



Partial Differential Equations



© Benoit Dionne, 2024 (University of Ottawa)

Adapted version from the notes for the courses MAT3320, MAT4130, MAT5133, and graduate courses in partial differential equations given at the University of Ottawa.

This document is available on the following sites.

uO Research: <http://hdl.handle.net/10393/49793>

GitHub: <https://github.com/BenoitDionne/PDE>



Unless otherwise stated, this book is made available under the terms of the license [Creative Commons Attribution - Non Commercial-Share Alike 4.0 International](https://creativecommons.org/licenses/by-nc-sa/4.0/) (CC BY-NC-SA 4.0)

Cover Page:

Rainbow, South of France, photo by Louise Oegema.

Ancient Theatre of Epidauros, Greece, photo by Jean Dionne.

Contents

Preface	1
Chapter 1 Method of Characteristics	9
1.1. Quasi-Linear First Order Equations	9
1.1.1. Lagrange Method	16
1.2. Envelope of a Family of Surfaces	18
1.3. General First Order Partial Differential Equations	19
1.3.1. Equations for x , y and u	20
1.3.2. Equations for p and q	23
1.3.3. initial conditions	25
1.4. Solution Generated as Envelope	35
1.5. Addendum	43
1.6. Exercises	44
Chapter 2 Classification of Partial Differential Equations	45
2.1. Linear Second Order Partial Differential Equation in the Plane	45
2.2. Characteristic curves	48
2.3. Classification of Linear Second Order Partial Differential Equation in the Plane	50
2.3.1. Elliptic Equations	51
2.3.2. Hyperbolic Equations	53
2.3.3. Parabolic Equations	58
2.4. More on the Classification of Partial Differential Equations	59
2.4.1. Classification of Second Order Semi-Linear Partial Differential Equations in the Plane	59
2.4.2. System of First Order Partial Differential Equations	60

2.4.3.	Classification of Linear Partial Differential Equations in $\mathbb{R}^n$	62
2.5.	Exercises	66
Chapter 3	Cauchy Problem	69
3.1.	Cauchy-Kovalevski's Theorem	69
3.2.	Lewy Example	72
3.3.	Exercises	73
Chapter 4	Review of Functional Analysis	75
4.1.	Hilbert Spaces	75
4.2.	General L^2 Spaces	79
4.2.1.	L^2 Spaces on the Real Line	80
4.2.2.	L^2 Spaces on the Unit Circle	81
4.2.3.	Pointwise Convergence of the Classical Fourier Series	82
4.3.	Classical Fourier Series of real Valued Functions	85
4.3.1.	Odd and Even Functions	87
4.3.2.	More Special Cases	90
4.3.3.	Fourier Series of Functions of Two Variables	91
4.4.	Operator Theory	91
4.4.1.	Bounded Operators	91
4.4.2.	Unbounded Operators	96
4.4.3.	Operators on Hilbert Spaces	98
4.5.	Lax-Milgram Theorem	102
4.6.	Exercises	108
Chapter 5	Review of some Special Functions	109
5.1.	Frobenius Method	109
5.2.	Sturm-Liouville Problems	119
5.3.	Legendre Polynomials $P_n(x)$ on $[-1, 1]$	123
5.4.	Associated Legendre Polynomials $P_n^m(x)$ on $] - 1, 1[$	128
5.5.	Gamma Function	129
5.6.	Bessel Functions	130
5.7.	Exercises	138

Chapter 6	Distributions and Strong Solutions	139
6.1.	Introduction	139
6.2.	Test Functions	140
6.3.	Distributions	141
6.4.	Derivatives of Distributions	144
6.5.	Convolution	149
6.6.	Fourier Transforms	155
6.7.	Fundamental Solutions	163
6.8.	Exercises	171
Chapter 7	Wave Equation in One-Dimension	173
7.1.	D'Alembert Method	173
7.1.1.	Over the Real Line	173
7.1.2.	Over a Semi-Finite Interval	177
7.1.3.	Over a Finite Interval	178
7.2.	Non-Homogeneous Wave Equation	180
7.3.	Strong Solution of the Wave equation	184
7.4.	Separation of variables	186
7.5.	Exercises	203
Chapter 8	Wave Equation in $\mathbb{R}^n$	205
8.1.	The Cauchy Problem	205
8.2.	Uniqueness of the Solution	206
8.3.	Laplace Operator	209
8.4.	Existence of the Solution	212
8.4.1.	One Dimensional Space	212
8.4.2.	Higher Odd Dimensional Spaces	212
8.4.3.	Even Dimensional Space	220
8.5.	Huygens Principle	221
8.6.	Nonhomogeneous Wave Equations	222
8.7.	Separation of Variables	224
8.7.1.	Cartesian Coordinates	224
8.7.2.	Polar, Cylindrical and Spherical Coordinates	226

8.8.	Exercises	230
Chapter 9	Laplace Equation	231
9.1.	Harmonic Functions	232
9.2.	Maximum Principle	236
9.3.	Fundamental Solutions	239
9.4.	Dirichlet and Neumann Problems	243
9.5.	Green's Functions	247
9.5.1.	Dirichlet Problem on the Half Space	256
9.5.2.	Dirichlet Problem in a Ball	261
9.6.	Harmonic Polynomials	270
9.7.	Reflection Principle	277
9.8.	Perron's Method	278
9.9.	Layer Potentials	285
9.9.1.	Uniqueness of solutions	293
9.9.2.	Necessary condition for the Neumann Problem	296
9.9.3.	Note on the Integration on Manifolds	297
9.9.4.	Integral Operators	302
9.9.5.	Double Layer Potential	308
9.9.6.	Single Layer Potential	318
9.9.7.	Existence of Solutions	328
9.10.	Separation of Variables	338
9.10.1.	Cartesian Coordinates	338
9.10.2.	Cylindrical and Spherical Coordinates	353
9.11.	Eigenvalue Problems for the Laplacian Operator	360
9.11.1.	Eigenvalue Problem in a Rectangle	360
9.11.2.	Eigenvalue Problem in a Disk	362
9.11.3.	Eigenvalue Problem on a Sphere	364
9.11.4.	Eigenvalue Problem in a Ball	366
9.12.	Exercises	368
Chapter 10	Heat Equation	369
10.1.	Cauchy Problem	370

10.1.1.	One Dimensional Heat (Diffusion) Equation	370
10.1.2.	Non-homogeneous Heat Equation, Diffusion with a Source	373
10.1.3.	Higher Dimensional Heat Equation	374
10.2.	Fundamental Solution	380
10.3.	Maximum Principle	382
10.4.	Separation of Variables	384
10.4.1.	Cartesian Coordinates	384
10.4.2.	Cylindrical and Spherical Coordinates	394
10.5.	Addendum	395
10.6.	Exercises	395
Chapter 11	Sobolev Spaces	397
11.1.	The $\mathbf{W}^{k,p}$ spaces	397
11.1.1.	Dense Subsets	401
11.1.2.	Localized Sobolev Spaces	413
11.2.	The $\mathbf{W}_0^{k,p}$ spaces	416
11.3.	The $\mathbf{H}^{k,p}$ spaces	420
11.4.	Sobolev Spaces on Manifolds	422
11.5.	Dual Spaces	423
11.6.	Extension	426
11.7.	Imbedding	433
11.7.1.	In $\mathbb{R}$	433
11.7.2.	In $\mathbb{R}^n$	435
11.8.	Trace Operators	442
11.8.1.	Appendix	456
11.9.	Nonexistence of Solutions	459
11.10.	Exercises	460
Chapter 12	Linear Elliptic Partial Differential Equations	461
12.1.	Introduction to Variational Methods	461
12.2.	Classical, Strong and Weak Solutions	466
12.2.1.	Dirichlet Problem	467
12.2.2.	Neumann Problem	470

12.3.	Adjoint Boundary Conditions	473
12.4.	Weakly Coercive Bilinear Mappings	476
12.5.	Existence and Uniqueness of Solutions	485
12.5.1.	Eigenvalue Problems for the Laplacian Operator (continued)	502
12.6.	Regularity of Weak Solutions	504
12.6.1.	Local Regularity	519
12.6.2.	Regularity at the Boundary	527
12.7.	Existence of Green's Functions	539
12.8.	Maximum Principle	541
12.8.1.	Strong Maximum Principle	541
12.8.2.	General Maximum Principle	548
12.9.	Addendum	552
12.10.	Exercises	552
Chapter 13	Shock Waves	555
13.1.	Discontinuous Solutions	556
13.2.	Strong Solutions	557
13.3.	Rankine-Hugoniot Formula	559
13.4.	Existence, Uniqueness and Asymptotic Behaviour of Solutions	564
13.4.1.	Rarefaction Waves	565
13.4.2.	Shock Waves	567
13.4.3.	Existence and Uniqueness	569
13.4.4.	Asymptotic Behaviour	578
13.5.	Systems of Conservation Laws	579
13.5.1.	Basic Concepts	589
13.5.2.	The Glimm Difference Scheme	597
13.5.3.	Entropy Pairs	600
13.5.4.	Viscosity Method	604
13.6.	Addendum	605
13.7.	Exercises	606
Bibliography		607
Index		611

Preface

This book is for an advanced undergraduate course on partial differential equations or for a beginning graduate course on partial differential equations.

There are excellent books on partial differential equations for graduate students (or advanced undergraduate students) who plan to pursue the study of partial differential equations for their thesis. However, these books often require a solid background in measure and integration, and functional analysis that students who are required to study partial differential equations do not have. As a consequence, it takes a lot of time for these students to study and understand these books because of their condensed style. This is often frustrating for students who do not plan to pursue their research in partial differential equations but are only taking a course on partial differential equation as an elective in their program. The present book covers most of the important topics found in books on partial differential equations, using most of the time the same approach, but provide a lot more details so that students do not have to struggle to understand the theory.

As we said above, university students now a day do not always learn as much measure and integration, and functional analysis as it used to be. Most of the excellent books on partial differential equations require a good knowledge of these topics. The present book is accessible to students with a minimal introduction to measure and integration, and functional analysis, and who are willing to accept some of the fundamentals results which are presented in Chapter 4. Authors often assume that their readers will fill in the standard details in analysis. Unfortunately, students will get stuck trying to understand the proofs or, worse, will completely miss the subtleties of some proofs. For instance, they will assume that all series and integrals converge, or that derivative and integral can always be interchanged. Every computation becomes a formal computation. We provide a lot of the details in analysis that are often assumed in books on partial differential equations. Hopefully, students will learn to appreciate the rigour of analysis and its *raison d'être*.

This book contains most of the material on partial differential equations usually taught at the undergraduate level for science and engineering students. The reason to include this material is that many mathematics students do not have the chance to study partial differential equations at the undergraduate level. Thus, their first contact with partial differential equations is at the graduate level where all the undergraduate material on partial differential equations is assumed. The sections covering the undergraduate material are accessible to readers with only a calculus background. The presentation in these sections is usually more formal. They are therefore appropriate for an introduction to partial differential equations

for science and engineering students. A first undergraduate course could possibly include:

- parts of Chapter 1;
- the introduction of Chapter 13 and Section 13.1;
- Chapter 2.
- Sections 4.2 and 4.3 (limited to the study of the classical Fourier series);
- part of Chapter 5;
- Chapter 7 excluding Section 7.3;
- Chapter 8 excluding Section 8.3 and the higher dimensional case in Section 8.4;
- Section 9.1, 9.2 and 9.10; and
- Chapter 10 excluding the higher dimension heat equation in Section 10.1 and Section 10.2.

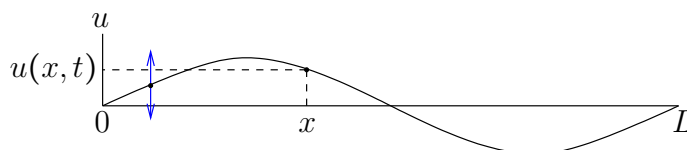
For the more advanced topics, the readers are expected to be familiar with all the concept of vector calculus in $\mathbb{R}^n$. These topics are still accessible to readers with a more standard vector calculus course in $\mathbb{R}^2$ and $\mathbb{R}^3$ if they assume that n is 2 or 3 when reading the sections covering these advanced topics. The readers may however find the presentation very technical in comparison to the simpler presentation which is limited to $n = 2$ or $n = 3$.

Except for Chapters 1 and 13, this book is about linear partial differential equations. So, there is no mention of reaction diffusion equations, solitons, inverse scattering, etc. In particular, there is no mention of semi-group theory. On this subject, we strongly recommend [12]. Energy methods are used a few times in the book but not as extensively as you will find in some books on partial differential equation.

It would be a sacrilege to not motivate heuristically the major partial differential equations that have had a major influence on the theory of partial differential equations. I am talking about the wave equation, the heat equation and the Laplace equation. We present them in the next sections.

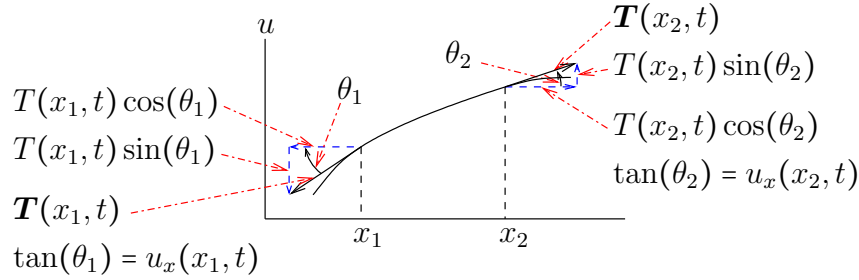
Wave equation

We consider a guitar or violin string of length L at rest. Since the string is really thin, we consider only the linear dimension of the string. Such strings are homogeneous. So, we may assume that the string is of constant density ρ . Moreover, we assume that the string is perfectly elastic when subject to very small vibration. Therefore, the tension in the string is tangent to the string. Moreover, because of the perfect elasticity, we also assume that each point on the string is only moving up and down in a vertical plane as illustrated in the figure below. There is no horizontal motion. We assume that the vertical motion is really small.



The vertical position of the string at a position x on the string and at a time t is denoted $u(x, t)$, and the vector tension in the string at that position and time t is denoted $\mathbf{T}(x, t)$. We consider a very small section of the string and apply Newton's law "force equals mass times acceleration" at the end points.

