2. Suppose that L is an $n \times n$ matrix (thought of as a linear transformation on $\mathbb{C}^n$) and $v_0 \in \mathbb{C}^n$. Let $v(t) := \sum_{n=0}^{\infty} \frac{t^n}{n!} L^n v_0$, then the sum converges and $v(t) = e^{tL}v_0$.

The following proposition generalizes the above example.

Proposition 73.3 (Evolution). *Suppose that $(X, \|\cdot\|)$ is a Banach space and $L \in B(X)$–the Banach space of bounded operators on X with the operator norm. Then*

$$e^{tL} = \sum_{n=0}^{\infty} \frac{t^n}{n!} L^n \tag{73.1}$$

is convergent in the norm topology on $B(X)$. Moreover, if $v_0 \in X$ and $v(t) := e^{tL}v_0$, then v is the unique function in $C^1(\mathbb{R}, X)$ solving the differential equation:

$$\dot{v}(t) = Lu(t) \text{ with } v(0) = v_0. \tag{73.2}$$

Proof. First notice that $\sum_{n=0}^{\infty} \frac{|t|^n}{n!} \|L\|^n = e^{|t|\|L\|} < \infty$ so that the sum in (73.1) exists in $B(X)$ and $\|e^{tL}\| \le e^{|t|\|L\|}$. Let us now check that $\frac{d}{dt}e^{tL} = Le^{tL} = e^{tL}L$. Using the mean value theorem we have,

$$\begin{aligned} e^{(t+h)L} - e^{tL} &= \sum_{n=0}^{\infty} \frac{1}{n!}\{(t+h)^n - t^n\}L^n \\ &= \sum_{n=1}^{\infty} \frac{n}{n!} c_n^{n-1}(h) hL^n = hL \sum_{n=0}^{\infty} \frac{1}{n!} c_{n+1}^n(h) L^n, \end{aligned}$$

where $c_n(h)$ is some number between t and $t+h$ for each n. Hence

$$\frac{e^{(t+h)L} - e^{tL}}{h} - Le^{tL} = L \sum_{n=0}^{\infty} \frac{1}{n!} [c_{n+1}^n(h) - t^n] L^n,$$

and thus

$$\|\frac{e^{(t+h)L} - e^{tL}}{h} - Le^{tL}\| \le \|L\| \sum_{n=1}^{\infty} \frac{1}{n!} |c_{n+1}^n(h) - t^n| \|L\|^n \to 0,$$

as $n \to \infty$ by the dominated convergence theorem.

Before continuing, let us prove the basic group property of e^{tL}, namely:

$$e^{tL}e^{sL} = e^{(t+s)L}. \tag{73.3}$$

To prove this equation, notice that

$$\frac{d}{dt} e^{-tL} e^{(t+s)L} = e^{-tL}(-L + L)e^{(t+s)L} = 0.$$

Thus $e^{-tL}e^{(t+s)L}$ is independent of t and hence

$$e^{-tL}e^{(t+s)L} = e^{sL}. \tag{73.4}$$

By choosing $s = 0$ we find that $e^{-tL}e^{tL} = I$, and by replacing t by $-t$ we can conclude that $e^{-tL} = (e^{tL})^{-1}$. This last observation combined with (73.4) proves (73.3).

Alternate Proof of Eq. (73.3).

$$\begin{aligned} e^{(t+s)L} &= \sum_{n=0}^{\infty} \frac{(t+s)^n}{n!} L^n = \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{t^k L^k s^{n-k} L^{n-k}}{k!(n-k)!} \\ &= \sum_{n=0}^{\infty} \sum_{\substack{k+\ell=n \\ b,\ell \ge 0}} \frac{t^k L^k s^\ell L^\ell}{k!\ell!} = e^{tL}e^{sL}, \end{aligned}$$

where the above manipulations are justified since,

$$\sum_{n=0}^{\infty} \sum_{\substack{k+\ell=n \\ b,\ell \ge 0}} \frac{|t|^k |s|^\ell \|L\|^k \|L\|^\ell}{k!l!} = e^{(|t|+|s|)\|L\|} < \infty.$$

Now clearly if

$$v(t) := e^{tL}v_0 := \sum_{n=0}^{\infty} \frac{t^n}{n!} L^n v_0$$

we have that $v \in C^1(\mathbb{R} \to X)$ and $\dot{v} = Lv$. Therefore v does solve Eq. (73.2). To see that this solution is unique, suppose that $v(t)$ is any solution to (73.2). Then

$$\frac{d}{dt} e^{-tL} v(t) = e^{-tL}(-L + L)v(t) = 0,$$

so that $e^{-tL}v(t)$ is constant and thus $e^{-tL}v(t) = v_0$. Therefore $v(t) = e^{tL}v_0$. ■

Theorem 73.4 (The Diagonal Case). *Consider $t > 0$ only now. Let $p \in [1, \infty)$, $(\Omega, \mathcal{F}, m)$ be a measure space and $a : \Omega \to \mathbb{R}$ be a measurable function such that a is bounded above by a constant $C < \infty$. Define $D(L) = \{f \in L^p(m) : af \in L^p\}$, and for $f \in D(L)$ set $Lf = af = M_a f$. (In general L is an unbounded operator.) Define $e^{tL} = M_e ta f$. (Note that $|e^{ta}| \le e^{tC} < \infty$, so e^{tL} is a bounded operator.) by D.C.T. So one has $e^{tL} \to I$ strongly as $t \downarrow 0$.*

1. *e^{tL} is a strongly continuous semi-group of bounded operators.*
2. *If $f \in D(L)$ then $\frac{d}{dt} e^{tL} f = e^{tL} L f$ in $L^p(m)$.*
3. *For all $f \in L^p$ and $t > 0$, $\frac{d}{dt} e^{tL} f = L e^{tL} f$ in $L^p(m)$.*

Proof. By the dominated convergence theorem,

$$\left\| e^{tL} f - f \right\|_{L^p}^p = \int_\Omega \left| \left(e^{ta} - 1\right) f \right|^p dm \to 0 \text{ as } t \downarrow 0,$$

which proves item 1. For item 2 we see using the fundamental theorem of calculus that

$$\begin{aligned} \left\| \left(\frac{e^{(t+h)L} - e^{tL}}{h} - e^{tL} L \right) f \right\|_{L^p} &= \left\| \frac{1}{h} \int_0^h (a e^{(t+\tau)a} - e^{ta} a) d\tau \cdot f \right\|_{L^p} \\ &= \left\| \frac{1}{h} \int_0^h (e^{(t+\tau)a} - e^{ta}) d\tau \cdot af \right\|_{L^p}. \end{aligned} \tag{73.5}$$

Since $af \in L^p$,

$$\left| \frac{1}{h} \int_0^h (e^{(t+\tau)a} - e^{ta}) d\tau \right| \le 2e^{(t+|h|)C},$$

and

$$\frac{1}{h} \int_0^h (e^{(t+\tau)a} - e^{ta}) d\tau \to 0 \text{ as } h \to 0,$$

the Dominated convergence shows that the last term in Eq. (73.5) tends to zero as $h \to 0$. The above computations also work at $t = 0$ provided h is restricted to be positive.

Item 3 follows by the same techniques as item 2. We need only notice that by basic calculus if $t > 0$ and $\tau \in (t/2, 3t/2)$ then

$$\left| a e^{(t+\tau)a} \right| \le \max\{(t/2)^{-1}, C e^{3C/2}\}.$$

■

Example 73.5 (Nilpotent Operators). Let $L : X \to X$ be a nilpotent operator, i.e., $\forall v \in D(L)$ there exists $n = n(v)$ such that $L^n v = 0$ Then

$$\dot{v}(t) = Lv(t) \text{ with } v(0) = v \in D(L)$$

has a solution (in $D(L)$) given by

$$v(t) = e^{tL} v := \sum_{n=0}^{\infty} \frac{t^n}{n!} L^n v.$$

A special case of the last example would be to take $X = \mathcal{D}(L)$ to be the space of polynomial functions on $\mathbb{R}^d$ and $Lp = \Delta p$.

Example 73.6 (Eigenvector Case). Let $L : X \to X$ be a linear operator and suppose that $D(L) = \text{span} X_0$, where X_0 is a subset of X consisting of eigenvectors for L, i.e., $\forall v \in X_0$ there exists $\lambda(v) \in \mathbf{C}$ such the $Lv = \lambda(v)v$. Then

$$\dot{v}(t) = Lv(t) \text{ with } v(0) = v \in D(L)$$

has a solution in $D(L)$ given by

$$v(t) = e^{tL} v := \sum_{n=0}^{\infty} \frac{t^n}{n!} L^n v.$$

More explicitly, if $v = \sum_{v=1}^n v_i$ with $v_i \in X_0$, then

$$e^{tL} v = \sum_{i=1}^n e^{t\lambda(v_i)} v_i.$$

The following examples will be covered in more detail in the exercises.

Example 73.7 (Translation Semi-group). Let $X := L^2(\mathbb{R}^d, d\lambda)$, $w \in \mathbb{R}^d$ and

$$(T_w(t) f)(x) := f(x + t).$$

Then $T_w(t)$ is a strongly continuous contraction semi-group. In fact $T_w(t)$ is unitary for all $t \in \mathbb{R}$.