Let $T(x, t) = \|\mathbf{T}(x, t)\|_2$ be the Euclidean length of the vector tension $\mathbf{T}(x, t)$ at position x and time t . So $T(x, t)$ is the magnitude of the tension at position x and time t .



Since there is no horizontal motion, we get that the difference between the horizontal component of the force at the end points is null; namely, $T(x_2, t) \cos(\theta_2) - T(x_1, t) \cos(\theta_1) = 0$. Since $\cos(\theta_1) = \frac{1}{\sqrt{1 + u_x^2(x_1, t)}}$ and $\cos(\theta_2) = \frac{1}{\sqrt{1 + u_x^2(x_2, t)}}$, we get

$$\frac{T(x_1, t)}{\sqrt{1 + u_x^2(x_1, t)}} = \frac{T(x_2, t)}{\sqrt{1 + u_x^2(x_2, t)}}. \quad (1)$$

For the vertical motion, the difference between the vertical component of the force at the end points is equal to the mass of the section of the string times its acceleration; namely, $T(x_2, t) \sin(\theta_2) - T(x_1, t) \sin(\theta_1) = \int_{x_1}^{x_2} \rho u_{tt}(x, t) dx$. Since $\sin(\theta_1) = \frac{u_x(x_1, t)}{\sqrt{1 + u_x^2(x_1, t)}}$ and $\sin(\theta_2) = \frac{u_x(x_2, t)}{\sqrt{1 + u_x^2(x_2, t)}}$, we get

$$\frac{T(x_2, t) u_x(x_2, t)}{\sqrt{1 + u_x^2(x_2, t)}} - \frac{T(x_1, t) u_x(x_1, t)}{\sqrt{1 + u_x^2(x_1, t)}} = \int_{x_1}^{x_2} \rho u_{tt}(x, t) dx. \quad (2)$$

Since we assume that the vertical motion is really small, we may assume that $u_x(x, t) \approx 0$ for all x and t . Hence, (1) yields $T(x_1, t) \approx T(x_2, t)$. Since we assume that this is true for all x_1, x_2 and t , we may conclude that the magnitude of the tension is constant along the string; namely, $T(x, t) \equiv T$, a constant, for all x and t .

With all our assumption, we may rewrite (2) as

$$T(u_x(x_2, t) - u_x(x_1, t)) = \rho \int_{x_1}^{x_2} u_{tt}(x, t) dx. \quad (3)$$

If we divide both sides of (3) by $x_2 - x_1$ and let $x_2 \rightarrow x_1$, we get from the definition of the derivative that $T u_{xx}(x_1, t) = \rho u_{tt}(x_1, t)$. Since x_1 is arbitrary, we get the famous wave equation

$$u_{tt} = c^2 u_{xx} \quad (4)$$

for $0 < x < L$ and $t > 0$, where $c = \sqrt{T/\rho}$. This partial differential equation is normally accompanied by initial and boundary conditions. For instance, the initial position of the string at time $t = 0$ may be given by $u(x, 0) = f(x)$ for $0 \leq x \leq L$. Moreover, the ends of the string may be fixed. So $u(0, t) = u(L, t) = 0$ for $t \geq 0$. There are many other choices of boundary conditions. Moreover, the partial differential equation in (4) can be modified to include damping. We will address some of these issues in this book.

A similar reasoning can be used to deduce the wave equation for the vibrating membrane of a drum; namely,

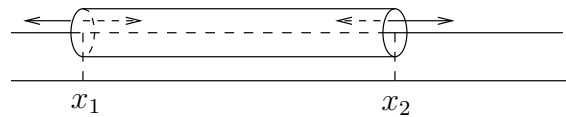
$$u_{tt} = c^2(u_{xx} + u_{yy})$$

for $(x, y) \in D$ and $t > 0$, where $c = \sqrt{T/\rho}$ and D is the region representing the membrane in $\mathbb{R}^2$.

Heat Equation

To study the diffusion of heat in a thin rod, it is easier and equivalent to consider the diffusion of dye in the water contained in a thin and long tube. According to Fick's law, the diffusion of the dye is from a region of higher density to a region of lower density, and the rate of diffusion is proportional to the concentration difference. For the heat equation, the diffusion is from the region of higher temperature to the region of lower temperature, and the rate of diffusion is proportional to the difference of temperature. We should point out that for the rod, we assume that diffusion is only possible at the ends of the rod, there is not diffusion along the rod.

Since the tube is really thin, we consider only the linear dimension of the tube of length L . We consider a small section of a tube of length L .



Let $u(x, t)$ be the density of dye at the position x and time t , the mass of dye in the little section is given by $M(x_1, x_2, t) = \int_{x_1}^{x_2} u(x, t) dx$. The variation of the mass of dye in the little section is due to dye flowing in or out at the ends of the section. Therefore,

$$\int_{x_1}^{x_2} u_t(x, t) dx = \frac{d}{dt} \int_{x_1}^{x_2} u(x, t) dx = k(u_x(x_2, t) - u_x(x_1, t)) ,$$

where k is a constant. If we divide both sides of this equation by $x_2 - x_1$ and let $x_2 \rightarrow x_1$, we get from the definition of the derivative that $u_t(x_1, t) = ku_{xx}(x_1, t)$. Since x_1 is arbitrary, we get the heat equation

$$u_t = ku_{xx} \tag{5}$$

for $0 < x < L$ and $t > 0$. It is called the heat equation because $u(x, t)$ traditionally represent the temperature in a thin and long rod at position x and time t . But, as we have just seen

and as it is typical in mathematics, the same partial differential equation can be used to study different phenomena. The partial differential equation in (5) is usually accompanied of initial and boundary conditions. The initial temperature in the rod may be given by $u(x, 0) = f(x)$ for $0 \leq x \leq L$. There may be boundary conditions at the end of the rod. For instance, the temperature at both ends of the rod may be fixed; so $u(0, t) = g_1(t)$ and $u(L, t) = g_2(t)$ for $t \geq 0$, where $g_1(t)$ is the temperature outside the left side of the rod at time t and $g_2(t)$ is the temperature outside the right side of the rod at time t . Or, there may be no diffusion at the end of the rod; so $u_x(0, t) = u_x(L, t) = 0$ for $t \geq 0$. Another possibility is that the rate of diffusion is proportional to the difference of temperature in the rod and outside the rod according to Newton's law of cooling. Thus, $u_x(0, t) = k(u(0, t) - g_1(t))$ and $u_x(L, t) = -k(u(L, t) - g_2(t))$ for $t \geq 0$, where $k > 0$ is a constant of proportionality.

A similar reasoning can be used to deduce the heat equation in a thin plate; namely,

$$u_t = k(u_{xx} + u_{yy})$$

for $(x, y) \in D$ and $t > 0$, where k is the constant of proportionality and D is the region representing the thin plate in $\mathbb{R}^2$.

Laplace Equation

The Laplace equation comes from the steady state case of the wave equation and the heat equation; namely, we assume that the function u does not depend on time. We get

$$u_{xx} + u_{yy} = 0$$

for $(x, y) \in D$, where D is an open subset of $\mathbb{R}^2$. The solutions of this equation are called “harmonic functions.” We will study these functions intensively in this book. The Laplace equation is usually accompanied of boundary conditions. For instance, $u = 0$ on the boundary ∂D of the region D .

Other Equations

Partial differential equations are ubiquitous in physics. Several examples are mentioned in [38]. To name a few of them, we have the Schrödinger equation that describe the motion of an electron around a proton, the Maxwell's equations in electromagnetism, the Navier-Stokes equation in fluid dynamics, and the Klein-Gordon equation in particle physics. For an excellent introduction to fluid mechanics, we strongly recommend [4]. As the previous list suggest, the first impetus for the study of partial differential equations came from physics. However, partial differential equations are beginning to play an important role in other scientific fields. For an introduction to the role play by partial differential equation (and integro-differential equations) in mathematical biology, we recommend [13]. For a nice introduction to the applications of partial differential equations in several area of science, economics, and so on, we recommend [32].

Suggested Exercises

We could have created exercises very similar to those found in the literature on partial differential equations. Moreover, some exercises are classical problems that are found in mostly all good textbooks on partial differential equations and thus should be included. So, instead of (re)producing some exercises on partial differential equation, we have chosen to refer the reader to the exercises that are found in some of the good textbooks on partial differential equations. Our choice of books is arbitrary and many other good textbooks could have been selected.

We have selected exercises that the reader can solve after having read the related chapter in the book and without having to read the book used as the source for the exercises. We have tried to avoid exercises that refer specifically to the content of the textbook used as the source for the exercises. There are a few exceptions where we have selected exercises that cover important material that has not been included in the present book. As for any book on partial differential equations, we could not cover every topics on the subject. We had to make a personal choice.

Some Terminology

The **k-norm** of $\mathbf{x} \in \mathbb{R}^n$ is defined by

$$\|\mathbf{x}\|_k = \begin{cases} \left(\sum_{i=1}^n |x_i|^k \right)^{1/k} & \text{if } k > 0 \\ \max_{1 \leq i \leq n} |x_i| & \text{if } k = \infty \end{cases}$$

We adopt the convention that $\|\cdot\|$ will denote the **2-norm** or **Euclidean norm** $\|\cdot\|_2$.

The open ball of radius r centred at $\mathbf{y} \in \mathbb{R}^n$ is denoted $B_r(\mathbf{y})$; namely, $B_r(\mathbf{y}) = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x} - \mathbf{y}\| < r\}$. We will also use this notation to denote open balls in any metric spaces.

Let V be a subset of $\mathbb{R}^n$ and $f : V \rightarrow \mathbb{R}$. The **supremum norm** of f on V is defined by $\|f\|_{\infty, V} = \sup_{\mathbf{x} \in V} |f(\mathbf{x})|$. We will simply write $\|f\|_{\infty}$ when it is clear from the context what is the set V used in the definition of the supremum.

If V be an open subset of $\mathbb{R}^n$ and $f : V \rightarrow \mathbb{R}$ is sufficiently differentiable, we define the differential operator D^{α} for $\alpha \in \mathbb{N}^n$ by

$$D^{\alpha} f(\mathbf{x}) = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \dots \partial x_n^{\alpha_n}}(\mathbf{x})$$

for $\mathbf{x} \in V$, where $|\alpha| = \sum_{j=1}^n \alpha_j$.

Definition

Let V be an open subset of $\mathbb{R}^n$. For $m \in \mathbb{N}$, $C^m(\overline{V})$ is the space of all $f \in C^m(V) \cap C(\overline{V})$ such that $D^\alpha f$ can be continuously extended to $\overline{V}$ for all multi-index $\alpha \in \mathbb{N}^n$ such that $|\alpha| \leq m$.

$C^\infty(\overline{V})$ is the space of all $f \in C^\infty(V) \cap C(\overline{V})$ such that $D^\alpha f$ can be continuously extended to $\overline{V}$ for all multi-index $\alpha \in \mathbb{N}^n$.

Definition

Let V be an open subset of $\mathbb{R}^n$ and k be a positive integer. The space of **bounded continuous functions**, denoted $C_b^k(\overline{V})$, is the set of all functions $f \in C^k(\overline{V})$ such that

$$\|f\|_{k,\infty,V} \equiv \max_{|\alpha| \leq k} \|D^\alpha f\|_{\infty,\overline{V}} < \infty .$$

It is not hard to prove that the space $C_b^k(\overline{V})$ is a Banach space with respect to the norm $\|f\|_{k,\infty,V}$.

Definition

Let V be an open subset of $\mathbb{R}^n$ and k be a positive integer. The space of **bounded uniformly continuous functions**, denoted $C_{bu}^k(\overline{V})$, is the set of all functions $f \in C_b^k(\overline{V})$ such that $D^\alpha f : \overline{V} \rightarrow \mathbb{R}$ is uniformly continuous for $|\alpha| \leq k$.

It is also not hard to prove that $C_{bu}^k(\overline{V})$ is a closed subset of $C_b^k(\overline{V})$ with respect to the norm $\|\cdot\|_{k,\infty,V}$.

Definition 0.0.1

If V is an open subset of $\mathbb{R}^n$ and $f : V \rightarrow \mathbb{R}$ is a sufficiently differentiable function, then the **gradient** of u is defined by $\nabla f(\mathbf{x}) = \left(\frac{\partial f}{\partial x_1}(\mathbf{x}), \frac{\partial f}{\partial x_2}(\mathbf{x}), \dots, \frac{\partial f}{\partial x_n}(\mathbf{x}) \right)$ and the

Laplacian of f is defined by $\Delta f(\mathbf{x}) = \frac{\partial^2 f}{\partial x_1^2}(\mathbf{x}) + \frac{\partial^2 f}{\partial x_2^2}(\mathbf{x}) + \dots + \frac{\partial^2 f}{\partial x_n^2}(\mathbf{x})$ for all $\mathbf{x} \in V$.

If $F : V \rightarrow \mathbb{R}^n$ is a sufficiently differentiable function, then the **divergence** of F is defined by $\operatorname{div} F(\mathbf{x}) = \nabla \cdot F(\mathbf{x}) = \frac{\partial F_1}{\partial x_1}(\mathbf{x}) + \frac{\partial F_2}{\partial x_2}(\mathbf{x}) + \dots + \frac{\partial F_n}{\partial x_n}(\mathbf{x})$ for all $\mathbf{x} \in V$.

We have that $\Delta f(\mathbf{x}) = \operatorname{div}(\nabla f(\mathbf{x}))$ if $f : V \rightarrow \mathbb{R}$ is sufficiently differentiable.

Warning: In this book, $\mathbb{N} = \{0, 1, 2, 3, \dots\}$.

Chapter 1

Method of Characteristics

1.1 Quasi-Linear First Order Partial Differential Equations

Consider the **partial differential equation**

$$a(x, y, u) \frac{\partial u}{\partial x} + b(x, y, u) \frac{\partial u}{\partial y} = c(x, y, u) , \quad (1.1.1)$$

where $a : \mathbb{R}^3 \rightarrow \mathbb{R}$ and $b : \mathbb{R}^3 \rightarrow \mathbb{R}$ are two given differentiable functions, and $u : \mathbb{R}^2 \rightarrow \mathbb{R}$ is an unknown differentiable function.

Equations of the form (1.1.1) are called **quasi-linear partial differential equations**. If a and b are independent of u , and $c(x, y, u) = p(x, y)u + q(x, y)$ for some differentiable functions $p : \mathbb{R}^2 \rightarrow \mathbb{R}$ and $q : \mathbb{R}^2 \rightarrow \mathbb{R}$, then the partial differential equation is called **linear**.

The goal is to find a differentiable solution, also called surface, $u = u(x, y)$ that satisfy (1.1.1) and contains a given curve Γ defined by

$$\Gamma = \{(x, y, u) = (x_0(s), y_0(s), u_0(s)) : s \in I\}$$

for some differentiable functions x_0 , y_0 and u_0 defined on an open interval I . The problems where the solutions (if any) of the partial differential equation must contain a given curve are called **Cauchy problems**.

Definition 1.1.1

A surface $u = u(x, y)$ satisfying (1.1.1) and containing a given curve Γ is called an **integral surface** for (1.1.1).

We may rewrite (1.1.1) as $(a, b, c) \cdot \left(\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}, -1 \right) = 0$. The vector $\left(\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}, -1 \right)$ evaluated at (x, y) is perpendicular to the surface $u = u(x, y)$ at $(x, y, u(x, y))$. Hence, the vector (a, b, c) evaluated at $(x, y, u(x, y))$ is in the tangent plan to the integral surface $u = u(x, y)$ at the point $(x, y, u(x, y))$.

If $\gamma = \{(x(t), y(t), u(t)) : t \in \mathbb{R}\}$ is a curve on the integral surface $u = u(x, y)$ then $(x'(t), y'(t), u'(t))$ is in the tangent plane (more precisely, parallel to the tangent plane) to the integral surface at the point $(x(t), y(t), u(t))$.

It follows from the previous discussion that we should consider the following family of curves. For each fixed value of $s \in I$, let γ_s be the curve defined by

$$\gamma_s = \{(x, y, u) = (x(t, s), y(t, s), u(t, s)) : t \in \mathbb{R}\} ,$$

where

$$\begin{aligned} \frac{\partial x}{\partial t}(t, s) &= a(x(t, s), y(t, s), u(t, s)) \\ \frac{\partial y}{\partial t}(t, s) &= b(x(t, s), y(t, s), u(t, s)) \\ \frac{\partial u}{\partial t}(t, s) &= c(x(t, s), y(t, s), u(t, s)) \end{aligned} \quad (1.1.2)$$

and

$$(x(0, s), y(0, s), u(0, s)) = (x_0(s), y_0(s), u_0(s)) . \quad (1.1.3)$$

For s fix, (1.1.2) is a system of **ordinary differential equations** depending on a parameter s and satisfying the initial condition (1.1.3). For this reason, in the literature, the partial derivatives are often replaced by the derivatives $\frac{dx}{dt}$, $\frac{dy}{dt}$ and $\frac{du}{dt}$.

Definition 1.1.2

The equations in (1.1.2) are called **characteristic equations** and, for each s , the curve γ_s is called a **characteristic curve**.

We end up with a family of curves $\{\gamma_s\}_{s \in I}$. If we assume that

$$\frac{\partial(x, y)}{\partial(t, s)} \bigg|_{(t, s)} = \det \begin{pmatrix} \frac{\partial x}{\partial t} & \frac{\partial x}{\partial s} \\ \frac{\partial y}{\partial t} & \frac{\partial y}{\partial s} \end{pmatrix} \bigg|_{(t, s)} \neq 0 , \quad (1.1.4)$$

then it follows from the Inverse Mapping Theorem that s and t may be expressed (locally) in function of x and y . Substituting these expressions of s and t in terms of x and y into $u = u(t, s)$ yields the equation of a surface $u = u(x, y)$ (Figure 1.1).

To simplify the notation, we use u to denote both the function of t and s , and the function of x and y . The context will determine which one is used. We have $u(x, y) = u(t, s)$ where $x = x(t, s)$ and $y = y(t, s)$.

Theorem 1.1.3

Let Γ be a curve defined by

$$\Gamma = \{(x, y, u) = (x_0(s), y_0(s), u_0(s)) : s \in I\} ,$$

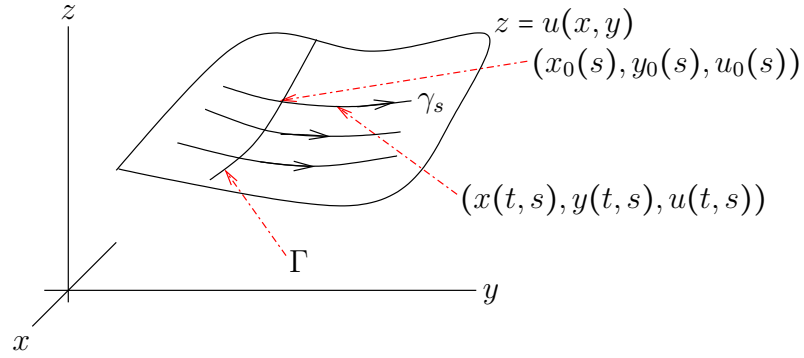


Figure 1.1: Schematic representation of the characteristic curves and the integral surface.

where I is an open interval and $x_0, y_0, u_0 \in C_b^1(I)$. For $s \in I$, let

$$\gamma_s = \{(x(t, s), y(t, s), u(t, s)) : t \in J_s\}$$

be the **orbit** of the solution $t \mapsto (x(t, s), y(t, s), u(t, s))$ of (1.1.2) with the initial condition (1.1.3), where $J_s \subset \mathbb{R}$ is the domain of the solution. If

$$\det \begin{pmatrix} a(x(t, s), y(t, s), u(t, s)) & \frac{\partial x}{\partial s}(t, s) \\ b(x(t, s), y(t, s), u(t, s)) & \frac{\partial y}{\partial s}(t, s) \end{pmatrix} \neq 0$$

for all (t, s) ¹, then the surface S given by the family of curves $\{\gamma_s\}_{s \in I}$ is the integral surface of (1.1.1) containing the curve Γ . There is only one integral surface containing Γ .

Proof.

i) Given $(\tilde{t}, \tilde{s}) \in J_s \times I$, it follows from the theorems of existence and uniqueness of solutions for ordinary differential equations that there exist two open intervals $I_b \subset I$ and $J_b \subset J_s$ such that $(\tilde{t}, \tilde{s}) \in J_b \times I_b$ and the solutions x and y are defined and continuously differentiable on $J_b \times I_b$. Since $J_b \times I_b$ is simply connected², we have that the surface

$$S_b = \{(x(t, s), y(t, s), u(t, s)) : (t, s) \in J_b \times I_b\}$$

is simply connected. The integral surface S is the union of the surfaces S_b . The surface S is covered by at most a countable collections of surfaces S_b .

¹From now on, we will often state “for all t and s ” to mean “for all t and s where $x(t, s)$, $y(t, s)$ and $u(t, s)$ are defined.”

²There is no “holes” in the surface. The reader may want to consult the chapter on homotopy theory in a good textbook on topology.

ii) As we mentioned before, the existence of a differentiable surface $u = u(x, y)$ comes from the Inverse Mapping Theorem because

$$\left. \frac{\partial(x, y)}{\partial(t, s)} \right|_{(\tilde{t}, \tilde{s})} = \det \left(\begin{array}{cc} \frac{\partial x}{\partial t} & \frac{\partial x}{\partial s} \\ \frac{\partial y}{\partial t} & \frac{\partial y}{\partial s} \end{array} \right) \bigg|_{(\tilde{t}, \tilde{s})} = \det \left(\begin{array}{cc} a(x(\tilde{t}, \tilde{s}), y(\tilde{t}, \tilde{s}), u(\tilde{t}, \tilde{s})) & \frac{\partial x}{\partial s}(\tilde{t}, \tilde{s}) \\ b(x(\tilde{t}, \tilde{s}), y(\tilde{t}, \tilde{s}), u(\tilde{t}, \tilde{s})) & \frac{\partial y}{\partial s}(\tilde{t}, \tilde{s}) \end{array} \right) \neq 0$$

by hypothesis. We can write s and t in a neighbourhood of $(\tilde{t}, \tilde{s})$ as functions of x and y . Substituting these expressions of s and t in $u = u(t, s)$ gives the equation $u = u(x, y)$ to describe the surface S .

iii) This surface $u = u(x, y)$ contains Γ by construction. In fact, $\Gamma = \{(x(0, s), y(0, s), u(0, s)) : s \in I\}$. Moreover, the points on the surface S are of the form

$$(x, y, u) = (x(t, s), y(t, s), u(t, s)) = (x(t, s), y(t, s), u(x(t, s), y(t, s)))$$

for some value of s and t . Hence,

$$\begin{aligned} c(x, y, u) &= c(x(t, s), y(t, s), u(t, s)) = \frac{\partial}{\partial t} u(t, s) = \frac{\partial}{\partial t} u(x(t, s), y(t, s)) \\ &= \frac{\partial u}{\partial x}(x(t, s), y(t, s)) \frac{\partial x}{\partial t}(t, s) + \frac{\partial u}{\partial y}(x(t, s), y(t, s)) \frac{\partial y}{\partial t}(t, s) \\ &= a(x(t, s), y(t, s), u(t, s)) \frac{\partial u}{\partial x}(x(t, s), y(t, s)) \\ &\quad + b(x(t, s), y(t, s), u(t, s)) \frac{\partial u}{\partial y}(x(t, s), y(t, s)) \\ &= a(x, y, u) \frac{\partial u}{\partial x}(x, y) + b(x, y, u) \frac{\partial u}{\partial y}(x, y) . \end{aligned}$$

Thus (1.1.1) is satisfied.

iv) To prove the uniqueness of the integral surface containing Γ , we show that a characteristic curve that starts on an integral surface $u = v(x, y)$ of (1.1.1) containing the curve Γ stays on the integral surface. By uniqueness of characteristic curves, it will follow that the integral surface $u = v(x, y)$ is the integral surface $u = u(x, y)$ that we have found above.

Let

$$g(t, s) = u(t, s) - v(x(t, s), y(t, s)) ,$$

where

$$\gamma_s = \{(x, y, u) = (x(t, s), y(t, s), u(t, s)) : t \in \mathbb{R}\}$$

is a characteristic curve, where we assume that s is fix. We have that (1.1.2) and (1.1.3) are satisfied by x , y and u . Hence,

$$\begin{aligned} \frac{\partial g}{\partial t}(t, s) &= \frac{\partial u}{\partial t}(t, s) - \frac{\partial v}{\partial x}(x(t, s), y(t, s)) \frac{\partial x}{\partial t}(t, s) - \frac{\partial v}{\partial y}(x(t, s), y(t, s)) \frac{\partial y}{\partial t}(t, s) \\ &= c(x(t, s), y(t, s), u(t, s)) - a(x(t, s), y(t, s), u(t, s)) \frac{\partial v}{\partial x}(x(t, s), y(t, s)) \end{aligned}$$

$$\begin{aligned}
& -b(x(t, s), y(t, s), u(t, s)) \frac{\partial v}{\partial y}(x(t, s), y(t, s)) \\
& = c(x(t, s), y(t, s), g(t, s) + v(x(t, s), y(t, s))) \\
& \quad - a(x(t, s), y(t, s), g(t, s) + v(x(t, s), y(t, s))) \frac{\partial v}{\partial x}(x(t, s), y(t, s)) \\
& \quad - b(x(t, s), y(t, s), g(t, s) + v(x(t, s), y(t, s))) \frac{\partial v}{\partial y}(x(t, s), y(t, s)) . \tag{1.1.5}
\end{aligned}$$

This is a first order ordinary differential equation for g with the initial condition

$$g(0, s) = u(0, s) - v(x(0, s), y(0, s)) = u_0(s) - v(x_0(s), y_0(s)) = 0$$

because $u = v(x, y)$ is a surface containing the curve Γ . By uniqueness of solutions for ordinary differential equations and the fact that v satisfies (1.1.1), we have that $g(t, s) = 0$ for all $t \in J_s$ because it is a solution of (1.1.5). Thus $v(x(t, s), y(t, s)) = u(t, s) = u(x(t, s), y(t, s))$ for all t and s since s is arbitrary; namely, $v = u$. ■

Remark 1.1.4

It is important to remember that, for each characteristic curve

$\gamma_s = \{(x(t, s), y(t, s), u(t, s)) : t \in J_s\}$, the vector $\left(\frac{\partial x}{\partial t}(t, s), \frac{\partial y}{\partial t}(t, s), \frac{\partial u}{\partial t}(t, s)\right)$ is in the tangent plane of the integral surface $u = u(x, y)$ (more precisely, parallel to the tangent plane) at $(x(t, s), y(t, s), u(t, s))$. ♠

Remark 1.1.5

If (1.1.4) is not satisfied at $(t, s) = (0, \tilde{s})$, then

$$\begin{aligned}
0 &= \frac{\partial(x, y)}{\partial(t, s)} \Big|_{(0, \tilde{s})} = \det \begin{pmatrix} \frac{\partial x}{\partial t} & \frac{\partial x}{\partial s} \\ \frac{\partial y}{\partial t} & \frac{\partial y}{\partial s} \end{pmatrix} \Big|_{(0, \tilde{s})} = \det \begin{pmatrix} a(x(0, \tilde{s}), y(0, \tilde{s}), u(0, \tilde{s})) & \frac{\partial x}{\partial s}(0, \tilde{s}) \\ b(x(0, \tilde{s}), y(0, \tilde{s}), u(0, \tilde{s})) & \frac{\partial y}{\partial s}(0, \tilde{s}) \end{pmatrix} \\
&= \det \begin{pmatrix} a(x_0(\tilde{s}), y_0(\tilde{s}), u_0(\tilde{s})) & x'_0(\tilde{s}) \\ b(x_0(\tilde{s}), y_0(\tilde{s}), u_0(\tilde{s})) & y'_0(\tilde{s}) \end{pmatrix} \\
&= a(x_0(\tilde{s}), y_0(\tilde{s}), u_0(\tilde{s}))y'_0(\tilde{s}) - b(x_0(\tilde{s}), y_0(\tilde{s}), u_0(\tilde{s}))x'_0(\tilde{s}) . \tag{1.1.6}
\end{aligned}$$

If Γ is on the integral surface $u = u(x, y)$, then $u_0(s) = u(x_0(s), y_0(s))$ and thus

$$u'_0(s) = \frac{\partial u}{\partial x}(x_0(s), y_0(s))x'_0(s) + \frac{\partial u}{\partial y}(x_0(s), y_0(s))y'_0(s) . \tag{1.1.7}$$

Moreover, we have from (1.1.1) that

$$a(x, y, u) \frac{\partial u}{\partial x}(x, y) + b(x, y, u) \frac{\partial u}{\partial y}(x, y) = c(x, y, u) \tag{1.1.8}$$

at $(x, y, u) = (x_0(s), y_0(s), u_0(s))$ for $s \in I$.