Example 73.8 (Rotation Semi-group). Suppose that $X := L^2(\mathbb{R}^d, d\lambda)$ and $O : \mathbb{R} \to O(d)$ is a one parameter semi-group of orthogonal operators. Set $(T_O(t) f)(x) := f(O(t)x)$ for all $f \in X$ and $x \in \mathbb{R}^d$. Then T_O is also a strongly continuous unitary semi-group.

73.2 General Theory of Contraction Semigroups

For this section, let $(X, \|\cdot\|)$ be a Banach space with norm $\|\cdot\|$. Also let $T := \{T(t)\}_{t>0}$ be a collection of bounded operators on X.

Definition 73.9. *Let X and T be as above.*

1. *T is a* ***semi-group*** *if $T(t+s) = T(t)T(s)$ for all $s, t > 0$.*
2. *A semi-group T is* ***strongly continuous*** *if $\lim_{t\downarrow 0} T(t)v = v$ for all $v \in X$. By convention if T is strongly continuous, set $T(0) \equiv I$–the identity operator on X.*
3. *A semi-group T is a* ***contraction semi-group*** *if $\|T(t)\| \le 1$ for $t > 0$.*

Definition 73.10. *Suppose that T is a contraction semi group. Set*

$$\mathcal{D}(L) := \{v \in X : \frac{d}{dt}|_{0+} T(t)v \text{ exists in } X\}$$

and for $v \in \mathcal{D}(L)$ set $Lv := \frac{d}{dt}|_{0+} T(t)v$. L is called the ***infinitesimal generator*** *of T.*

Proposition 73.11. *Let T be a strongly continuous contraction semi-group, then*

1. *For all $v \in X$, $t \in [0, \infty) \to T(t)v \in X$ is continuous.*
2. *$D(L)$ is dense linear subspace of X.*
3. *Suppose that $v : [0, \infty) \to X$ is a continuous, then $w(t) := T(t)v(t)$ is also continuous on $[0, \infty)$.*

Proof. By assumption $v(t) := T(t)v$ is continuous at $t = 0$. For $t > 0$ and $h > 0$,

$$\|v(t+h) - v(t)\| = \|T(t)(T(h) - I)v\| \le \|v(h) - v\| \to 0 \text{ as } h \downarrow 0.$$

Similarly if $h \in (0, t)$,

$$\|v(t-h) - v(t)\| = \|T(t-h)(I - T(h))v\| \le \|v - v(h)\| \to 0 \text{ as } h \downarrow 0.$$

This proves the first item.

Let $v \in X$ set $v_s := \int_0^s T(\sigma)v d\sigma$, where, since $\sigma \to T(\sigma)v$ is continuous, the integral may be interpreted as X–valued Riemann integral. Note

$$\|\frac{1}{s}v_s - v\| = \|\frac{1}{s}\int_0^s (T(\sigma)v - v)d\sigma\| \le \frac{1}{|s|}\left|\int_0^s \|T(\sigma)v - v\| d\sigma\right| \to 0$$

as $s \downarrow 0$, so that $\mathcal{D} := \{v_s : s > 0 \text{ and } v \in X\}$ is dense in X. Moreover,

$$\begin{aligned}\frac{d}{dt}\Big|_{0+} T(t)v_s &= \frac{d}{dt}\Big|_{0+} \int_0^s T(t+\sigma)v d\sigma \\ &= \frac{d}{dt}\Big|_{0+} \int_t^{t+s} T(\tau)v d\tau = T(s)v - v.\end{aligned}$$

Therefore $v_s \in \mathcal{D}(L)$ and $Lv_s = T(s)v - v$. In particular, $\mathcal{D} \subset \mathcal{D}(L)$ and hence $\mathcal{D}(L)$ is dense in X. It is easily checked that $\mathcal{D}(L)$ is a linear subspace of X.

Finally if $v : [0, \infty) \to X$ is a continuous function and $w(t) := T(t)v(t)$, then for $t > 0$ and $h \in (-t, \infty)$,

$$w(t+h) - w(t) = (T(t+h) - T(t))v(t) + T(t+h)(v(t+h) - v(t))$$

The first term goes to zero as $h \to 0$ by item 1 and the second term goes to zero since v is continuous and $\|T(t+h)\| \le 1$. The above argument also works with $t = 0$ and $h \ge 0$. ■

Definition 73.12 (Closed Operators). *A linear operator L on X is said to be closed if $\Gamma(L) := \{(v, Lv) \in X \times X : v \in \mathcal{D}(L)\}$ is closed in the Banach space $X \times X$. Equivalently, L is closed iff for all sequences $\{v_n\}_{n=1}^\infty \subset \mathcal{D}(L)$ such that $\lim_{n\to\infty} v_n =: v$ exists and $\lim_{n\to\infty} Lv_n =: w$ exists implies that $v \in \mathcal{D}(L)$ and $Lv = w$.*

Roughly speaking, a closed operator is the next best thing to a bounded (i.e. continuous) operator. Indeed, the definition states that L is closed iff

$$\lim_{n\to\infty} Lv_n = L \lim_{n\to\infty} v_n \tag{73.6}$$

provided **both** limits in (73.6) exist. While L is continuous iff Eq. (73.6) holds whenever $\lim_{n\to\infty} v_n$ exists: part of the assertion being that the limit on the left side of Eq. (73.6) should exist.

Proposition 73.13 (L is Closed). *Let L be the infinitesimal generator of a contraction semi-group, then L is* ***closed***.

Proof. Suppose that $v_n \in D(L)$, $v_n \to v$, and $Lv_n \to w$ in X as $n \to \infty$. Then, using the fundamental theorem of calculus,

$$\begin{aligned}\frac{T(t)v - v}{t} = \lim_{n\to\infty} \frac{T(t)v_n - v_n}{t} &= \lim_{n\to\infty} \frac{1}{t}\int_0^t T(\tau)Lv_n d\tau \\ &\to \frac{1}{t}\int_0^t T(\tau)w d\tau.\end{aligned}$$

Therefore $v \in D(L)$ and $Lv = w$. ■

Theorem 73.14 (Solution Operator).

1. *For any $t > 0$ and $v \in \mathcal{D}(L)$, $T(t)v \in \mathcal{D}(L)$ and $\frac{d}{dt}T(t)v = LT(t)v$.*
2. *Moreover if $v \in \mathcal{D}(L)$, then $\frac{d}{dt}T(t)v = T(t)Lv$.*

Proof. $T(t) : D(L) \to D(L)$ and $\frac{d}{dt}T(t)v = LT(t)v = T(t)Lv$. Suppose that $v \in \mathcal{D}(L)$, then

$$\frac{T(t+h) - T(t)}{h}v = \frac{(T(h) - I)}{h}T(t)v = \frac{T(t)(T(h) - I)}{h}v.$$

Letting $h \downarrow 0$ in the last set of equalities show that $T(t)v \in \mathcal{D}(L)$ and

$$\frac{d}{dh}|_{0+}T(t+h)v = LT(t)v = T(t)Lv.$$

For the derivative from below we will use,

$$\frac{T(t-h)-T(t)}{h}v = \frac{(I-T(h))}{h}T(t-h)v = \frac{T(t-h)(I-T(h))}{h}v, \quad (73.7)$$

which is valid for $t>0$ and $h \in [0,\infty)$. Set $u(h) := h^{-1}(I-T(h))v$ if $h>0$ and $u(0) := Lv$. Then $u : [0,\infty) \to X$ is continuous. Hence by same argument as in the proof of item 3 of Proposition 73.11), $h \to T(t-h)w(h)$ is continuous at $h=0$ and hence

$$\frac{T(t-h)(I-T(h))}{h}v = T(t-h)w(h) \to T(t-0)w(0) = T(t)Lv \text{ as } h \downarrow 0.$$

Thus it follows from Eq. (73.7) that, for all $t>0$,

$$\frac{d}{dh}|_{0-}T(t+h)v = LT(t)v = T(t)Lv.$$

■

Definition 73.15 (Evolution Equation). *Let T be a strongly continuous contraction semi-group with infinitesimal generator L. A function $v : [0,\infty) \to X$ is said to solve the differential equation*

$$\dot{v}(t) = Lv(t) \quad (73.8)$$

if

1. *$v(t) \in \mathcal{D}(L)$ for all $t \geq 0$,*
2. *$v \in C([0,\infty) \to X) \cap C^1((0,\infty) \to X)$, and*
3. *Eq. (73.8) holds for all $t>0$.*

Theorem 73.16 (Evolution Equation). *Let T be a strongly continuous contraction semi-group with infinitesimal generator L. The for all $v_0 \in \mathcal{D}(L)$, there is a unique solution to (73.8) such that $v(0) = v_0$.*

Proof. We have already shown existence. Namely by Theorem 73.14 and Proposition 73.11, $v(t) := T(t)v_0$ solves (73.8.

For uniqueness let v be any solution of (73.8). Fix $\tau > 0$ and set $w(t) := T(\tau - t)v(t)$. By item 3 of Proposition 73.11, w is continuous for $t \in [0,\tau]$. We will now show that w is also differentiable on $(0,\tau)$ and that $\dot{w} := 0$.