Hence, from (1.1.7) and (1.1.8), we have

$$\begin{aligned}
& b(x_0(\tilde{s}), y_0(\tilde{s}), u_0(\tilde{s}))u'_0(\tilde{s}) - c(x_0(\tilde{s}), y_0(\tilde{s}), u_0(\tilde{s}))y'_0(\tilde{s}) \\
&= b(x_0(\tilde{s}), y_0(\tilde{s}), u_0(\tilde{s})) \left(\frac{\partial u}{\partial x}(x_0(\tilde{s}), y_0(\tilde{s}))x'_0(\tilde{s}) + \frac{\partial u}{\partial y}(x_0(\tilde{s}), y_0(\tilde{s}))y'_0(\tilde{s}) \right) \\
&\quad - \left(a(x, y, u) \frac{\partial u}{\partial x}(x, y) + b(x, y, u) \frac{\partial u}{\partial y}(x, y) \right) \Big|_{(x,y,u)=(x_0(\tilde{s}), y_0(\tilde{s}), u_0(\tilde{s}))} y'_0(\tilde{s}) \\
&= \left(b(x_0(\tilde{s}), y_0(\tilde{s}), u_0(\tilde{s}))x'_0(\tilde{s}) - a(x_0(\tilde{s}), y_0(\tilde{s}), u_0(\tilde{s}))y'_0(\tilde{s}) \right) \frac{\partial u}{\partial x}(x_0(\tilde{s}), y_0(\tilde{s})) \\
&= 0, \tag{1.1.9}
\end{aligned}$$

where the last identity comes from (1.1.6). Similarly,

$$\begin{aligned}
& a(x_0(\tilde{s}), y_0(\tilde{s}), u_0(\tilde{s}))u'_0(\tilde{s}) - c(x_0(\tilde{s}), y_0(\tilde{s}), u_0(\tilde{s}))x'_0(\tilde{s}) \\
&= a(x_0(\tilde{s}), y_0(\tilde{s}), u_0(\tilde{s})) \left(\frac{\partial u}{\partial x}(x_0(\tilde{s}), y_0(\tilde{s}))x'_0(\tilde{s}) + \frac{\partial u}{\partial y}(x_0(\tilde{s}), y_0(\tilde{s}))y'_0(\tilde{s}) \right) \\
&\quad - \left(a(x, y, u) \frac{\partial u}{\partial x}(x, y) + b(x, y, u) \frac{\partial u}{\partial y}(x, y) \right) \Big|_{(x,y,u)=(x_0(\tilde{s}), y_0(\tilde{s}), u_0(\tilde{s}))} x'_0(\tilde{s}) \\
&= \left(a(x_0(\tilde{s}), y_0(\tilde{s}), u_0(\tilde{s}))y'_0(\tilde{s}) - b(x_0(\tilde{s}), y_0(\tilde{s}), u_0(\tilde{s}))x'_0(\tilde{s}) \right) \frac{\partial u}{\partial y}(x_0(\tilde{s}), y_0(\tilde{s})) \\
&= 0. \tag{1.1.10}
\end{aligned}$$

Hence, from (1.1.6), (1.1.9) and (1.1.10), we deduce that

$$(a(x, y, u), b(x, y, u), c(x, y, u)) \Big|_{(x,y,u)=(x_0(\tilde{s}), y_0(\tilde{s}), u_0(\tilde{s}))} \times (x'_0(\tilde{s}), y'_0(\tilde{s}), u'_0(\tilde{s})) = 0$$

and thus these two vectors are parallel.

If (1.1.4) is not satisfied at $(t, s) = (0, \tilde{s})$ and Γ is a characteristic curve, then there may be infinitely integral surfaces, one for each initial curve Γ^* intersecting Γ transversely that we chose. However, if (1.1.4) is not satisfied at $(t, s) = (0, \tilde{s})$ and Γ is not a characteristic, then there is no integral surface. ♠

Example 1.1.6

Solve the partial differential equation

$$\frac{\partial u}{\partial x} + \frac{\partial u}{\partial y} = 2$$

with the condition $u(x, 0) = x^2$.

The characteristic equations are $x'(t) = 1$, $y'(t) = 1$ and $u'(t) = 2$. The initial curve is given by $x(0) = s$, $y(0) = 0$ and $u(0) = s^2$.

The equation $x'(t) = 1$ with $x(0) = s$ gives $x = x(t, s) = t + s$, the equation $y'(t) = 1$ with $y(0) = 0$ gives $y = y(t, s) = t$, and the equation $u'(t) = 2$ with $u(0) = s^2$ gives $u = u(t, s) =$

$2t + s^2$. We can solve $x = t + s$ and $y = t$ in terms of s and t to get $t = y$ and $s = x - y$. If we substitute these values of s and t in the expression for u , we get the integral surface $u = u(x, y) = 2y + (x - y)^2$. ♣

Example 1.1.7

Solve the partial differential equation

$$x \frac{\partial u}{\partial x} + \frac{\partial u}{\partial y} = u$$

with the condition $u(x, 0) = \ln(x)$ for $x > 0$.

The characteristic equations are $x'(t) = x(t)$, $y'(t) = 1$ and $u'(t) = u(t)$. The initial curve is given by $x(0) = s$, $y(0) = 0$ and $u(0) = \ln(s)$.

The equation $x'(t) = x(t)$ with $x(0) = s$ gives $x = x(t, s) = se^t$, the equation $y'(t) = 1$ with $y(0) = 0$ gives $y = y(t, s) = t$, and the equation $u'(t) = u(t)$ with $u(0) = \ln(s)$ gives $u = u(t, s) = \ln(s)e^t$. We can solve $x = se^t$ and $y = t$ in terms of s and t to get $t = y$ and $s = xe^{-y}$. If we substitute these values of s and t in the expression for u , we get the integral surface

$$u = u(x, y) = \ln(xe^{-y})e^y = \ln(x)e^y - ye^y.$$

♣

Example 1.1.8

Solve the partial differential equation

$$\frac{\partial u}{\partial x} + \frac{\partial u}{\partial y} = u$$

with the condition $u(x, 0) = \cos(x)$.

The characteristic equations are $x'(t) = 1$, $y'(t) = 1$ and $u'(t) = u(t)$. The initial curve is given by $x(0) = s$, $y(0) = 0$ and $u(0) = \cos(s)$.

The equation $x'(t) = 1$ with $x(0) = s$ gives $x = x(t, s) = t + s$, the equation $y'(t) = 1$ with $y(0) = 0$ gives $y = y(t, s) = t$, and the equation $u'(t) = u(t)$ with $u(0) = \cos(s)$ gives $u = u(t, s) = \cos(s)e^t$. We can solve $x = t + s$ and $y = t$ in terms of s and t to get $t = y$ and $s = x - y$. If we substitute these values of s and t in the expression for u , we get the integral surface $u = u(x, y) = \cos(x - y)e^y$. ♣

Example 1.1.9

Solve the partial differential equation

$$x^2 \frac{\partial u}{\partial x} + y^2 \frac{\partial u}{\partial y} = u^2$$

with the condition $u(x, 2x) = 1$.

The characteristic equations are $x'(t) = x^2(t)$, $y'(t) = y^2(t)$ and $u'(t) = u^2(t)$. The initial curve is given by $x(0) = s$, $y(0) = 2s$ and $u(0) = 1$.

The equation $x'(t) = x^2(t)$ with $x(0) = s$ gives

$$x = x(t, s) = \frac{s}{1 - st} . \quad (1.1.11)$$

Similarly, the equation $y'(t) = y^2(t)$ with $y(0) = 2s$ gives

$$y = y(t, s) = \frac{2s}{1 - 2st} . \quad (1.1.12)$$

Finally, the equation $u'(t) = u^2(t)$ with $u(0) = 1$ gives $u = u(t, s) = \frac{1}{1-t}$. We can solve (1.1.11) and (1.1.12) for s and t in terms of x and y to get $t = \frac{1}{x} - \frac{2}{y}$ and $s = \frac{xy}{2y - 2x}$. If we substitute these values of s and t in the expression for u , we get the integral surface

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

♣

1.1.1 Lagrange Method

The Lagrange method produce the **general solution** (namely, a family of integral surfaces that are not required to contain the initial curve Γ) of a quasi-Linear first order ordinary differential equation like (1.1.1). We assume that the characteristic curves are given by the intersection of two families of surfaces of the form

$$\phi(x, y, u) = \alpha \quad \text{and} \quad \psi(x, y, u) = \beta .$$

The characteristic curve associated to the initial condition $(x_0(s), y_0(s), u_0(s)) \in \Gamma$ is the intersection of the surfaces $\phi(x, y, u) = \alpha_s$ and $\psi(x, y, u) = \beta_s$. for some pair (α_s, β_s) as illustrated in Figure 1.2.

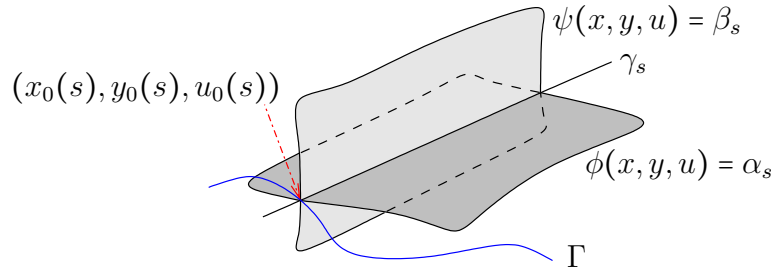


Figure 1.2: The intersection of the two surfaces $\phi(x, y, u) = \alpha_s$ and $\psi(x, y, u) = \beta_s$ is the characteristic curve γ_s corresponding to the initial condition $(x_0(s), y_0(s), u_0(s))$. We have $F(\alpha_s, \beta_s) = 0$.

The set $\{(\alpha_s, \beta_s) : s \in I\}$ is a curve in the α, β plane. If this curve is given by $F(\alpha, \beta) = 0$, the integral surface of the quasi-linear equation containing Γ is given by

$$F(\phi(x, y, u), \psi(x, y, u)) = 0.$$

The problem is to find ϕ and ψ . In some rare cases, we can find two linearly independent vector valued functions $v_1 : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ and $v_2 : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ such that $v_1(x, y, u)$ and $v_2(x, y, u)$ are orthogonal to $(a(x, y, u), b(x, y, u), c(x, y, u))$ for all (x, y, u) , and $\text{curl}(v_1) = \text{curl}(v_2) = 0$. Since v_1 and v_2 are curl free, there exist differentiable functions $\phi : \mathbb{R}^3 \rightarrow \mathbb{R}$ and $\psi : \mathbb{R}^3 \rightarrow \mathbb{R}$ such that $\nabla \phi = v_1$ and $\nabla \psi = v_2$. We have that ϕ and ψ are constant along the characteristic curves because

$$\begin{aligned} \frac{\partial}{\partial t} \phi(x(t, s), y(t, s), u(t, s)) &= \nabla \phi(x(t, s), y(t, s), u(t, s)) \cdot \left(\frac{\partial x}{\partial t}(t, s), \frac{\partial y}{\partial t}(t, s), \frac{\partial u}{\partial t}(t, s) \right) \\ &= v_1(x(t, s), y(t, s), u(t, s)) \cdot (a(x, y, u), b(x, y, u), c(x, y, u)) \Big|_{(x, y, u) = (x(t, s), y(t, s), u(t, s))} = 0 \end{aligned}$$

and, similarly,

$$\frac{\partial}{\partial t} \psi(x(t, s), y(t, s), u(t, s)) = 0$$

for all t and s .

Example 1.1.10

Find the integral surface of the partial differential equation

$$-y \frac{\partial u}{\partial x} + x \frac{\partial u}{\partial y} = 0$$

with the condition $u(x, x) = e^x$.

We have

$$(a(x, y, u), b(x, y, u), c(x, y, u)) = (-y, x, 0).$$

The vectors $v_1(x, y, u) = (x, y, 0)$ and $v_2(x, y, u) = (0, 0, 1)$ are two linearly independent vector valued functions satisfying $\text{curl}(v_1) = \text{curl}(v_2) = 0$ and orthogonal to $(a(x, y, u), b(x, y, u), c(x, y, u))$ for all (x, y, u) . We have that $\phi(x, y, u) = (x^2 + y^2)/2$ satisfies $\nabla \phi = v_1$ and $\psi(x, y, u) = u$ satisfies $\nabla \psi = v_2$. The integral surfaces of the partial differential equation are of the form

$$F(x^2 + y^2, u) = 0 \tag{1.1.13}$$

for functions $F : \mathbb{R}^2 \rightarrow \mathbb{R}$.

Suppose that $\frac{\partial F}{\partial \beta}(\alpha, \beta) \neq 0$, then we can solve $F(\alpha, \beta) = 0$ for β as a function of α . Namely, there exists a function $f : \mathbb{R} \rightarrow \mathbb{R}$ such that $F(\alpha, f(\alpha)) = 0$. It then follows from (1.1.13) that $u(x, y) = f(x^2 + y^2)$. From the initial condition $u(x, x) = e^x$, we get $e^x = f(2x^2)$. With $z = 2x^2$, we get $f(z) = e^{\sqrt{z/2}}$ for $z \geq 0$. We finally get the integral surface $u(x, y) = e^{\sqrt{(x^2 + y^2)/2}}$.

♣

1.2 Envelope of a Family of Surfaces

Our analysis of the general first order partial differential equation in the next sections will use the concept of the envelope of a family of surfaces. This is the subject of this section.

We consider the family of surfaces described by

$$F(x, y, z, \lambda) = z - G(x, y, \lambda) = 0 \quad , \quad \lambda \in \mathbb{R} . \quad (1.2.1)$$

Namely, the collection of surfaces S_λ for $\lambda \in \mathbb{R}$ given by

$$S_\lambda = \{(x, y, z) : F(x, y, z, \lambda) = 0\} .$$

The differential equation

$$\frac{\partial G}{\partial \lambda}(x, y, \lambda) = 0 \quad (1.2.2)$$

combined with (1.2.1) yields a curve γ_λ for each fixed value of λ . The **envelope** of the family of surfaces is the union of these curves. The word “envelope” refers generally to a surface that wrap up something. Is our definition of the envelope of a family of surfaces a real surface in $\mathbb{R}^3$?