To simplify notation let $P(t) := T(t-t)$ and for fixed $t \in (0,\tau)$ and $h>0$ sufficiently small let $\epsilon(h) := h^{-1}(v(t+h)-v(t)) - \dot{v}(t)$. Since v is differentiable, $\epsilon(h) \to 0$ as $h \to 0$. Therefore,

$$\begin{aligned}\frac{w(t+h)-w(t)}{h} &= \frac{1}{h}\left[P(t+h)v(t+h) - P(t)v(t)\right] \\ &= \frac{(P(t+h)-P(t))}{h}v(t) + P(t+h)\frac{(v(t+h)-v(t))}{h} \\ &= \frac{(P(t+h)-P(t))}{h}v(t) + P(t+h)(\dot{v}(t)+\epsilon(h)) \\ &\to -P(t)Lv(t) + P(t)\dot{v}(t) \text{ as } h \to 0,\end{aligned}$$

wherein we have used $\|P(t+h)\epsilon(h)\| \leq \|\epsilon(h)\| \to 0$ as $h \to 0$. Hence we have shown that

$$\dot{w}(t) = -P(t)Lv(t) + P(t)\dot{v}(t) = -P(t)Lv(t) + P(t)Lv(t) = 0$$

Therefore $w(t) = T(t-t)v(t)$ is constant or $(0,t)$ and hence by continuity of $w(\tau) = w(0)$, i.e.

$$v(\tau) = w(\tau) = w(0) = T(\tau)v(0) = T(\tau)v_0.$$

This proves uniqueness. ■

Corollary 73.17. *Suppose that T and $\hat{T}$ are two strongly continuous contraction semi-groups on a Banach space X which have the same infinitesimal generators L. Then $T = \hat{T}$.*

Proof. Let $v_0 \in D(L)$ then $v(t) = T(t)v$ and $\hat{v}(t) = \hat{T}(t)\hat{v}$ both solve Eq. (73.8 with initial condition v_0. By Theorem 73.16, $v = \hat{v}$ which implies that $T(t)v_0 = \hat{T}(t)v_0$, i.e., $T = \hat{T}$.

Because of the last corollary the following notion is justified. ■

Notation 73.18 *If T is a strongly continuous contraction semi-group with infinitesimal generator L, we will write $T(t)$ as e^{tL}.*

Remark 73.19. Since T is a contraction, L should be "negative." Thus, working informally,

$$\int_0^\infty e^{-t\lambda}e^{tL}dt = \frac{1}{L-\lambda}e^{t(L-\lambda)}|_{t=0}^\infty = \frac{1}{\lambda - L} = (\lambda - L)^{-1}.$$

Theorem 73.20. *Suppose $T = e^{tL}$ is a strongly continuous contraction semi-group with infinitesimal generator L. For any $\lambda > 0$ the integral*

$$\int_0^\infty e^{-t\lambda}e^{tL}dt =: R_\lambda \quad (73.9)$$

exists as a $B(X)$–valued improper Riemann integral.[1] *Moreover, $(\lambda - L) : \mathcal{D}(L) \to X$ is an invertible operator, $(\lambda - L)^{-1} = R_\lambda$, and $\|R_\lambda\| \leq \lambda^{-1}$.*

[1] This may also be interpreted as a Bochner integral, since $T(t)$ is continuous and thus has separable range in $B(X)$.

Proof. First notice that

$$\int_0^\infty e^{-t\lambda}\|e^{tL}\|dt \le \int_0^\infty e^{-t\lambda}dt = 1/\lambda.$$

Therefore the integral in Eq. (73.9) exists and the result, R_λ, satisfies $\|R_\lambda\| \le \lambda^{-1}$. So we now must show that $R_\lambda = (\lambda - L)^{-1}$.

Let $v \in X$ and $h > 0$, then

$$e^{hL}R_\lambda v = \int_0^\infty e^{-t\lambda}e^{(t+h)L}vdt = \int_h^\infty e^{-(t-h)\lambda}e^{tL}vdt = e^{h\lambda}\int_h^\infty e^{-t\lambda}e^{tL}vdt. \tag{73.10}$$

Therefore

$$\frac{d}{dh}\Big|_{0+} e^{hL}R_\lambda v = -v + \int_0^\infty \lambda e^{-t\lambda}e^{tL}vdt = -v + \lambda R_\lambda v,$$

which shows that $R_\lambda v \in \mathcal{D}(L)$ and that $LR_\lambda v = -v+\lambda R_\lambda v$. So $(\lambda-L)R_\lambda = I$. Similarly,

$$R_\lambda e^{hL}v = e^{h\lambda}\int_h^\infty e^{-t\lambda}e^{tL}vdt \tag{73.11}$$

and hence if $v \in \mathcal{D}(L)$, then

$$R_\lambda Lv = \frac{d}{dh}\Big|_{0+} R_\lambda e^{hL}v = -v + \lambda R_\lambda v.$$

Hence $R_\lambda(\lambda - L) = I_{\mathcal{D}(L)}$. ■

Before continuing it will be useful to record some properties of the resolvent operators $R_\lambda := (\lambda - L)^{-1}$. Again working formally for the moment, if $\lambda, \mu \in (0, \infty)$, then we expect

$$R_\lambda - R_\mu = \frac{1}{\lambda - L} - \frac{1}{\mu - L} = \frac{\mu - L - (\lambda - L)}{(\lambda - L)(\mu - L)} = (\mu - \lambda)R_\lambda R_\mu.$$

For each $\lambda > 0$ define $L_\lambda := \lambda L R_\lambda$. Working again formally we have that

$$L_\lambda = \frac{\lambda L}{\lambda - L} = \frac{\lambda(L - \lambda + \lambda)}{\lambda - L} = -\lambda + \lambda^2 R_\lambda$$

and

$$\frac{d}{d\lambda}L_\lambda = \frac{L(\lambda - L) - \lambda L}{(\lambda - L)^2} = -\frac{L^2}{(\lambda - L)^2} = -LR_\lambda LR_\lambda.$$

These equations will be verified in the following lemma.

Lemma 73.21. *Let $L : X \to X$ be an operator on X such that for all $\lambda > 0$, $\lambda - L$ is invertible with a bounded inverse R_λ or $(\lambda - L)^{-1}$. Set $L_\lambda := \lambda L R_\lambda$. Then*

1. *for $\lambda, \mu \in (0, \infty)$,*
$$R_\lambda - R_\mu = (\mu - \lambda)R_\lambda R_\mu, \tag{73.12}$$
and in particular R_λ and R_μ commute,
2. *$L_\lambda = -\lambda + \lambda^2 R_\lambda$, and*
3. *$\frac{d}{d\lambda}L_\lambda = -LR_\lambda LR_\lambda$.*

Proof. Since $\lambda - L$ is invertible, $\lambda - L$ is injective. So in order to verify Eq. (73.12 it suffices to verify:

$$(\lambda - L)(R_\lambda - R_\mu) = (\mu - \lambda)(\lambda - L)R_\lambda R_\mu. \tag{73.13}$$

Now

$$(\lambda - L)(R_\lambda - R_\mu) = I - (\lambda - \mu + \mu - L)R_\mu = I - (\lambda - \mu)R_\mu - I = -(\lambda - \mu)R_\mu,$$

while

$$(\mu - \lambda)(\lambda - L)R_\lambda R_\mu = (\mu - \lambda)R_\mu.$$

Clearly the last two equations show that Eq. (73.13) holds. The second item is easily verified since, $L_\lambda = \lambda(L - \lambda + \lambda)R_\lambda = -\lambda + \lambda^2 R_\lambda$.

For the third item, first recall that $\lambda \to R_\lambda$ is continuous in the operator norm topology (in fact analytic). To see this let us first work informally,

$$R_{\lambda+h} = \frac{1}{\lambda + h - L} = \frac{1}{(\lambda - L + h)} = \frac{1}{(\lambda - L)(I + h(\lambda - L)^{-1})} = R_\lambda(I+hR_\lambda)^{-1}. \tag{73.14}$$

To verify this last equation, first notice that for sufficiently small h, $\|hR_\lambda\| < 1$, so that $\sum_{n=0}^\infty \|-hR_\lambda\|^n < \infty$ and hence $(I + hR_\lambda)$ is invertible and

$$(I + hR_\lambda)^{-1} = \sum_{n=0}^\infty (-hR_\lambda)^n.$$

To verify the ends of Eq. (73.14) are equal it suffices to verify that $R_{\lambda+h}(I + hR_\lambda) = R_\lambda$, i.e., $R_{\lambda+h} - R_\lambda = -hR_{\lambda+h}R_\lambda$. But this last equation follows directly form (73.12). Therefore, we have shown that for h sufficiently close to zero,

$$R_{\lambda+h} = R_\lambda \sum_{n=0}^\infty (-hR_\lambda)^n.$$

Differentiating this last equation at $h = 0$ shows that

$$\frac{d}{d\lambda}R_\lambda = \frac{d}{dh}|_0 R_{\lambda+h} = -R_\lambda^2. \tag{73.15}$$

We now may easily compute:

$$\frac{d}{d\lambda}L_\lambda = \frac{d}{d\lambda}(-\lambda + \lambda^2 R_\lambda) = -I + 2\lambda R_\lambda - \lambda^2 R_\lambda^2 = -(\lambda R_\lambda - I)^2.$$

This finishes the proof since,

$$\lambda R_\lambda - I = \lambda R_\lambda - (\lambda - L)R_\lambda = LR_\lambda$$

We now show that L_λ is a good approximation to L when $\lambda \to \infty$. ∎

Proposition 73.22. *Let L be an operator on X such that for each* $\lambda \in (0,\infty)$, $R_\lambda := (\lambda - L)^{-1}$ *exists as a bounded operator and* $\|(\lambda - L)^{-1}\| \leq \lambda^{-1}$. *Then* $\lambda R_\lambda \to I$ *strongly as* $\lambda \to \infty$ *and for all* $v \in \mathcal{D}(L)$

$$\lim_{\lambda\to\infty} L_\lambda v = Lv. \tag{73.16}$$

Proof. First notice, informally, that

$$\lambda R_\lambda = \frac{\lambda}{\lambda - L} = \frac{\lambda - L + L}{\lambda - L} = I + R_\lambda L.$$

So we expect that

$$\lambda R_\lambda|_{\mathcal{D}(L)} = I + R_\lambda L. \tag{73.17}$$

(This last equation is easily verified by applying $(\lambda - L)$ to both sides of the equation.) Hence, for $v \in \mathcal{D}(L)$,

$$\lambda R_\lambda v = v + R_\lambda L v,$$

and $\|R_\lambda L v\| \leq \|Lv\|/\lambda$. Thus $\lim_{\lambda\to\infty} \lambda R_\lambda v = v$ for all $v \in \mathcal{D}(L)$. Using the fact that $\|\lambda R_\lambda\| \leq 1$ and a standard 3ϵ–argument, it follows that λR_λ converges strongly to I as $\lambda \to \infty$. Finally, for $v \in \mathcal{D}(L)$,

$$L_\lambda v = \lambda L R_\lambda v = \lambda R_\lambda L v \to Lv \text{ as } \lambda \to \infty.$$

See Dynkin, ∎

Lemma 73.23. *Suppose that A and B are commuting bounded operators on a Banach space, X, such that* $\|e^{tA}\|$ *and* $\|e^{tB}\|$ *are bounded by* 1 *for all* $t > 0$, *then*

$$\|(e^{tA} - e^{tB})v\| \leq t\|(A - B)v\| \text{ for all } v \in X. \tag{73.18}$$

Proof. The fundamental theorem of calculus implies that

$$e^{-tA}e^{tB} - I = \int_0^t \frac{d}{d\tau} e^{-\tau A}e^{\tau B} d\tau = \int_0^t e^{-\tau A}(-A + B)e^{\tau B} d\tau,$$

and hence, by multiplying on the left by e^{tA},

$$\begin{aligned} e^{tB} - e^{tA} &= \int_0^t e^{(t-\tau)A}(-A + B)e^{\tau B} d\tau \\ &= \int_0^t e^{(t-\tau)A}e^{\tau B}(-A + B) d\tau, \end{aligned}$$

wherein the last line we have used the fact that A and B commute. Eq. (73.18) is an easy consequence of this equation. ∎

Theorem 73.24 (Hile-Yosida). *A closed linear operator L or a Banach space X generates a contraction semi-group iff for all* $\lambda \in (0,\infty)$,

1. $(\lambda - L)^{-1}$ *exists as a bounded operator and*
2. $\|(\lambda - L)^{-1}\| \leq \lambda^{-1}$ *for all* $\lambda > 0$.

Proof.

$$T_\lambda(t) := e^{tL_\lambda} = e^{tL_\lambda} := \sum_{n=0}^\infty (tL_\lambda)^n/n!.$$

The outline of the proof is: i) show that $T_\lambda(t)$ is a contraction for all $t > 0$, ii) show for $t > 0$ that $T_\lambda(t)$ converges strongly to an operator $T(t)$, iii) show $T(t)$ is a strongly continuous contraction semi-group, and iv) show that L is the generator of T.

Step i) Using $L_\lambda = -\lambda + \lambda^2 R_\lambda$, we find that $e^{tL_\lambda} = e^{-t\lambda}e^{t\lambda^2 R_\lambda}$. Hence

$$\|T_\lambda(t)\| = \|e^{tL_\lambda}\| \leq e^{-t\lambda}e^{t\lambda^2\|R_\lambda\|} \leq e^{-t\lambda}e^{t\lambda^2\lambda^{-1}} = 1.$$

Step ii) Let $\alpha, \mu > 0$ and $v \in \mathcal{D}(L)$, then by Lemma 73.23 and Proposition 73.22,

$$\|(T_\alpha(t) - T_\mu(t))v\| \leq t\|L_\alpha v - L_\mu v\| \to 0 \text{ as } \alpha, \mu \to \infty.$$

This shows, for all $v \in \mathcal{D}(L)$, that $\lim_{\alpha\to\infty} T_\alpha(t)v$ exists uniformly for t in compact subsets of $[0,\infty)$ For general $v \in X$, $w \in \mathcal{D}(L)$, $\tau > 0$, and $0 \leq t \leq \tau$, we have

$$\begin{aligned} \|(T_\alpha(t) - T_\mu(t))v\| &\leq \|(T_\alpha(t) - T_\mu(t))w\| + \|(T_\alpha(t) - T_\mu(t))(v - w)\| \\ &\leq \|(T_\alpha(t) - T_\mu(t))w\| + 2\|v - w\|. \\ &\leq \tau\|L_\alpha w - L_\mu w\| + 2\|v - w\|. \end{aligned}$$

Thus

$$\limsup_{\alpha,\mu\to\infty} \sup_{t\in[0,\tau]} \|(T_\alpha(t) - T_\mu(t))v\| \leq 2\|v - w\| \to 0 \text{ as } w \to v.$$

Hence for each $v \in X$, $T(t)v := \lim_{\alpha\to\infty} T_\alpha(t)v$ exists uniformly for t in compact sets of $[0,\infty)$.

Step iii) It is now easily follows that $\|T(t)\| \leq 1$ and that $t \to T(t)$ is strongly continuous. Moreover,

$$T(t+s)v = \lim_{\alpha\to\infty} T_\alpha(t+s)v = \lim_{\alpha\to\infty} T_\alpha(t)T_\alpha(s)v.$$

Letting $\epsilon(\alpha) := T_\alpha(s)v - T(s)v$, it follows that

$$T(t+s)v = \lim_{\alpha\to\infty} T_\alpha(t)(T(s)v + \epsilon(\alpha)) = T(t)T(s)v + \lim_{\alpha\to\infty} T_\alpha(t)\epsilon(\alpha).$$

This shows that T is also satisfies the semi-group property, since $\|T_\alpha(t)\epsilon(\alpha)\| \leq \|\epsilon(\alpha)\|$ and $\lim_{\alpha\to\infty} \epsilon(\alpha) = 0$.

Step iv) Let $\tilde{L}$ denote the infinitesimal generator of T. We wish to show that $\tilde{L} = L$. To this end, let $v \in \mathcal{D}(L)$, then

$$T_\lambda(t)v = v + \int_0^t e^{\tau L_\lambda} L_\lambda v d\tau.$$

Letting $\lambda \to \infty$ in this last equation shows that

$$T(t)v = v + \int_0^t T(\tau)Lv\, d\tau,$$

and hence $\frac{d}{dt}|_{0+}T(t)v$ exists and is equal to Lv. That is $\mathcal{D}(L) \subset \mathcal{D}(\tilde{L})$ and $L = \tilde{L}$ on $\mathcal{D}(L)$.

To finish the proof we must show that $\mathcal{D}(\tilde{L}) \subset \mathcal{D}(L)$. Suppose that $\tilde{v} \in \mathcal{D}(\tilde{L})$ and $\lambda > 0$ and let $v := (\lambda - L)^{-1}(\lambda - \tilde{L})\tilde{v}$. Since $\mathcal{D}(L) \subset \mathcal{D}(\tilde{L})$, $(\lambda - \tilde{L})\tilde{v} = (\lambda - L)v = (\lambda - \tilde{L})v$ and because $\lambda - \tilde{L}$ is invertible, $\tilde{v} = v \in \mathcal{D}(L)$. ■

Theorem 73.25. *Let L be a closed operator or Hilbert space H. Then L generates a contraction semi-group $T(t)$ iff there exits $\lambda_0 > 0$ such that $Ran(L + \lambda_0) = H$ and $\mathrm{Re}(Lv, v) \le 0$ for all $v \in D(L)$.*

Proof. ($\Rightarrow$) If $T(t) = e^{tL}$ is a contraction semigroup, then, for all $v \in D(L)$, $\|e^{tL}v\|^2 \le \|v\|^2$ with equality at $t = 0$. So it is permissible to differentiate the inequality at $t = 0$ to find $2\,\mathrm{Re}(Lv, v) \le 0$. The remaining assertions in this direction follows from Theorem 73.24. ($\Longleftarrow$) If $\mathrm{Re}(Lv, v) \le 0$ and $\lambda > 0$ then

$$\|(\lambda - L)v\|^2 = \lambda^2\|v\|^2 - 2\lambda\,\mathrm{Re}(Lv, v) + \|Lv\|^2 \ge \lambda^2\|v\|^2$$

which implies $\lambda - L$ is 1-1 on $D(L)$ and $\mathrm{Ran}(\lambda - L)$ is closed. ■

Theorem 73.26. *Let $P_t : X \to X$ be a contraction semi-group. $D(L) = \left\{v \in X | Lv = \frac{d}{dt}|_{0t}P_t v\right\}$ exists Then for all $v \in D(L)$ $\quad v(t) = P_t v$ is the unique solution to*

$$\dot{v}(t) = Lv(t)v(0) = v.$$

Lemma 73.27. *$D(L)$ is dense in X.*

Proof. Let $\varphi \in C_c^\infty((0, \infty))$ set

$$P_\phi v := \int_0^\infty \varphi(s)P_s v ds$$

then

$$P_t P_\varphi v = \int_0^\infty \varphi(s)P_{t+s} v ds = \int_0^\infty \varphi(s-t)P_s v ds = P_{\varphi(0-t)}v$$

$$\frac{P_t P_\varphi v - P_\varphi v}{t} = \int_0^\infty \frac{\varphi(s-t) - \varphi(s)}{t} P_s v ds$$
$$\to \int_0^\infty -\varphi'(s)P_s v ds \frac{d}{dt}|_0 P_t P_\varphi v = -P_{\varphi'}v$$

so $P_\varphi v \in D(L)$ and $LP_\varphi v = -P_{\varphi'}v$ Now $\varphi \to \delta_0$. ■

Proof of Theorem Key point is to prove uniqueness. Let $P_t^* =$ transpose semi group. Let $B = \frac{d}{dt}|_{0+}P_t^*$ denote the adjoint generator.