Suppose that $\frac{\partial G}{\partial \lambda}(x_1, y_1, \lambda_1) = 0$ and $\frac{\partial^2 G}{\partial \lambda^2}(x_1, y_1, \lambda_1) \neq 0$, then we may use the Implicit Function Theorem to find open neighbourhoods $V \subset \mathbb{R}^2$ of (x_1, y_1) and $W \subset \mathbb{R}^3$ of (x_1, y_1, λ_1) , and a unique continuously differentiable function $\lambda : V \rightarrow \mathbb{R}$ such that $\lambda(x_1, y_1) = \lambda_1$, $(x, y, \lambda(x, y)) \in W$ for all $(x, y) \in V$ and

$$\frac{\partial G}{\partial \lambda}(x, y, \lambda(x, y)) = 0 \quad (1.2.3)$$

for $(x, y) \in V$. Moreover, all solutions of (1.2.2) in W are given by this function λ .

For λ_2 closed to λ_1 , the curve γ_{λ_2} is defined by

$$\begin{aligned} \gamma_{\lambda_2} &= \{(x, y, z) : (x, y) \in V, \lambda(x, y) = \lambda_2 \text{ and } F(x, y, z, \lambda_2) = 0\} \\ &= \{(x, y, G(x, y, \lambda(x, y))) : (x, y) \in V \text{ and } \lambda(x, y) = \lambda_2\} . \end{aligned}$$

The union of these curves for λ_2 close to λ_1 form a “patch” of the envelope E for the family of surfaces in (1.2.1). If we substitute $\lambda = \lambda(x, y)$ in (1.2.1), we get the equation

$$F(x, y, z, \lambda(x, y)) = z - G(x, y, \lambda(x, y)) = 0$$

of a local “patch” of the envelope E .

Note that γ_{λ_2} is the intersection of the surface S_{λ_2} with the envelope E of the family of surface.

Remark 1.2.1

We have from (1.2.3) that

$$\frac{\partial}{\partial x} G(x, y, \lambda(x, y)) = \frac{\partial G}{\partial x}(x, y, \lambda(x, y)) + \frac{\partial G}{\partial \lambda}(x, y, \lambda(x, y)) \frac{\partial \lambda}{\partial x}(x, y) = \frac{\partial G}{\partial x}(x, y, \lambda(x, y))$$

and

$$\frac{\partial}{\partial y}G(x, y, \lambda(x, y)) = \frac{\partial G}{\partial y}(x, y, \lambda(x, y)) + \frac{\partial G}{\partial \lambda}(x, y, \lambda(x, y)) \frac{\partial \lambda}{\partial y}(x, y) = \frac{\partial G}{\partial y}(x, y, \lambda(x, y)) .$$

Hence, if we let $H(x, y, z) = F(x, y, z, \lambda(x, y))$, a normal to the envelope E at a point $(\tilde{x}, \tilde{y}, \tilde{z} = G(\tilde{x}, \tilde{y}, \lambda(\tilde{x}, \tilde{y})))$ of E is given by

$$\begin{aligned} \nabla H(\tilde{x}, \tilde{y}, \tilde{z}) &= \left(\frac{\partial H}{\partial x}(\tilde{x}, \tilde{y}, \tilde{z}), \frac{\partial H}{\partial y}(\tilde{x}, \tilde{y}, \tilde{z}), \frac{\partial H}{\partial z}(\tilde{x}, \tilde{y}, \tilde{z}) \right) \\ &= \left(-\frac{\partial}{\partial x}G(\tilde{x}, \tilde{y}, \lambda(\tilde{x}, \tilde{y})), -\frac{\partial}{\partial y}G(\tilde{x}, \tilde{y}, \lambda(\tilde{x}, \tilde{y})), 1 \right) \\ &= \left(-\frac{\partial G}{\partial x}(\tilde{x}, \tilde{y}, \lambda(\tilde{x}, \tilde{y})), -\frac{\partial G}{\partial y}(\tilde{x}, \tilde{y}, \lambda(\tilde{x}, \tilde{y})), 1 \right) . \end{aligned}$$

This is also a normal vector to the surface described by $F(x, y, z, \tilde{\lambda}) = z - G(x, y, \tilde{\lambda}) = 0$ at the point $(\tilde{x}, \tilde{y}, \tilde{z})$ of this surface where $\lambda(\tilde{x}, \tilde{y}) = \tilde{\lambda}$; namely, a point $(\tilde{x}, \tilde{y}, \tilde{z}) \in \gamma_{\tilde{\lambda}}$. ♠

Example 1.2.2

Find the envelope of the family of spheres

$$x^2 + (y - \lambda)^2 + z^2 = 4 .$$

If we assume that $z \geq 0$, we can write the equation of this family of spheres as

$$F(x, y, x, \lambda) = z - \sqrt{4 - x^2 - (y - \lambda)^2} = 0 . \quad (1.2.4)$$

Thus $G(x, y, \lambda) = \sqrt{4 - x^2 - (y - \lambda)^2}$ in (1.2.1). To find the envelope, we need to consider (1.2.2) for this function G ; namely,

$$\frac{\partial G}{\partial \lambda}(x, y, \lambda) = \frac{(y - \lambda)}{\sqrt{4 - x^2 - (y - \lambda)^2}} = 0 .$$

Thus $\lambda = y$. If we substitute $\lambda = y$ in (1.2.4), we get $z = \sqrt{4 - x^2}$. Similarly, for $z < 0$, we get $z = -\sqrt{4 - x^2}$. Therefore, the envelope is the surface $x^2 + z^2 = 4$. It is the cylinder of radius 2 whose axis is the y -axis. ♣

1.3 General First Order Partial Differential Equations

We consider partial differential equation of the form

$$F\left(x, y, u, \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}\right) = 0 , \quad (1.3.1)$$

where $F : \mathbb{R}^5 \rightarrow \mathbb{R}$ is sufficiently differentiable.

The goal is to find a differentiable surface $u = u(x, y)$ that satisfies (1.3.1) and contains a given curve Γ defined by

$$\{(x, y, u) = (\tilde{x}(s), y_0(s), u_0(s)) : s \in I\} ,$$

where I is an open interval, and x_0, y_0, u_0 are differentiable functions. This is the **Cauchy problem** for the general first-order partial differential equation. As for the quasi-linear partial differential equation, we have the following definition.

Definition 1.3.1

A surface $u = u(x, y)$ satisfying (1.3.1) and containing a given curve Γ is called an **integral surface** for (1.3.1).

As for the quasi-linear partial differential equation, we will find a family of curves $\{\gamma_s\}_{s \in I}$ that generates an integral surface $u = u(x, y)$ for (1.3.1) and where each curve γ_s intersects transversely Γ at $(x_0(s), y_0(s), u_0(s))$. Unfortunately, the ordinary differential equation that determine this family of curves cannot be read directly from the partial differential equation as it was done for the quasi-linear partial differential equation.

We did not have to explicitly consider the tangent plane to the integral surface for the quasi-linear partial differential equation but we have to do so for the general partial differential equation. Suppose that $u = u(x, y)$ is a surface satisfying (1.3.1) and let

$$(\tilde{x}, \tilde{y}, \tilde{u}) = (\tilde{x}, \tilde{y}, u(\tilde{x}, \tilde{y}))$$

be a point on this surface. The equation of the tangent plane to the integral surface at $(\tilde{x}, \tilde{y}, \tilde{u})$ is

$$\frac{\partial u}{\partial x}(\tilde{x}, \tilde{y})(x - \tilde{x}) + \frac{\partial u}{\partial y}(\tilde{x}, \tilde{y})(y - \tilde{y}) - (u - \tilde{u}) = 0 . \quad (1.3.2)$$

(1.3.1) gives a relation between $\frac{\partial u}{\partial x}(\tilde{x}, \tilde{y})$ and $\frac{\partial u}{\partial y}(\tilde{x}, \tilde{y})$; namely,

$$F\left(\tilde{x}, \tilde{y}, \tilde{u}, \frac{\partial u}{\partial x}(\tilde{x}, \tilde{y}), \frac{\partial u}{\partial y}(\tilde{x}, \tilde{y})\right) = 0 .$$

Combined with the derivative along the curve Γ , we will be able to determine the tangent plane to the integral surface.

1.3.1 Equations for x, y and u

The comments of the previous paragraph can be used to deduce conditions that must be satisfied by the curves γ_s with $s \in I$. The equation $F(\tilde{x}, \tilde{y}, \tilde{u}, p, q) = 0$ represents a curve in the p, q plane. We get a one-parameter family of tangent planes defined by

$$p(x - \tilde{x}) + q(y - \tilde{y}) - (u - \tilde{u}) = 0 \quad \text{and} \quad F(\tilde{x}, \tilde{y}, \tilde{u}, p, q) = 0 .$$

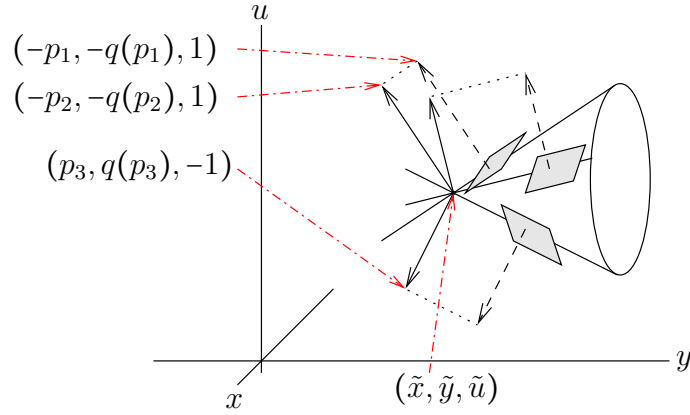


Figure 1.3: Assuming that we can write the solution of the equation $F(\tilde{x}, \tilde{y}, \tilde{u}, p, q) = 0$ as $q = q(p)$, the figure represent some of the tangent planes containing the point $(\tilde{x}, \tilde{y}, \tilde{u})$ and enveloping the Monge Cone.

This family of tangent planes, all containing the point $(\tilde{x}, \tilde{y}, \tilde{u})$, generates a “cone” with its summit at $(\tilde{x}, \tilde{y}, \tilde{u})$. This cone is called a **Monge Cone**. To visualize this cone, think of the region that is enveloped by this one-parameter family of tangent planes all containing the point $(\tilde{x}, \tilde{y}, \tilde{u})$. This is illustrate in Figure 1.3.

Let

$$S = \{(p, q) : F(\tilde{x}, \tilde{y}, \tilde{u}, p, q) = 0\} .$$

We assume that F in (1.3.1) is not degenerate on S ; namely

$$\left(\frac{\partial F}{\partial p}(\tilde{x}, \tilde{y}, \tilde{u}, p, q), \frac{\partial F}{\partial q}(\tilde{x}, \tilde{y}, \tilde{u}, p, q) \right) \neq (0, 0)$$

for all $(p, q) \in S$. We may assume without loss of generality that

$$\frac{\partial F}{\partial q}(\tilde{x}, \tilde{y}, \tilde{u}, p, q) \neq 0$$

locally along the curve S . We can then write $q = q(p)$ locally.

If we derive with respect to p the equation

$$p(x - \tilde{x}) + q(y - \tilde{y}) = u - \tilde{u} \tag{1.3.3}$$

of the tangent plane at $(\tilde{x}, \tilde{y}, \tilde{u})$, we get

$$(x - \tilde{x}) + \frac{dq}{dp}(y - \tilde{y}) = 0 .$$

By the Implicit Function Theorem, since we assume that $\frac{\partial F}{\partial q}(\tilde{x}, \tilde{y}, \tilde{u}, p, q) \neq 0$, we have

$$\frac{dq}{dp} = - \frac{\frac{\partial F}{\partial p}(\tilde{x}, \tilde{y}, \tilde{u}, p, q)}{\frac{\partial F}{\partial q}(\tilde{x}, \tilde{y}, \tilde{u}, p, q)} .$$

Hence,

$$\frac{(x - \tilde{x})}{\frac{\partial F}{\partial p}(\tilde{x}, \tilde{y}, \tilde{u}, p, q)} = \frac{(y - \tilde{y})}{\frac{\partial F}{\partial q}(\tilde{x}, \tilde{y}, \tilde{u}, p, q)} . \quad (1.3.4)$$

Moreover, if we divide (1.3.3) by $(y - \tilde{y})$, we get

$$\frac{u - \tilde{u}}{y - \tilde{y}} = p \frac{x - \tilde{x}}{y - \tilde{y}} + q = p \frac{\frac{\partial F}{\partial p}(\tilde{x}, \tilde{y}, \tilde{u}, p, q)}{\frac{\partial F}{\partial q}(\tilde{x}, \tilde{y}, \tilde{u}, p, q)} + q .$$

Thus

$$\left(\frac{u - \tilde{u}}{y - \tilde{y}} \right) \frac{\partial F}{\partial q}(\tilde{x}, \tilde{y}, \tilde{u}, p, q) = p \frac{\partial F}{\partial p}(\tilde{x}, \tilde{y}, \tilde{u}, p, q) + q \frac{\partial F}{\partial q}(\tilde{x}, \tilde{y}, \tilde{u}, p, q) .$$

This finally gives

$$\frac{u - \tilde{u}}{p \frac{\partial F}{\partial p}(\tilde{x}, \tilde{y}, \tilde{u}, p, q) + q \frac{\partial F}{\partial q}(\tilde{x}, \tilde{y}, \tilde{u}, p, q)} = \frac{y - \tilde{y}}{\frac{\partial F}{\partial q}(\tilde{x}, \tilde{y}, \tilde{u}, p, q)} . \quad (1.3.5)$$

Equations (1.3.4) and (1.3.5) yield the standard representation of the line defined by (1.3.3). This line is in the tangent plane of the integral surface $u = u(x, y)$ at $(\tilde{x}, \tilde{y}, \tilde{u})$ if

$$(p, q) = \left(\frac{\partial u}{\partial x}(\tilde{x}, \tilde{y}), \frac{\partial u}{\partial y}(\tilde{x}, \tilde{y}) \right) .$$

The direction of this line at $(\tilde{x}, \tilde{y}, \tilde{u})$ is

$$\left(\frac{\partial F}{\partial p}(\tilde{x}, \tilde{y}, \tilde{u}, p, q), \frac{\partial F}{\partial q}(\tilde{x}, \tilde{y}, \tilde{u}, p, q), p \frac{\partial F}{\partial p}(\tilde{x}, \tilde{y}, \tilde{u}, p, q) + q \frac{\partial F}{\partial q}(\tilde{x}, \tilde{y}, \tilde{u}, p, q) \right) .$$

As for the quasi-linear first order partial differential equation, this suggests that the integral surface is the union of the family of curves $\{\gamma_s\}_{s \in I}$, where each curve γ_s is defined by

$$\gamma_s = \{(x, y, u) = (x(t, s), y(t, s), u(t, s)) : t \in J_s\}$$

such that

$$\begin{aligned}
\frac{\partial x}{\partial t}(t, s) &= \frac{\partial F}{\partial p}(x(t, s), y(t, s), u(t, s), p(t, s), q(t, s)) , \\
\frac{\partial y}{\partial t}(t, s) &= \frac{\partial F}{\partial q}(x(t, s), y(t, s), u(t, s), p(t, s), q(t, s)) , \\
\frac{\partial u}{\partial t}(t, s) &= p(t, s) \frac{\partial F}{\partial p}(x(t, s), y(t, s), u(t, s), p(t, s), q(t, s)) \\
&\quad + q(t, s) \frac{\partial F}{\partial q}(x(t, s), y(t, s), u(t, s), p(t, s), q(t, s))
\end{aligned} \tag{1.3.6}$$

and

$$(x(0, s), y(0, s), u(0, s)) = (x_0(s), y_0(s), u_0(s)) ,$$

where p and q are unknown functions to be determined later and J_s is the domain of the solution for s given. To determine p and q , we need two additional differential equations related to the tangent planes of the integral surface sought.