Claim $B = L^*$.

$$\varphi \in D(B) \Rightarrow \frac{d}{dt}|_0\langle P_t^*\varphi, v\rangle = \langle B\varphi, v\rangle$$
$$\frac{d}{dt}|_0\langle \varphi, P_t v\rangle = \langle \varphi, Lv\rangle \qquad \forall v \in D(Lt$$

$\Rightarrow \varphi \in D(L^*)$ and $B\varphi = L\varphi$ so that $B \subseteq L^*$. Suppose $\varphi \in D(L^*) \Rightarrow \exists\psi \ni L^*\varphi = \psi$. For example $\langle L^*\varphi, v\rangle = \langle \psi, v\rangle \qquad \forall v \in D(L)$ for example $\langle \varphi, Lv\rangle = \langle \psi, v\rangle$

$$\frac{d}{dt}\langle \varphi, P_t v\rangle = \langle \psi, P_t v\rangle.$$

So $\langle \varphi, P_t v\rangle - \langle \varphi, v\rangle = \int_0^t \langle P_t^*\psi, v\rangle dt \qquad \forall v \in D(L)$. So $P_t^*\varphi = \varphi + \int_0^t P_t^*\psi dt$. So $\frac{d}{dt}|_{0+}P_t^*\varphi = \psi \Rightarrow \varphi \in D(L^*)$ and $L^*\varphi = \varphi$.

Uniqueness: Let $\dot{v} = Lv, \quad v(0) = v_0$. Let $\varphi \in D(L^*), T > 0$.

$$\frac{d}{dt}\langle P_{T-t}^*\varphi, v(t)\rangle = \langle -L^*P_{T-t}^*\varphi, v(t)\rangle + \langle P_{T-t}^*\varphi, Lv\rangle = 0$$

$\Rightarrow \langle P_{T-t}^*\varphi, v(T)\rangle =$ construction therefore $\langle P_T^*\varphi, v_0\rangle = \langle \varphi, v(T)\rangle \qquad \forall \varphi \in D(L^*)\langle \varphi'', P_T v_0\rangle$. Since $D(L^*)$ is dense $v(T) = P_T v_0$.

Part XIX

Appendices

A

Multinomial Theorems and Calculus Results

Given a multi-index $\alpha \in \mathbb{Z}_+^n$, let $|\alpha| = \alpha_1 + \cdots + \alpha_n$, $\alpha! := \alpha_1! \cdots \alpha_n!$,

$$x^\alpha := \prod_{j=1}^n x_j^{\alpha_j} \text{ and } \partial_x^\alpha = \left(\frac{\partial}{\partial x}\right)^\alpha := \prod_{j=1}^n \left(\frac{\partial}{\partial x_j}\right)^{\alpha_j}.$$

We also write

$$\partial_v f(x) := \frac{d}{dt} f(x+tv)|_{t=0}.$$

A.1 Multinomial Theorems and Product Rules

For $a = (a_1, a_2, \dots, a_n) \in \mathbb{C}^n$, $m \in \mathbb{N}$ and $(i_1, \dots, i_m) \in \{1, 2, \dots, n\}^m$ let $\hat{\alpha}_j(i_1, \dots, i_m) = \#\{k : i_k = j\}$. Then

$$\left(\sum_{i=1}^n a_i\right)^m = \sum_{i_1,\dots,i_m=1}^n a_{i_1} \dots a_{i_m} = \sum_{|\alpha|=m} C(\alpha) a^\alpha$$

where

$$C(\alpha) = \#\{(i_1, \dots, i_m) : \hat{\alpha}_j(i_1, \dots, i_m) = \alpha_j \text{ for } j = 1, 2, \dots, n\}$$

I claim that $C(\alpha) = \frac{m!}{\alpha!}$. Indeed, one possibility for such a sequence $(a_1, \dots, a_{i_m})$ for a given α is gotten by choosing

$$(\overbrace{a_1, \dots, a_1}^{\alpha_1}, \overbrace{a_2, \dots, a_2}^{\alpha_2}, \dots, \overbrace{a_n, \dots, a_n}^{\alpha_n}).$$

Now there are $m!$ permutations of this list. However, only those permutations leading to a distinct list are to be counted. So for each of these $m!$ permutations we must divide by the number of permutation which just rearrange the

groups of a_i's among themselves for each i. There are $\alpha! := \alpha_1! \cdots \alpha_n!$ such permutations. Therefore, $C(\alpha) = m!/\alpha!$ as advertised. So we have proved

$$\left(\sum_{i=1}^n a_i\right)^m = \sum_{|\alpha|=m} \frac{m!}{\alpha!} a^\alpha. \tag{A.1}$$

Now suppose that $a, b \in \mathbb{R}^n$ and α is a multi-index, we have

$$(a+b)^\alpha = \sum_{\beta \le \alpha} \frac{\alpha!}{\beta!(\alpha-\beta)!} a^\beta b^{\alpha-\beta} = \sum_{\beta+\delta=\alpha} \frac{\alpha!}{\beta!\delta!} a^\beta b^\delta \tag{A.2}$$

Indeed, by the standard Binomial formula,

$$(a_i + b_i)^{\alpha_i} = \sum_{\beta_i \le \alpha_i} \frac{\alpha_i!}{\beta_i!(\alpha_i - \beta_i)!} a^{\beta_i} b^{\alpha_i - \beta_i}$$

from which Eq. (A.2) follows. Eq. (A.2) generalizes in the obvious way to

$$(a_1 + \dots + a_k)^\alpha = \sum_{\beta_1+\dots+\beta_k=\alpha} \frac{\alpha!}{\beta_1! \cdots \beta_k!} a_1^{\beta_1} \dots a_k^{\beta_k} \tag{A.3}$$

where $a_1, a_2, \dots, a_k \in \mathbb{R}^n$ and $\alpha \in \mathbb{Z}_+^n$.

Now let us consider the product rule for derivatives. Let us begin with the one variable case (write $d^n f$ for $f^{(n)} = \frac{d^n}{dx^n} f$) where we will show by induction that

$$d^n(fg) = \sum_{k=0}^n \binom{n}{k} d^k f \cdot d^{n-k} g. \tag{A.4}$$

Indeed assuming Eq. (A.4) we find

$$\begin{aligned} d^{n+1}(fg) &= \sum_{k=0}^n \binom{n}{k} d^{k+1} f \cdot d^{n-k} g + \sum_{k=0}^n \binom{n}{k} d^k f \cdot d^{n-k+1} g \\ &= \sum_{k=1}^{n+1} \binom{n}{k-1} d^k f \cdot d^{n-k+1} g + \sum_{k=0}^n \binom{n}{k} d^k f \cdot d^{n-k+1} g \\ &= \sum_{k=1}^{n+1} \left[\binom{n}{k-1} + \binom{n}{k}\right] d^k f \cdot d^{n-k+1} g + d^{n+1} f \cdot g + f \cdot d^{n+1} g. \end{aligned}$$

Since

$$\begin{aligned} \binom{n}{k-1} + \binom{n}{k} &= \frac{n!}{(n-k+1)!(k-1)!} + \frac{n!}{(n-k)!k!} \\ &= \frac{n!}{(k-1)!\,(n-k)!} \left[\frac{1}{(n-k+1)} + \frac{1}{k}\right] \\ &= \frac{n!}{(k-1)!\,(n-k)!} \frac{n+1}{(n-k+1)\,k} = \binom{n+1}{k} \end{aligned}$$

the result follows.