If we assume for now that we have also determined p and q , and that

$$\left. \frac{\partial(x, y)}{\partial(t, s)} \right|_{(t, s)} = \det \begin{pmatrix} \frac{\partial x}{\partial t} & \frac{\partial x}{\partial s} \\ \frac{\partial y}{\partial t} & \frac{\partial y}{\partial s} \end{pmatrix} \bigg|_{(t, s)} \neq 0 , \tag{1.3.7}$$

then we can use the Inverse Mapping Theorem to express locally s and t as functions of x and y . Substituting these expressions of s and t into $u = u(s, t)$, $p = p(t, s)$ and $q = q(s, t)$ yields $u = u(x, y)$, $p = p(x, y)$ and $q = q(x, y)$. As we did for the quasi-linear first order partial differential equation, we keep u , p and q to denote the functions of s and t , as well as the functions of x and y . We have $u(x, y) = u(t, s)$ for $x = x(s, t)$ and $y = y(s, t)$, and similar equalities for p and q .

1.3.2 Equations for p and q

The functions p and q in (1.3.6) must satisfy

$$p(t, s) = \frac{\partial u}{\partial x}(x(t, s), y(t, s)) \quad \text{and} \quad q(t, s) = \frac{\partial u}{\partial y}(x(t, s), y(t, s)) \tag{1.3.8}$$

if we want to satisfy the consistency conditions

$$p(x, y) = \frac{\partial u}{\partial x}(x, y) \quad \text{and} \quad q(x, y) = \frac{\partial u}{\partial y}(x, y) .$$

If we assume that (1.3.8) is satisfied and derive the equations in (1.3.8) with respect to t , we get

$$\frac{\partial p}{\partial t}(t, s) = \frac{\partial^2 u}{\partial x^2}(x(t, s), y(t, s)) \frac{\partial x}{\partial t}(t, s) + \frac{\partial^2 u}{\partial y \partial x}(x(t, s), y(t, s)) \frac{\partial y}{\partial t}(t, s)$$

$$\begin{aligned}
&= \frac{\partial^2 u}{\partial x^2}(x(t, s), y(t, s)) \frac{\partial F}{\partial p}(x(t, s), y(t, s), u(t, s), p(t, s), q(t, s)) \\
&\quad + \frac{\partial^2 u}{\partial y \partial x}(x(t, s), y(t, s)) \frac{\partial F}{\partial q}(x(t, s), y(t, s), u(t, s), p(t, s), q(t, s)) \quad (1.3.9)
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial q}{\partial t}(t, s) &= \frac{\partial^2 u}{\partial x \partial y}(x(t, s), y(t, s)) \frac{\partial x}{\partial t}(t, s) + \frac{\partial^2 u}{\partial y^2}(x(t, s), y(t, s)) \frac{\partial y}{\partial t}(t, s) \\
&= \frac{\partial^2 u}{\partial x \partial y}(x(t, s), y(t, s)) \frac{\partial F}{\partial p}(x(t, s), y(t, s), u(t, s), p(t, s), q(t, s)) \\
&\quad + \frac{\partial^2 u}{\partial y^2}(x(t, s), y(t, s)) \frac{\partial F}{\partial q}(x(t, s), y(t, s), u(t, s), p(t, s), q(t, s)) \quad (1.3.10)
\end{aligned}$$

Moreover, if u is an integral surface, we have $F(x, y, u(x, y), p(x, y), q(x, y)) = 0$ with $p(x, y) = \frac{\partial u}{\partial x}(x, y)$ and $q(x, y) = \frac{\partial u}{\partial y}(x, y)$. If we derive $F(x, y, u(x, y), p(x, y), q(x, y)) = 0$ with respect to x , we get

$$\frac{\partial F}{\partial x} + \frac{\partial F}{\partial u} \frac{\partial u}{\partial x} + \frac{\partial F}{\partial p} \frac{\partial p}{\partial x} + \frac{\partial F}{\partial q} \frac{\partial q}{\partial x} = 0$$

at $(x, y, u(x, y), p(x, y), q(x, y))$. Hence,

$$-\frac{\partial F}{\partial x} - \frac{\partial F}{\partial u} \frac{\partial u}{\partial x} = \frac{\partial F}{\partial p} \frac{\partial p}{\partial x} + \frac{\partial F}{\partial q} \frac{\partial q}{\partial x} = \frac{\partial^2 u}{\partial x^2} \frac{\partial F}{\partial p} + \frac{\partial^2 u}{\partial y \partial x} \frac{\partial F}{\partial q} \quad (1.3.11)$$

at $(x, y, u(x, y), p(x, y), q(x, y))$. Similarly, if we derive $F(x, y, u(x, y), p(x, y), q(x, y)) = 0$ with respect to y , we get

$$\frac{\partial F}{\partial y} + \frac{\partial F}{\partial u} \frac{\partial u}{\partial y} + \frac{\partial F}{\partial p} \frac{\partial p}{\partial y} + \frac{\partial F}{\partial q} \frac{\partial q}{\partial y} = 0$$

at $(x, y, u(x, y), p(x, y), q(x, y))$. Hence,

$$-\frac{\partial F}{\partial y} - \frac{\partial F}{\partial u} \frac{\partial u}{\partial y} = \frac{\partial F}{\partial p} \frac{\partial p}{\partial y} + \frac{\partial F}{\partial q} \frac{\partial q}{\partial y} = \frac{\partial^2 u}{\partial x \partial y} \frac{\partial F}{\partial p} + \frac{\partial^2 u}{\partial y^2} \frac{\partial F}{\partial q} \quad (1.3.12)$$

at $(x, y, u(x, y), p(x, y), q(x, y))$. If we substitute (1.3.11) and (1.3.12) into (1.3.9) and (1.3.10) respectively, and use the relations $p(x, y) = p(t, s)$, $q(x, y) = q(t, s)$ and $u(x, y) = u(t, s)$ for $(x, t) = (x(t, s), y(t, s))$, we get

$$\begin{aligned}
\frac{\partial p}{\partial t}(t, s) &= -\frac{\partial F}{\partial x}(x(t, s), y(t, s), u(t, s), p(t, s), q(t, s)) \\
&\quad - p(t, s) \frac{\partial F}{\partial u}(x(t, s), y(t, s), u(t, s), p(t, s), q(t, s)) \quad , \\
\frac{\partial q}{\partial t}(t, s) &= -\frac{\partial F}{\partial y}(x(t, s), y(t, s), u(t, s), p(t, s), q(t, s)) \\
&\quad - q(t, s) \frac{\partial F}{\partial u}(x(t, s), y(t, s), u(t, s), p(t, s), q(t, s)) \quad .
\end{aligned} \quad (1.3.13)$$

(1.3.6) and (1.3.13) define a system of five ordinary differential equations for the five unknown functions $x(t, s)$, $y(t, s)$, $u(t, s)$, $p(t, s)$ and $q(t, s)$.

1.3.3 initial conditions

The only thing missing for the system of differential equations (1.3.6) and (1.3.13) is a set of initial conditions. Since Γ must be on the integral surface, we require $x(0, s) = x_0(s)$, $y(0, s) = y_0(s)$ and $u(0, s) = u_0(s)$ for $s \in I$. Moreover, if (1.3.8) is true, then

$$p_0(s) = p(0, s) = p(x(0, s), y(0, s)) = \frac{\partial u}{\partial x}(x_0(s), y_0(s))$$

and

$$q_0(s) = q(0, s) = q(x(0, s), y(0, s)) = \frac{\partial u}{\partial y}(x_0(s), y_0(s))$$

must satisfy

$$F(x_0(s), y_0(s), u_0(s), p_0(s), q_0(s)) = 0 \quad (1.3.14)$$

since

$$F\left(x_0(s), y_0(s), u(x_0(s), y_0(s)), \frac{\partial u}{\partial x}(x_0(s), y_0(s)), \frac{\partial u}{\partial y}(x_0(s), y_0(s))\right) = 0$$

for all s .

Moreover,

$$\frac{du_0}{ds}(s) = \frac{d}{ds}u(x_0(s), y_0(s)) = \frac{\partial u}{\partial x}(x_0(s), y_0(s)) \frac{dx_0}{ds}(s) + \frac{\partial u}{\partial y}(x_0(s), y_0(s)) \frac{dy_0}{ds}(s)$$

yields

$$\frac{du_0}{ds}(s) = p_0(s) \frac{dx_0}{ds}(s) + q_0(s) \frac{dy_0}{ds}(s) . \quad (1.3.15)$$

The initial conditions for p and q at $t = 0$ should therefore be $p_0(s)$ and $q_0(s)$, the solution of (1.3.14) and (1.3.15).

Definition 1.3.2

The equations in (1.3.6) and (1.3.13) are called the **characteristic equations** for the general partial differential equation in (1.3.1). A solution of the characteristic equations is called a **characteristic strip**. The **characteristic curve** are the projection in the x, y, u space of the solutions of the characteristic equations with the initial condition

$$(x(0, s), y(0, s), u(0, s), p(0, s), q(0, s)) = (x_0(s), y_0(s), u_0(s), p_0(s), q_0(s)) ,$$

where $p_0(s)$ and $q_0(s)$ are given by

$$F(x_0(s), y_0(s), u_0(s), p_0(s), q_0(s)) = 0 \quad (1.3.16)$$

and

$$\frac{du_0}{ds}(s) = p_0(s) \frac{dx_0}{ds}(s) + q_0(s) \frac{dy_0}{ds}(s) \quad (1.3.17)$$

for $s \in I$.

Remark 1.3.3

The system of equations (1.3.16) and (1.3.17) may have no solution, a unique solution or many solutions. Namely, there may be no choice, only one choice or many choices possible for $p_0(s)$ and $q_0(s)$. Therefore, there may be zero, one or many integral surfaces containing the initial curve Γ . ♠

Theorem 1.3.4

Let Γ be a curve defined by

$$\{(x, y, u) = (x_0(s), y_0(s), u_0(s)) : s \in I\} ,$$

where I is an open interval and $x_0, y_0, u_0 \in C_b^1(I)$. Let $p_0(s)$ and $q_0(s)$ be the solutions, if any, of

$$F(x_0(s), y_0(s), u_0(s), p_0(s), q_0(s)) = 0 \quad (1.3.18)$$

and

$$\frac{du_0}{ds}(s) = p_0(s) \frac{dx_0}{ds}(s) + q_0(s) \frac{dy_0}{ds}(s) \quad (1.3.19)$$

for $s \in I$. (see Remark 1.3.6 below.) For $s \in I$ fixed, let

$$\gamma_s = \{(x(t, s), y(t, s), u(t, s)) : t \in J_s\}$$

be the projection in the x, y, u space of the orbit of the solution

$$(x, y, u, p, q) = (x(t, s), y(t, s), u(t, s), p(t, s), q(t, s))$$

of (1.3.6) and (1.3.13) with the initial condition

$$(x(0, s), y(0, s), u(0, s), p(0, s), q(0, s)) = (x_0(s), y_0(s), u_0(s), p_0(s), q_0(s)) , \quad (1.3.20)$$

where $J_s \subset \mathbb{R}$ is the domain of the solution. If

$$\det \begin{pmatrix} \frac{\partial F}{\partial p}(x(t, s), y(t, s), u(t, s), p(t, s), q(t, s)) & \frac{\partial x}{\partial s}(t, s) \\ \frac{\partial F}{\partial q}(x(t, s), y(t, s), u(t, s), p(t, s), q(t, s)) & \frac{dy}{ds}(t, s) \end{pmatrix} \neq 0 \quad (1.3.21)$$

for all (t, s) , then the surface S given by the union of the family of curves $\{\gamma_s\}_{s \in I}$ is the integral surface for (1.3.1) containing the curve Γ . There is only one integral surface containing Γ .