Now consider the multi-variable case

$$\partial^\alpha (fg) = \left(\prod_{i=1}^n \partial_i^{\alpha_i} \right) (fg) = \prod_{i=1}^n \left[\sum_{k_i=0}^{\alpha_i} \binom{\alpha_i}{k_i} \partial_i^{k_i} f \cdot \partial_i^{\alpha_i - k_i} g \right]$$
$$= \sum_{k_1=0}^{\alpha_1} \cdots \sum_{k_n=0}^{\alpha_n} \prod_{i=1}^n \binom{\alpha_i}{k_i} \partial^k f \cdot \partial^{\alpha-k} g = \sum_{k \le \alpha} \binom{\alpha}{k} \partial^k f \cdot \partial^{\alpha-k} g$$

where $k = (k_1, k_2, \dots, k_n)$ and

$$\binom{\alpha}{k} := \prod_{i=1}^n \binom{\alpha_i}{k_i} = \frac{\alpha!}{k!(\alpha-k)!}.$$

So we have proved

$$\partial^\alpha (fg) = \sum_{\beta \le \alpha} \binom{\alpha}{\beta} \partial^\beta f \cdot \partial^{\alpha-\beta} g. \tag{A.5}$$

A.2 Taylor's Theorem

Theorem A.1. *Suppose $X \subset \mathbb{R}^n$ is an open set, $x : [0,1] \to X$ is a C^1 – path, and $f \in C^N(X, \mathbb{C})$. Let $v_s := x(1) - x(s)$ and $v = v_1 = x(1) - x(0)$, then*

$$f(x(1)) = \sum_{m=0}^{N-1} \frac{1}{m!} \left(\partial_v^m f \right) (x(0)) + R_N \tag{A.6}$$

where

$$R_N = \frac{1}{(N-1)!} \int_0^1 \left(\partial_{\dot{x}(s)} \partial_{v_s}^{N-1} f \right) (x(s)) ds = \frac{1}{N!} \int_0^1 \left(-\frac{d}{ds} \partial_{v_s}^N f \right) (x(s)) ds. \tag{A.7}$$

and $0! := 1$.

Proof. By the fundamental theorem of calculus and the chain rule,

$$f(x(t)) = f(x(0)) + \int_0^t \frac{d}{ds} f(x(s)) ds = f(x(0)) + \int_0^t \left(\partial_{\dot{x}(s)} f \right) (x(s)) ds \tag{A.8}$$

and in particular,

$$f(x(1)) = f(x(0)) + \int_0^1 \left(\partial_{\dot{x}(s)} f \right) (x(s)) ds.$$

This proves Eq. (A.6) when $N = 1$. We will now complete the proof using induciton on N.

Applying Eq. (A.8) with f replaced by $\frac{1}{(N-1)!} \left(\partial_{\dot{x}(s)} \partial_{v_s}^{N-1} f \right)$ gives

$$\frac{1}{(N-1)!} \left(\partial_{\dot{x}(s)} \partial_{v_s}^{N-1} f \right) (x(s)) = \frac{1}{(N-1)!} \left(\partial_{\dot{x}(s)} \partial_{v_s}^{N-1} f \right) (x(0))$$
$$+ \frac{1}{(N-1)!} \int_0^s \left(\partial_{\dot{x}(s)} \partial_{v_s}^{N-1} \partial_{\dot{x}(t)} f \right) (x(t)) dt$$
$$= -\frac{1}{N!} \left(\frac{d}{ds} \partial_{v_s}^N f \right) (x(0)) - \frac{1}{N!} \int_0^s \left(\frac{d}{ds} \partial_{v_s}^N \partial_{\dot{x}(t)} f \right)$$

wherein we have used the fact that mixed partial derivatives commute to show $\frac{d}{ds} \partial_{v_s}^N f = N \partial_{\dot{x}(s)} \partial_{v_s}^{N-1} f$. Integrating this equation on $s \in [0,1]$ shows, using the fundamental theorem of calculus,

$$R_N = \frac{1}{N!} \left(\partial_v^N f \right) (x(0)) - \frac{1}{N!} \int_{0 \le t \le s \le 1} \left(\frac{d}{ds} \partial_{v_s}^N \partial_{\dot{x}(t)} f \right) (x(t)) ds dt$$
$$= \frac{1}{N!} \left(\partial_v^N f \right) (x(0)) + \frac{1}{(N+1)!} \int_{0 \le t \le 1} \left(\partial_{w_t}^N \partial_{\dot{x}(t)} f \right) (x(t)) dt$$
$$= \frac{1}{N!} \left(\partial_v^N f \right) (x(0)) + R_{N+1}$$

which completes the inductive proof. ■

Remark A.2. Using Eq. (A.1) with a_i replaced by $v_i \partial_i$ (although $\{v_i \partial_i\}_{i=1}^n$ are not complex numbers they are commuting symbols), we find

$$\partial_v^m f = \left(\sum_{i=1}^n v_i \partial_i \right)^m f = \sum_{|\alpha|=m} \frac{m!}{\alpha!} v^\alpha \partial^\alpha.$$

Using this fact we may write Eqs. (A.6) and (A.7) as

$$f(x(1)) = \sum_{|\alpha| \le N-1} \frac{1}{\alpha!} v^\alpha \partial^\alpha f(x(0)) + R_N$$

and

$$R_N = \sum_{|\alpha|=N} \frac{1}{\alpha!} \int_0^1 \left(-\frac{d}{ds} v_s^\alpha \partial^\alpha f \right) (x(s)) ds.$$

Corollary A.3. *Suppose $X \subset \mathbb{R}^n$ is an open set which contains $x(s) = (1 - s)x_0 + s x_1$ for $0 \le s \le 1$ and $f \in C^N(X, \mathbb{C})$. Then*

$$f(x_1) = \sum_{m=0}^{N-1} \frac{1}{m!} \left(\partial_v^m f \right) (x_0) + \frac{1}{N!} \int_0^1 \left(\partial_v^N f \right) (x(s)) d\nu_N(s) \tag{A.9}$$
$$= \sum_{|\alpha|<N} \frac{1}{\alpha!} \partial^\alpha f(x(0)) (x_1 - x_0)^\alpha + \sum_{\alpha : |\alpha|=N} \frac{1}{\alpha!} \left[\int_0^1 \partial^\alpha f(x(s)) d\nu_N(s) \right] (x_1 - \tag{A.10}$$

where $v := x_1 - x_0$ *and* $d\nu_N$ *is the probability measure on* $[0,1]$ *given by*

$$d\nu_N(s) := N(1-s)^{N-1}ds. \tag{A.11}$$

If we let $x = x_0$ *and* $y = x_1 - x_0$ *(so* $x + y = x_1$*) Eq. (A.10) may be written as*

$$f(x+y) = \sum_{|\alpha|<N} \frac{\partial_x^\alpha f(x)}{\alpha!} y^\alpha + \sum_{\alpha:|\alpha|=N} \frac{1}{\alpha!}\left(\int_0^1 \partial_x^\alpha f(x+sy)d\nu_N(s)\right) y^\alpha. \tag{A.12}$$

Proof. This is a special case of Theorem A.1. Notice that

$$v_s = x(1) - x(s) = (1-s)(x_1 - x_0) = (1-s)v$$

and hence

$$R_N = \frac{1}{N!}\int_0^1 \left(-\frac{d}{ds}(1-s)^N \partial_v^N f\right)(x(s))ds = \frac{1}{N!}\int_0^1 \left(\partial_v^N f\right)(x(s))N(1-s)^{N-1}ds$$

■

Example A.4. Let $X = (-1,1) \subset \mathbb{R}$, $\beta \in \mathbb{R}$ and $f(x) = (1-x)^\beta$. The reader should verify

$$f^{(m)}(x) = (-1)^m \beta(\beta-1)\dots(\beta-m+1)(1-x)^{\beta-m}$$

and therefore by Taylor's theorem (Eq. (**??**) with $x = 0$ and $y = x$)

$$(1-x)^\beta = 1 + \sum_{m=1}^{N-1} \frac{1}{m!}(-1)^m \beta(\beta-1)\dots(\beta-m+1)x^m + R_N(x) \tag{A.13}$$

where

$$\begin{aligned} R_N(x) &= \frac{x^N}{N!}\int_0^1 (-1)^N \beta(\beta-1)\dots(\beta-N+1)(1-sx)^{\beta-N}d\nu_N(s) \\ &= \frac{x^N}{N!}(-1)^N\beta(\beta-1)\dots(\beta-N+1)\int_0^1 \frac{N(1-s)^{N-1}}{(1-sx)^{N-\beta}}ds. \end{aligned}$$

Now for $x \in (-1,1)$ and $N > \beta$,

$$0 \le \int_0^1 \frac{N(1-s)^{N-1}}{(1-sx)^{N-\beta}}ds \le \int_0^1 \frac{N(1-s)^{N-1}}{(1-s)^{N-\beta}}ds = \int_0^1 N(1-s)^{\beta-1}ds = \frac{N}{\beta}$$

and therefore,

$$|R_N(x)| \le \frac{|x|^N}{(N-1)!}\left|(\beta-1)\dots(\beta-N+1)\right| =: \rho_N.$$

Since

$$\limsup_{N\to\infty} \frac{\rho_{N+1}}{\rho_N} = |x| \cdot \limsup_{N\to\infty} \frac{N-\beta}{N} = |x| < 1$$

and so by the Ratio test, $|R_N(x)| \le \rho_N \to 0$ (exponentially fast) as $N \to \infty$. Therefore by passing to the limit in Eq. (A.13) we have proved

$$(1-x)^\beta = 1 + \sum_{m=1}^\infty \frac{(-1)^m}{m!}\beta(\beta-1)\dots(\beta-m+1)x^m \tag{A.14}$$

which is valid for $|x| < 1$ and $\beta \in \mathbb{R}$. An important special cases is $\beta = -1$ in which case, Eq. (A.14) becomes $\frac{1}{1-x} = \sum_{m=0}^\infty x^m$, the standard geometric series formula. Another another useful special case is $\beta = 1/2$ in which case Eq. (A.14) becomes

$$\begin{aligned} \sqrt{1-x} &= 1 + \sum_{m=1}^\infty \frac{(-1)^m}{m!}\frac{1}{2}(\frac{1}{2}-1)\dots(\frac{1}{2}-m+1)x^m \\ &= 1 - \sum_{m=1}^\infty \frac{(2m-3)!!}{2^m m!}x^m \text{ for all } |x| < 1. \end{aligned} \tag{A.15}$$

B

Zorn's Lemma and the Hausdorff Maximal Principle

Definition B.1. *A partial order $\leq$ on X is a relation with following properties*

(i) If $x \leq y$ and $y \leq z$ then $x \leq z$.
(ii) If $x \leq y$ and $y \leq x$ then $x = y$.
(iii) $x \leq x$ for all $x \in X$.