Proof.

i) The proof that S is a nice surface is identical to the proof given in (i) of the proof of Proposition 1.1.3

ii) The existence of a surface $u = u(x, y)$ comes from the Inverse Mapping Theorem because

$$\begin{aligned} \left. \frac{\partial(x, y)}{\partial(t, s)} \right|_{(\tilde{t}, \tilde{s})} &= \det \begin{pmatrix} \frac{\partial x}{\partial t} & \frac{\partial x}{\partial s} \\ \frac{\partial y}{\partial t} & \frac{\partial y}{\partial s} \end{pmatrix} \Big|_{(\tilde{t}, \tilde{s})} \\ &= \det \begin{pmatrix} \frac{\partial F}{\partial p}(x(\tilde{t}, \tilde{s}), y(\tilde{t}, \tilde{s}), u(\tilde{t}, \tilde{s}), p(\tilde{t}, \tilde{s}), q(\tilde{t}, \tilde{s})) & \frac{\partial x}{\partial s}(\tilde{t}, \tilde{s}) \\ \frac{\partial F}{\partial q}(x(\tilde{t}, \tilde{s}), y(\tilde{t}, \tilde{s}), u(\tilde{t}, \tilde{s}), p(\tilde{t}, \tilde{s}), q(\tilde{t}, \tilde{s})) & \frac{\partial y}{\partial s}(\tilde{t}, \tilde{s}) \end{pmatrix} \neq 0 \end{aligned}$$

by hypothesis. We can write s and t in a neighbourhood of $(\tilde{t}, \tilde{s})$ as functions of x and y . Substituting these expressions of s and t in $u = u(t, s)$ gives the equation $u = u(x, y)$ to describe the surface S . We can also substitute these expressions of s and t in $p = p(t, s)$ and $q = q(t, s)$ to get $p = p(x, y)$ and $q = q(x, y)$.

iii) We claim that

$$F(x(t, s), y(t, s), u(t, s), p(t, s), q(t, s)) = 0 \quad (1.3.22)$$

for all t and s . We have from (1.3.6) and (1.3.13) that

$$\begin{aligned} \frac{\partial}{\partial t} F(x, y, u, p, q) &= \frac{\partial F}{\partial x}(x, y, u, p, q) \frac{\partial x}{\partial t} + \frac{\partial F}{\partial y}(x, y, u, p, q) \frac{\partial y}{\partial t} + \frac{\partial F}{\partial u}(x, y, u, p, q) \frac{\partial u}{\partial t} \\ &\quad + \frac{\partial F}{\partial p}(x, y, u, p, q) \frac{\partial p}{\partial t} + \frac{\partial F}{\partial q}(x, y, u, p, q) \frac{\partial q}{\partial t} \\ &= \frac{\partial F}{\partial x}(x, y, u, p, q) \frac{\partial F}{\partial p}(x, y, u, p, q) + \frac{\partial F}{\partial y}(x, y, u, p, q) \frac{\partial F}{\partial q}(x, y, u, p, q) \\ &\quad + \frac{\partial F}{\partial u}(x, y, u, p, q) \left(p \frac{\partial F}{\partial p}(x, y, u, p, q) + q \frac{\partial F}{\partial q}(x, y, u, p, q) \right) \\ &\quad + \frac{\partial F}{\partial p}(x, y, u, p, q) \left(-\frac{\partial F}{\partial x}(x, y, u, p, q) - p \frac{\partial F}{\partial u}(x, y, u, p, q) \right) \\ &\quad + \frac{\partial F}{\partial q}(x, y, u, p, q) \left(-\frac{\partial F}{\partial y}(x, y, u, p, q) - q \frac{\partial F}{\partial u}(x, y, u, p, q) \right) = 0 \end{aligned}$$

for all (t, s) . Thus $F(x(t, s), y(t, s), u(t, s), p(t, s), q(t, s))$ is constant with respect to t whatever $s \in I$. Since

$$F(x(0, s), y(0, s), u(0, s), p(0, s), q(0, s)) = F(x_0(s), y_0(s), u_0(s), p_0(s), q_0(s)) = 0$$

for $s \in I$ by hypothesis (1.3.18), we have that $F(x(t, s), y(t, s), u(t, s), p(t, s), q(t, s)) = 0$ for all $t \in J_s$ and $s \in I$.

iv) We need to prove that our new surface $u = u(x, y)$ satisfies

$$p(x, y) = \frac{\partial u}{\partial x}(x, y) \quad \text{and} \quad q(x, y) = \frac{\partial u}{\partial y}(x, y) .$$

To do this, it is enough to show that

$$\frac{\partial u}{\partial t}(t, s) = p(t, s) \frac{\partial x}{\partial t}(t, s) + q(t, s) \frac{\partial y}{\partial t}(t, s) \quad (1.3.23)$$

and

$$\frac{\partial u}{\partial s}(t, s) = p(t, s) \frac{\partial x}{\partial s}(t, s) + q(t, s) \frac{\partial y}{\partial s}(t, s) , \quad (1.3.24)$$

because the system

$$\begin{aligned} \frac{\partial u}{\partial t}(t, s) &= A(t, s) \frac{\partial x}{\partial t}(t, s) + B(t, s) \frac{\partial y}{\partial t}(t, s) \\ \frac{\partial u}{\partial s}(t, s) &= A(t, s) \frac{\partial x}{\partial s}(t, s) + B(t, s) \frac{\partial y}{\partial s}(t, s) \end{aligned}$$

as the solution

$$\begin{aligned} A(t, s) &= \frac{\partial u}{\partial x}(x(t, s), y(t, s)) \\ B(t, s) &= \frac{\partial u}{\partial y}(x(t, s), y(t, s)) \end{aligned}$$

according the Chain Rule, and the solution is (locally) unique since $\left. \frac{\partial(x, y)}{\partial(t, s)} \right|_{(t, s)} \neq 0$.

(1.3.23) is a direct consequence of the three equations in (1.3.6). To prove (1.3.24), let

$$R(t, s) = \frac{\partial u}{\partial s}(t, s) - p(t, s) \frac{\partial x}{\partial s}(t, s) - q(t, s) \frac{\partial y}{\partial s}(t, s) .$$

We have

$$\begin{aligned} \frac{\partial R}{\partial t} &= \frac{\partial^2 u}{\partial t \partial s} - \frac{\partial p}{\partial t} \frac{\partial x}{\partial s} - p \frac{\partial^2 x}{\partial t \partial s} - \frac{\partial q}{\partial t} \frac{\partial y}{\partial s} - q \frac{\partial^2 y}{\partial t \partial s} \\ &= \frac{\partial}{\partial s} \left(\frac{\partial u}{\partial t} - p \frac{\partial x}{\partial t} - q \frac{\partial y}{\partial t} \right) - \frac{\partial p}{\partial t} \frac{\partial x}{\partial s} - \frac{\partial q}{\partial t} \frac{\partial y}{\partial s} + \frac{\partial p}{\partial s} \frac{\partial x}{\partial t} + \frac{\partial q}{\partial s} \frac{\partial y}{\partial t} \\ &= - \frac{\partial p}{\partial t} \frac{\partial x}{\partial s} - \frac{\partial q}{\partial t} \frac{\partial y}{\partial s} + \frac{\partial p}{\partial s} \frac{\partial x}{\partial t} + \frac{\partial q}{\partial s} \frac{\partial y}{\partial t} \end{aligned}$$

for all (t, s) due to (1.3.23). We then get from (1.3.6) and (1.3.13) that

$$\begin{aligned} \frac{\partial R}{\partial t} &= \left(\frac{\partial F}{\partial x}(x, y, u, p, q) + p \frac{\partial F}{\partial u}(x, y, u, p, q) \right) \frac{\partial x}{\partial s} + \left(\frac{\partial F}{\partial y}(x, y, u, p, q) \right. \\ &\quad \left. + q \frac{\partial F}{\partial u}(x, y, u, p, q) \right) \frac{\partial y}{\partial s} + \frac{\partial F}{\partial p}(x, y, u, p, q) \frac{\partial p}{\partial s} + \frac{\partial F}{\partial q}(x, y, u, p, q) \frac{\partial q}{\partial s} \\ &= \frac{\partial F}{\partial u}(x, y, u, p, q) \left(p \frac{\partial x}{\partial s} + q \frac{\partial y}{\partial s} \right) + \frac{\partial F}{\partial x}(x, y, u, p, q) \frac{\partial x}{\partial s} + \frac{\partial F}{\partial y}(x, y, u, p, q) \frac{\partial y}{\partial s} \\ &\quad + \frac{\partial F}{\partial p}(x, y, u, p, q) \frac{\partial p}{\partial s} + \frac{\partial F}{\partial q}(x, y, u, p, q) \frac{\partial q}{\partial s} \\ &= \frac{\partial F}{\partial u}(x, y, u, p, q) \left(p \frac{\partial x}{\partial s} + q \frac{\partial y}{\partial s} - \frac{\partial u}{\partial s} \right) + \frac{\partial F}{\partial x}(x, y, u, p, q) \frac{\partial x}{\partial s} + \frac{\partial F}{\partial y}(x, y, u, p, q) \frac{\partial y}{\partial s} \\ &\quad + \frac{\partial F}{\partial u}(x, y, u, p, q) \frac{\partial u}{\partial s} + \frac{\partial F}{\partial p}(x, y, u, p, q) \frac{\partial p}{\partial s} + \frac{\partial F}{\partial q}(x, y, u, p, q) \frac{\partial q}{\partial s} \end{aligned}$$

$$= -\frac{\partial F}{\partial u}(x, y, u, p, q) R + \frac{\partial}{\partial s} F(x, y, u, p, q)$$

for all (t, s) . It follows from (1.3.22) that

$$\frac{\partial}{\partial s} F(x(t, s), y(t, s), u(t, s), p(t, s), q(t, s)) = 0$$

for all (t, s) . Thus, we get the separable ordinary differential equation

$$\frac{\partial R}{\partial t}(t, s) = -\frac{\partial F}{\partial u}(x(t, s), y(t, s), u(t, s), p(t, s), q(t, s)) R(t, s) .$$

Its solutions is $R(t, s) = Ce^{H(t, s)}$, where C is a constant and

$$H(t, s) = -\int_0^t \frac{\partial F}{\partial u}(x(\tau, s), y(\tau, s), u(\tau, s), p(\tau, s), q(\tau, s)) d\tau .$$

From (1.3.19), we have $R(0, s) = 0$ for $s \in I$. Thus the constant C is null and $R(t, s) = 0$ for all $t \in J_s$ and $s \in I$. This prove (1.3.24).

▼ The surface $u = u(x, y)$ contains Γ by construction. Moreover, (1.3.1) is satisfied by $u = u(x, y)$ because (1.3.22) is true, and $u(t, s) = u(x, y)$, $p(t, s) = p(x, y) = \frac{\partial u}{\partial x}(x, y)$ and $q(t, s) = q(x, y) = \frac{\partial u}{\partial y}(x, y)$ for $x = x(t, s)$ and $y = y(t, s)$. ■

Remark 1.3.5

Given $p_0(s)$ and $q_0(s)$ for $s \in I$, the integral surface containing Γ is unique. ♠

Remark 1.3.6

In the previous proposition, we may only assume that (1.3.18), (1.3.19) and (1.3.21) are satisfied for one value $s = \tilde{s}$. In fact, the existence of $p_0(s)$ and $q_0(s)$ for $s \neq \tilde{s}$ is not even required.

Let

$$\begin{aligned} G(p, q, s) &= \frac{du_0}{ds}(s) - p \frac{dx_0}{ds}(s) - q \frac{dy_0}{ds}(s) \\ H(p, q, s) &= F(x_0(s), y_0(s), u_0(s), p, q) \end{aligned}$$

Since

$$\begin{aligned} & \left. \frac{\partial(G.H)}{\partial(p, q)} \right|_{(p, q, s) = (p_0(\tilde{s}), q_0(\tilde{s}), \tilde{s})} \\ &= \det \begin{pmatrix} -\frac{dx_0}{ds}(s) & -\frac{dy_0}{ds}(s) \\ \frac{\partial F}{\partial p}(x_0(s), y_0(s), u_0(s), p, q) & \frac{\partial F}{\partial q}(x_0(s), y_0(s), u_0(s), p, q) \end{pmatrix} \Bigg|_{(p, q, s) = (p_0(\tilde{s}), q_0(\tilde{s}), \tilde{s})} \\ & \neq 0 \end{aligned}$$

by hypothesis, we may use the Implicit Function Theorem to find $p = p_0(s)$ and $q = q_0(s)$ for s near $\tilde{s}$ such that $G(p_0(s), q_0(s), s) = 0$ and $H(p_0(s), q_0(s), s) = 0$; namely, (1.3.18) and (1.3.19) are true for s near $\tilde{s}$. By continuity, (1.3.21) will therefore be true for s close enough to $\tilde{s}$. By taking a small enough interval I containing $\tilde{s}$, we have the hypothesis of the previous proposition. ♠

Remark 1.3.7

In the previous proposition, not only do we find the characteristic curves but we also find the equation of the tangent plane of the integral surface along the characteristic curves; namely,

$$p(t, s)(x - x(t, s)) + q(t, s)(y - y(t, s)) = (u - u(t, s)) .$$

This is the reason why the solutions of the characteristic equations are called characteristic strips. ♠

Example 1.3.8

Solve the first order non-linear partial differential equation

$$\left(\frac{\partial u}{\partial x}\right)^2 + \frac{\partial u}{\partial y} = 0$$

with the condition $u(x, 0) = x$.

This is a partial differential equation of the form (1.3.1) with $F(x, y, u, p, q) = p^2 + q$. Hence, the characteristic equations are

$$\begin{aligned} \frac{\partial x}{\partial t} &= \frac{\partial F}{\partial p} = 2p \\ \frac{\partial y}{\partial t} &= \frac{\partial F}{\partial q} = 1 \\ \frac{\partial u}{\partial t} &= p \frac{\partial F}{\partial p} + q \frac{\partial F}{\partial q} = 2p^2 + q \\ \frac{\partial p}{\partial t} &= -\frac{\partial F}{\partial x} - p \frac{\partial F}{\partial u} = 0 \\ \frac{\partial q}{\partial t} &= -\frac{\partial F}{\partial y} - q \frac{\partial F}{\partial u} = 0 \end{aligned}$$

with the initial condition $x(0, s) = x_0(s) = s$, $y(0, s) = y_0(s) = 0$ and $u(0, s) = u_0(s) = s$, and where $p_0(s)$ and $q_0(s)$ are the solutions of

$$F(x_0(s), y_0(s), u_0(s), p_0(s), q_0(s)) = p_0^2(s) + q_0(s) = 0$$

and

$$\frac{du_0}{ds}(s) - p_0(s) \frac{dx_0}{ds}(s) - q_0(s) \frac{dy_0}{ds}(s) = 1 - p_0(s) = 0 .$$

We find $p_0(s) = -q_0(s) = 1$ for all s . Starting from the last characteristic equation above, we find $q(t, s) = -1$, $p(t, s) = 1$, $u(t, s) = t + s$, $y(t, s) = t$ and $x(t, s) = 2t + s$. Solving the equations for x and y , we get $t = y$ and $s = x - 2y$. Substituting these expressions of s and t in u , p and

q , we find $u(x, y) = x - y$, $p(x, y) = 1$ and $q(x, y) = -1$. Note that $\frac{\partial u}{\partial x}(x, y) = p(x, y) = 1$ and $\frac{\partial u}{\partial y}(x, y) = q(x, y) = -1$ for all x and y as expected. $\clubsuit$

Example 1.3.9

Consider the wave equation

$$\frac{\partial^2 u}{\partial t^2} = c^2(x, y) \Delta u ,$$

where $c(x, y)$ is the speed of a wave at the point (x, y) in the plane. As we will see, solutions of the wave equations are linear combinations of terms of the form $u_\omega = e^{i\omega t} \psi(x, y)$. If $\omega = kc_0$ and $c(x, y) = c_0/n(x, y)$, where c_0 is the average speed of a wave in the medium, then the wave equation gives

$$\Delta \psi + k^2 n^2(x, y) \psi = 0 . \quad (1.3.25)$$

The function n is called the **refraction index** and the number k is called the **wave number**. The **wavelength** is $1/(2\pi k)$. In particular, in the case of a lens, the wavelength is not more than one micron while the other lengths (e.g. thickness and curvature of the lens) are several millimetres. Rescaling with respect to the largest length in the system, we may assume that k is very large compared to the other lengths in the system.