Example B.2. Let Y be a set and $X = \mathcal{P}(Y)$. There are two natural partial orders on X.

1. Ordered by inclusion, $A \leq B$ is $A \subset B$ and
2. Ordered by reverse inclusion, $A \leq B$ if $B \subset A$.

Definition B.3. *Let $(X, \leq)$ be a partially ordered set we say X is **linearly** a **totally** ordered if for all $x, y \in X$ either $x \leq y$ or $y \leq x$. The real numbers $\mathbb{R}$ with the usual order $\leq$ is a typical example.*

Definition B.4. *Let $(X, \leq)$ be a partial ordered set. We say $x \in X$ is a **maximal** element if for all $y \in X$ such that $y \geq x$ implies $y = x$, i.e. there is no element larger than x. An **upper bound** for a subset E of X is an element $x \in X$ such that $x \geq y$ for all $y \in E$.*

Example B.5. Let

$$X = \left\{ a = \{1\}\ b = \{1, 2\}\ c = \{3\}\ d = \{2, 4\}\ e = \{2\} \right\}$$

ordered by set inclusion. Then b and d are maximal elements despite that fact that $b \nleq a$ and $a \nleq b$. We also have

- If $E = \{a, e, c\}$, then E has **no** upper bound.

Definition B.6. • *If $E = \{a, e\}$, then b is an upper bound.*
• *$E = \{e\}$, then b and d are upper bounds.*

Theorem B.7. *The following are equivalent.*

1. ***The axiom of choice:*** *to each collection, $\{X_\alpha\}_{\alpha \in A}$, of non-empty sets there exists a "choice function," $x : A \to \coprod_{\alpha \in A} X_\alpha$ such that $x(\alpha) \in X_\alpha$ for all $\alpha \in A$, i.e. $\prod_{\alpha \in A} X_\alpha \neq \emptyset$.*
2. ***The Hausdorff Maximal Principle:*** *Every partially ordered set has a **maximal** (relative to the inclusion order) linearly ordered subset.*
3. ***Zorn's Lemma:*** *If X is partially ordered set such that every linearly ordered subset of X has an upper bound, then X has a maximal element.*[1]

Proof. $(2 \Rightarrow 3)$ Let X be a partially ordered subset as in 3 and let $\mathcal{F} = \{E \subset X : E \text{ is linearly ordered}\}$ which we equip with the inclusion partial ordering. By 2. there exist a maximal element $E \in \mathcal{F}$. By assumption, the linearly ordered set E has an upper bound $x \in X$. The element x is maximal, for if $y \in Y$ and $y \geq x$, then $E \cup \{y\}$ is still an linearly ordered set containing E. So by maximality of E, $E = E \cup \{y\}$, i.e. $y \in E$ and therefore $y \leq x$ showing which combined with $y \geq x$ implies that $y = x$.[2]

$(3 \Rightarrow 1)$ Let $\{X_\alpha\}_{\alpha \in A}$ be a collection of non-empty sets, we must show $\prod_{\alpha \in A} X_\alpha$ is not empty. Let $\mathcal{G}$ denote the collection of functions $g : D(g) \to \coprod_{\alpha \in A} X_\alpha$ such that $D(g)$ is a subset of A, and for all $\alpha \in D(g)$, $g(\alpha) \in X_\alpha$. Notice that $\mathcal{G}$ is not empty, for we may let $\alpha_0 \in A$ and $x_0 \in X_\alpha$ and then set $D(g) = \{\alpha_0\}$ and $g(\alpha_0) = x_0$ to construct an element of $\mathcal{G}$. We now put a partial order on $\mathcal{G}$ as follows. We say that $f \leq g$ for $f, g \in \mathcal{G}$ provided that $D(f) \subset D(g)$ and $f = g|_{D(f)}$. If $\Phi \subset \mathcal{G}$ is a linearly ordered set, let $D(h) = \cup_{g \in \Phi} D(g)$ and for $\alpha \in D(g)$ let $h(\alpha) = g(\alpha)$. Then $h \in \mathcal{G}$ is an upper bound for Φ. So by Zorn's Lemma there exists a maximal element $h \in \mathcal{G}$. To finish the proof we need only show that $D(h) = A$. If this were not the case, then let $\alpha_0 \in A \setminus D(h)$ and $x_0 \in X_{\alpha_0}$. We may now define $D(\tilde{h}) = D(h) \cup \{\alpha_0\}$ and

$$\tilde{h}(\alpha) = \begin{cases} h(\alpha) \text{ if } \alpha \in D(h) \\ x_0 \ \text{ if } \ \alpha = \alpha_0. \end{cases}$$

[1] If X is a countable set we may prove Zorn's Lemma by induction. Let $\{x_n\}_{n=1}^\infty$ be an enumeration of X, and define $E_n \subset X$ inductively as follows. For $n = 1$ let $E_1 = \{x_1\}$, and if E_n have been chosen, let $E_{n+1} = E_n \cup \{x_{n+1}\}$ if x_{n+1} is an upper bound for E_n otherwise let $E_{n+1} = E_n$. The set $E = \cup_{n=1}^\infty E_n$ is a linearly ordered (you check) subset of X and hence by assumption E has an upper bound, $x \in X$. I claim that his element is maximal, for if there exists $y = x_m \in X$ such that $y \geq x$, then x_m would be an upper bound for E_{m-1} and therefore $y = x_m \in E_m \subset E$. That is to say if $y \geq x$, then $y \in E$ and hence $y \leq x$, so $y = x$. (Hence we may view Zorn's lemma as a " jazzed" up version of induction.)

[2] Similalry one may show that $3 \Rightarrow 2$. Let $\mathcal{F} = \{E \subset X : E \text{ is linearly ordered}\}$ and order $\mathcal{F}$ by inclusion. If $\mathcal{M} \subset \mathcal{F}$ is linearly ordered, let $E = \cup \mathcal{M} = \bigcup_{A \in \mathcal{M}} A$. If $x, y \in E$ then $x \in A$ and $y \in B$ for some $A, B \subset \mathcal{M}$. Now $\mathcal{M}$ is linearly ordered by set inclusion so $A \subset B$ or $B \subset A$ i.e. $x, y \in A$ or $x, y \in B$. Sinse A and B are linearly order we must have either $x \leq y$ or $y \leq x$, that is to say E is linearly ordered. Hence by 3. there exists a maximal element $E \in \mathcal{F}$ which is the assertion in 2.

Then $h \le \tilde{h}$ while $h \ne \tilde{h}$ violating the fact that h was a maximal element.

($1 \Rightarrow 2$) Let $(X, \le)$ be a partially ordered set. Let $\mathcal{F}$ be the collection of linearly ordered subsets of X which we order by set inclusion. Given $x_0 \in X$, $\{x_0\} \in \mathcal{F}$ is linearly ordered set so that $\mathcal{F} \ne \emptyset$.

Fix an element $P_0 \in \mathcal{F}$. If P_0 is not maximal there exists $P_1 \in \mathcal{F}$ such that $P_0 \subsetneq P_1$. In particular we may choose $x \notin P_0$ such that $P_0 \cup \{x\} \in \mathcal{F}$. The idea now is to keep repeating this process of adding points $x \in X$ until we construct a maximal element P of $\mathcal{F}$. We now have to take care of some details.

We may assume with out loss of generality that $\tilde{\mathcal{F}} = \{P \in \mathcal{F} : P \text{ is not maxima}$ is a non-empty set. For $P \in \tilde{\mathcal{F}}$, let $P^* = \{x \in X : P \cup \{x\} \in \mathcal{F}\}$. As the above argument shows, $P^* \ne \emptyset$ for all $P \in \tilde{\mathcal{F}}$. Using the axiom of choice, there exists $f \in \prod_{P \in \tilde{\mathcal{F}}} P^*$. We now define $g : \mathcal{F} \to \mathcal{F}$ by

$$g(P) = \begin{cases} P & \text{if } P \text{ is maximal} \\ P \cup \{f(x)\} & \text{if } P \text{ is not maximal.} \end{cases} \tag{B.1}$$

The proof is completed by Lemma B.8 below which shows that g must have a fixed point $P \in \mathcal{F}$. This fixed point is maximal by construction of g. ■

Lemma B.8. *The function $g : \mathcal{F} \to \mathcal{F}$ defined in Eq. (B.1) has a fixed point.*[3]

Proof. The **idea of the proof** is as follows. Let $P_0 \in \mathcal{F}$ be chosen arbitrarily. Notice that $\Phi = \left\{g^{(n)}(P_0)\right\}_{n=0}^{\infty} \subset \mathcal{F}$ is a linearly ordered set and it is therefore easily verified that $P_1 = \bigcup_{n=0}^{\infty} g^{(n)}(P_0) \in \mathcal{F}$. Similarly we may repeat the process to construct $P_2 = \bigcup_{n=0}^{\infty} g^{(n)}(P_1) \in \mathcal{F}$ and $P_3 = \bigcup_{n=0}^{\infty} g^{(n)}(P_2) \in \mathcal{F}$, etc. etc. Then take $P_\infty = \cup_{n=0}^{\infty} P_n$ and start again with P_0 replaced by P_∞. Then keep going this way until eventually the sets stop increasing in size, in which case we have found our fixed point. The problem with this strategy is that we may never win. (This is very reminiscent of constructing measurable sets and the way out is to use measure theoretic like arguments.)

Let us now start the **formal proof.** Again let $P_0 \in \mathcal{F}$ and let $\mathcal{F}_1 = \{P \in \mathcal{F} : P_0 \subset P\}$. Notice that $\mathcal{F}_1$ has the following properties:

1. $P_0 \in \mathcal{F}_1$.
2. If $\Phi \subset \mathcal{F}_1$ is a totally ordered (by set inclusion) subset then $\cup \Phi \in \mathcal{F}_1$.
3. If $P \in \mathcal{F}_1$ then $g(P) \in \mathcal{F}_1$.

Let us call a general subset $\mathcal{F}' \subset \mathcal{F}$ satisfying these three conditions a tower and let

[3] Here is an easy proof if the elements of $\mathcal{F}$ happened to all be finite sets and there existed a set $P \in \mathcal{F}$ with a maximal number of elements. In this case the condition that $P \subset g(P)$ would imply that $P = g(P)$, otherwise $g(P)$ would have more elements than P.

$$\mathcal{F}_0 = \cap \{\mathcal{F}' : \mathcal{F}' \text{ is a tower}\}.$$

Standard arguments show that $\mathcal{F}_0$ is still a tower and clearly is the smallest tower containing P_0. (Morally speaking $\mathcal{F}_0$ consists of all of the sets we were trying to constructed in the "idea section" of the proof.)

We now claim that $\mathcal{F}_0$ is a linearly ordered subset of $\mathcal{F}$. To prove this let $\Gamma \subset \mathcal{F}_0$ be the linearly ordered set

$$\Gamma = \{C \in \mathcal{F}_0 : \text{ for all } A \in \mathcal{F}_0 \text{ either } A \subset C \text{ or } C \subset A\}.$$

Shortly we will show that $\Gamma \subset \mathcal{F}_0$ is a tower and hence that $\mathcal{F}_0 = \Gamma$. That is to say $\mathcal{F}_0$ is linearly ordered. Assuming this for the moment let us finish the proof. Let $P \equiv \cup \mathcal{F}_0$ which is in $\mathcal{F}_0$ by property 2 and is clearly the largest element in $\mathcal{F}_0$. By 3. it now follows that $P \subset g(P) \in \mathcal{F}_0$ and by maximality of P, we have $g(P) = P$, the desired fixed point. So to finish the proof, we must show that Γ is a tower.

First off it is clear that $P_0 \in \Gamma$ so in particular Γ is not empty. For each $C \in \Gamma$ let

$$\Phi_C := \{A \in \mathcal{F}_0 : \text{either } A \subset C \text{ or } g(C) \subset A\}.$$

We will begin by showing that $\Phi_C \subset \mathcal{F}_0$ is a tower and therefore that $\Phi_C = \mathcal{F}_0$.

1. $P_0 \in \Phi_C$ since $P_0 \subset C$ for all $C \in \Gamma \subset \mathcal{F}_0$. 2. If $\Phi \subset \Phi_C \subset \mathcal{F}_0$ is totally ordered by set inclusion, then $A_\Phi := \cup \Phi \in \mathcal{F}_0$. We must show $A_\Phi \in \Phi_C$, that is that $A_\Phi \subset C$ or $C \subset A_\Phi$. Now if $A \subset C$ for all $A \in \Phi$, then $A_\Phi \subset C$ and hence $A_\Phi \in \Phi_C$. On the other hand if there is some $A \in \Phi$ such that $g(C) \subset A$ then clearly $g(C) \subset A_\Phi$ and again $A_\Phi \in \Phi_C$.

3. Given $A \in \Phi_C$ we must show $g(A) \in \Phi_C$, i.e. that

$$g(A) \subset C \text{ or } g(C) \subset g(A). \tag{B.2}$$

There are three cases to consider: either $A \subsetneq C$, $A = C$, or $g(C) \subset A$. In the case $A = C$, $g(C) = g(A) \subset g(A)$ and if $g(C) \subset A$ then $g(C) \subset A \subset g(A)$ and Eq. (B.2) holds in either of these cases. So assume that $A \subsetneq C$. Since $C \in \Gamma$, either $g(A) \subset C$ (in which case we are done) or $C \subset g(A)$. Hence we may assume that

$$A \subsetneq C \subset g(A).$$

Now if C were a proper subset of $g(A)$ it would then follow that $g(A) \setminus A$ would consist of at least two points which contradicts the definition of g. Hence we must have $g(A) = C \subset C$ and again Eq. (B.2) holds, so Φ_C is a tower.

It is now easy to show Γ is a tower. It is again clear that $P_0 \in \Gamma$ and Property 2. may be checked for Γ in the same way as it was done for Φ_C above. For Property 3., if $C \in \Gamma$ we may use $\Phi_C = \mathcal{F}_0$ to conclude for all $A \in \mathcal{F}_0$, either $A \subset C \subset g(C)$ or $g(C) \subset A$, i.e. $g(C) \in \Gamma$. Thus Γ is a tower and we are done. ■

References

1. Isaac Chavel, *Riemannian geometry—a modern introduction*, Cambridge University Press, Cambridge, 1993. MR 95j:53001
2. ———, *The Laplacian on Riemannian manifolds*, Spectral theory and geometry (Edinburgh, 1998), Cambridge Univ. Press, Cambridge, 1999, pp. 30–75. MR 2001c:58029
3. E. B. Davies, *Heat kernels and spectral theory*, Cambridge Tracts in Mathematics, vol. 92, Cambridge University Press, Cambridge, 1990. MR 92a:35035
4. ———, *Heat kernel bounds for higher order elliptic operators*, Journées "Équations aux Dérivées Partielles" (Saint-Jean-de-Monts, 1995), École Polytech., Palaiseau, 1995, pp. Exp. No. III, 11. MR 96i:35020
5. Enrico Giusti, *Minimal surfaces and functions of bounded variation*, Birkhäuser Verlag, Basel, 1984. MR 87a:58041
6. Ronald B. Guenther and John W. Lee, *Partial differential equations of mathematical physics and integral equations*, Dover Publications Inc., Mineola, NY, 1996, Corrected reprint of the 1988 original. MR 97e:35001
7. Serge Lang, *Real and functional analysis*, third ed., Graduate Texts in Mathematics, vol. 142, Springer-Verlag, New York, 1993. MR 94b:00005
8. Vladimir G. Maz'ja, *Sobolev spaces*, Springer Series in Soviet Mathematics, Springer-Verlag, Berlin, 1985, Translated from the Russian by T. O. Shaposhnikova. MR 87g:46056
9. A. Pazy, *Semigroups of linear operators and applications to partial differential equations*, Applied Mathematical Sciences, vol. 44, Springer-Verlag, New York, 1983. MR 85g:47061
10. Michael Reed and Barry Simon, *Methods of modern mathematical physics. II. Fourier analysis, self-adjointness*, Academic Press [Harcourt Brace Jovanovich Publishers], New York, 1975. MR 58 #12429b
11. ———, *Methods of modern mathematical physics. I*, second ed., Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1980, Functional analysis. MR 85e:46002
12. Laurent Saloff-Coste, *Aspects of Sobolev-type inequalities*, London Mathematical Society Lecture Note Series, vol. 289, Cambridge University Press, Cambridge, 2002. MR 2003c:46048

13. Robert Schatten, *Norm ideals of completely continuous operators*, Second printing. Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 27, Springer-Verlag, Berlin, 1970. MR 41 #2449
14. B. Simon, *Convergence in trace ideals*, Proc. Amer. Math. Soc. **83** (1981), no. 1, 39–43. MR 82h:47042
15. Barry Simon, *Trace ideals and their applications*, London Mathematical Society Lecture Note Series, vol. 35, Cambridge University Press, Cambridge, 1979. MR 80k:47048
16. S. L. Sobolev and E. R. Dawson, *Partial differential equations of mathematical physics*, Translated from the third Russian edition by E. R. Dawson; English translation edited by T. A. A. Broadbent, Pergamon Press, Oxford, 1964. MR 31 #2478
17. Elias M. Stein, *Singular integrals and differentiability properties of functions*, Princeton Mathematical Series, No. 30, Princeton University Press, Princeton, N.J., 1970. MR 44 #7280
18. Frank W. Warner, *Foundations of differentiable manifolds and Lie groups*, Graduate Texts in Mathematics, vol. 94, Springer-Verlag, New York, 1983, Corrected reprint of the 1971 edition. MR 84k:58001