We seek solutions of (1.3.25) of the form $\psi(x, y) = \rho(x, y, k) e^{ik\phi(x, y)}$. If we substitute this expression of ψ in (1.3.25) and divide the result by $e^{ik\phi(x, y)}$, we get

$$\rho_{xx} + \rho_{yy} + 2ik(\rho_x \phi_x + \rho_y \phi_y) + ik\rho(\phi_{xx} + \phi_{yy}) - k^2\rho(\phi_x^2 + \phi_y^2) + k^2 n^2 \rho = 0 .$$

Thus,

$$\rho(\phi_x^2 + \phi_y^2 - n^2) = \frac{1}{k^2}(\rho_{xx} + \rho_{yy}) + \frac{1}{k}(2i(\rho_x \phi_x + \rho_y \phi_y) + i\rho(\phi_{xx} + \phi_{yy})) = O\left(\frac{1}{k}\right) .$$

Since we assume that k is much larger than the order lengths in the system, we have approximately

$$\phi_x^2 + \phi_y^2 - n^2 = 0 . \quad (1.3.26)$$

This partial differential equation is called the **eikonal equation** and was first proposed by W. R. Hamilton in the nineteenth century. This equation is fundamental in geometrical optics. The surfaces defined by $\phi = a$, where a is a constant, are the **wavefronts**. More information on the eikonal equation can be found in [16, 22, 24].

a) We will solve the eikonal equation for n constant (i.e. when the speed $c(x, y)$ of a wave is independent of the position in the medium) and under the initial conditions $\phi(x, 2x) = 1$.

The eikonal equation is of the form $F(x, y, u, p, q) = p^2 + q^2 - n^2 = 0$ for $u = \phi$. Hence, the

characteristic equations are

$$\begin{aligned}
\frac{\partial x}{\partial t} &= \frac{\partial F}{\partial p}(x, y, u, p, q) = 2p \\
\frac{\partial y}{\partial t} &= \frac{\partial F}{\partial q}(x, y, u, p, q) = 2q \\
\frac{\partial u}{\partial t} &= p \frac{\partial F}{\partial p}(x, y, u, p, q) + q \frac{\partial F}{\partial q}(x, y, u, p, q) = 2(p^2 + q^2) \\
\frac{\partial p}{\partial t} &= -\frac{\partial F}{\partial x}(x, y, u, p, q) - p \frac{\partial F}{\partial u}(x, y, u, p, q) = 0 \\
\frac{\partial q}{\partial t} &= -\frac{\partial F}{\partial y}(x, y, u, p, q) - q \frac{\partial F}{\partial u}(x, y, u, p, q) = 0
\end{aligned} \tag{1.3.27}$$

We have $\Gamma = \{(x_0(s), y_0(s), u_0(s)) : s \in I\}$, where $x_0(s) = s$, $y_0(s) = 2s$ and $u_0(s) = 1$ for $s \in I = \mathbb{R}$. As for $p_0(s)$ and $q_0(s)$, they satisfy

$$F(x_0(s), y_0(s), u_0(s), p_0(s), q_0(s)) = p_0^2(s) + q_0^2(s) - n^2 = 0$$

and

$$0 = \frac{du_0}{ds}(s) = p_0(s) \frac{dx_0}{ds}(s) + q_0(s) \frac{dy_0}{ds}(s) = p_0(s) + 2q_0(s)$$

for $s \in \mathbb{R}$. We get $p_0(s) = -2q_0(s)$ and $q_0(s) = \pm n/\sqrt{5}$.

We have respectively that

$$\begin{aligned}
\frac{\partial p}{\partial t}(t, s) &= 0 \quad \text{with} \quad p(0, s) = -2n/\sqrt{5} \quad \Rightarrow \quad p(t, s) = -2n/\sqrt{5}, \\
\frac{\partial q}{\partial t}(t, s) &= 0 \quad \text{with} \quad q(0, s) = n/\sqrt{5} \quad \Rightarrow \quad q(t, s) = n/\sqrt{5}, \\
\frac{\partial x}{\partial t}(t, s) &= 2p(t, s) = -4n/\sqrt{5} \quad \text{with} \quad x(0, s) = s \quad \Rightarrow \quad x(t, s) = -4nt/\sqrt{5} + s, \\
\frac{\partial y}{\partial t}(t, s) &= 2q(t, s) = 2n/\sqrt{5} \quad \text{with} \quad y(0, s) = 2s \quad \Rightarrow \quad y(t, s) = 2nt/\sqrt{5} + 2s
\end{aligned}$$

and

$$\frac{\partial u}{\partial t}(t, s) = 2(p^2(t, s) + q^2(t, s)) = 2n^2 \quad \text{with} \quad u(0, s) = 1 \quad \Rightarrow \quad u(t, s) = 2n^2t + 1.$$

We solve $x = -4nt/\sqrt{5} + s$ and $y = 2nt/\sqrt{5} + 2s$ in terms of s and t to get $t = (y - 2x)/(2n\sqrt{5})$ and $s = (x + 2y)/5$. If we substitute these values of s and t in the expression for u , we get the integral surface $u(x, y) = (n(y - 2x))/\sqrt{5} + 1$. A similar reasoning with $p_0(s) = 2n/\sqrt{5}$ and $q_0(s) = -n/\sqrt{5}$ gives $u(x, y) = -(n(y - 2x))/\sqrt{5} + 1$.

b) The possible values of $p_0(s)$, $q_0(s)$ on the curve $\Gamma = \{(x_0(s), y_0(s), u_0(s)) : s \in \mathbb{R}\}$ are given by

$$F(x_0(s), y_0(s), u_0(s), p_0(s), q_0(s)) = p_0^2(s) + q_0^2(s) - n^2 = 0$$

and

$$u'_0(s) = \frac{du_0}{ds}(s) = p_0(s) \frac{dx_0}{ds}(s) + q_0(s) \frac{dy_0}{ds}(s) = x'_0(s)p_0(s) + y'_0(s)q_0(s) .$$

The first equation says that $\|(p_0(s), q_0(s))\|_2 = n$ and the second says that the projection of $(x'_0(s), y'_0(s))$ on $(p_0(s), q_0(s))$ is

$$\frac{\langle (x'_0(s), y'_0(s)), (p_0(s), q_0(s)) \rangle}{\|(p_0(s), q_0(s))\|} = \frac{u'_0(s)}{n} .$$

For this to be possible, the norm of the vector $(x'_0(s), y'_0(s))$ must be greater or equal than $|u'_0(s)|/n$ as can be seen in Figure 1.4 (or from Schwarz inequality). Thus, we need

$$(x'_0(s))^2 + (y'_0(s))^2 \geq \left(\frac{u'_0(s)}{n} \right)^2 . \quad (1.3.28)$$

When (1.3.28) is satisfied, we say that Γ is spacelike and the Cauchy problem has at least one solution. When (1.3.28) is not satisfied, we say that Γ is timelike and the Cauchy problem does not have any solution.

In part (a) above, we had $x_0(s) = s$, $y_0(s) = 2s$ and $u_0(s) = 1$. So (1.3.28) is strictly satisfied and the curve Γ is spacelike. This explains the two solutions that we have found.

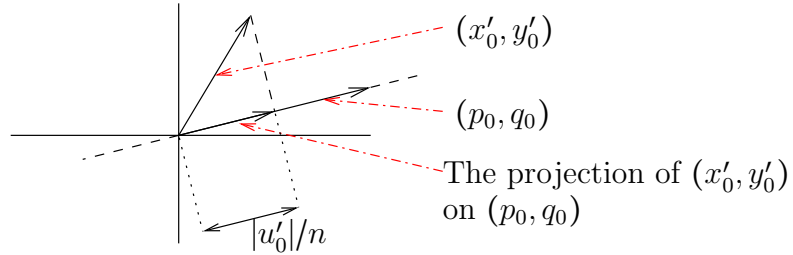


Figure 1.4: The projection of the vector $(x'_0(s), y'_0(s))$ on the vector $(p_0(s), q_0(s))$ is of norm smaller or equal to $\|(p_0(s), q_0(s))\|$.

c) If we consider the curve $\Gamma = \{(x_0(s), y_0(s), u_0(s)) : s \in I\}$, where $x_0(s) = \cos(s)$, $y_0(s) = \sin(s)$ and $u_0(s) = 0$ for $s \in I = \mathbb{R}$. Since (1.3.28) is strictly satisfied, Γ is spacelike and we should find at least one solution; in fact, we will find two solutions.

$p_0(s)$ and $q_0(s)$ satisfy

$$F(x_0(s), y_0(s), u_0(s), p_0(s), q_0(s)) = p_0^2(s) + q_0^2(s) - n^2 = 0$$

and

$$0 = \frac{du_0}{ds}(s) = p_0(s) \frac{dx_0}{ds}(s) + q_0(s) \frac{dy_0}{ds}(s) = -p_0(s) \sin(s) + q_0(s) \cos(s)$$

for $s \in \mathbb{R}$. We get $p_0(s) = n \cos(s)$ and $q_0(s) = n \sin(s)$, or $p_0(s) = -n \cos(s)$ and $q_0(s) = -n \sin(s)$.

We still have the characteristic equation given in (1.3.27).

$$\begin{aligned}\frac{\partial p}{\partial t}(t, s) &= 0 \quad \text{with} \quad p(0, s) = n \cos(s) \quad \Rightarrow \quad p(t, s) = n \cos(s) , \\ \frac{\partial q}{\partial t}(t, s) &= 0 \quad \text{with} \quad q(0, s) = n \sin(s) \quad \Rightarrow \quad q(t, s) = n \sin(s) , \\ \frac{\partial x}{\partial t}(t, s) &= 2p(t, s) = 2n \cos(s) \quad \text{with} \quad x(0, s) = \cos(s) \quad \Rightarrow \quad x(t, s) = (2nt + 1) \cos(s) , \\ \frac{\partial y}{\partial t}(t, s) &= 2q(t, s) = 2n \sin(s) \quad \text{with} \quad y(0, s) = \sin(s) \quad \Rightarrow \quad y(t, s) = (2nt + 1) \sin(s)\end{aligned}$$

and

$$\frac{\partial u}{\partial t}(t, s) = 2(p^2(t, s) + q^2(t, s)) = 2n^2 \quad \text{with} \quad u(0, s) = 0 \quad \Rightarrow \quad u(t, s) = 2n^2 t .$$

We can solve $x = (2nt + 1) \cos(s)$ and $y = (2nt + 1) \sin(s)$ in terms of s and t to get

$$t = \frac{\sqrt{x^2 + y^2}}{2n} - \frac{1}{2n} \quad \text{and} \quad s = \begin{cases} \arctan\left(\frac{y}{x}\right) & x > 0 \\ \frac{\pi}{2} & x = 0, y > 0 \\ \arctan\left(\frac{y}{x}\right) + \pi & x < 0 \\ \frac{3\pi}{2} & x = 0, y < 0 \end{cases}$$

If we substitute these values of s and t in the expression for u , we get the integral surface

$$u(x, y) = n\sqrt{x^2 + y^2} - n .$$

A similar reasoning with $p_0(s) = -n \cos(s)$ and $q_0(s) = -n \sin(s)$ yields

$$u(x, y) = -n\sqrt{x^2 + y^2} + n .$$

We have that $x^2 + y^2 = (1 \pm 2nt)^2$ represents the wavefront because u is constant on these circles. $x^2 + y^2 = (1 + 2nt)^2$ is the wavefront moving away of the origin as t increases while $x^2 + y^2 = (1 - 2nt)^2$ is the wavefront moving toward the origin to reach it at $t = 1/2n$. ♣

Remark 1.3.10

The eikonal equation is of the form $F(x, y, u, p, q) = p^2 + q^2 - n^2 = 0$. Because of the special form of this equation, we can describe the Monge Cones for this problem. They are cones with an axis parallel to the u axis and such that the angle θ between the axis and the side of the cone satisfies $\cot(\theta) = n$ (Figure 1.5). Effectively, each vector $(p, q, -1)$ with $p^2 + q^2 = n^2$ is perpendicular to a tangent plane enveloping the Monge Cone. Moreover, the angle ψ between the vectors $(p, q, -1)$ and the vertical axis satisfies $\tan(\psi) = \sqrt{p^2 + q^2}$. Since a rotation by $\pi/2$ in the vertical plane containing the vector $(p, q, -1)$ and the vertical axis transforms this vector into a vector in the direction of the line of intersection between the Monge Cone and the tangent plane perpendicular to $(p, q, -1)$, the angle θ satisfies $\cot(\theta) = \sqrt{p^2 + q^2} = n$. ♠

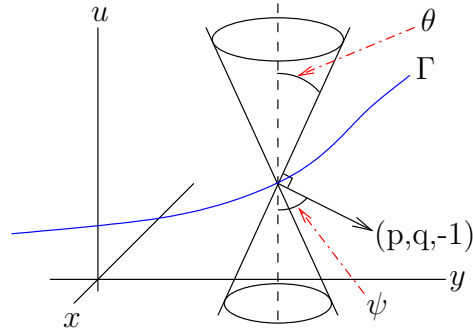


Figure 1.5: Schematic representation of a Monge Cone for the eikonal equation. The curve Γ is spacelike. The curve Γ is inside the cone when Γ is timelike.

1.4 Solution Generated as Envelope

Before addressing the subject of this section, we review what we know about the envelope of a family of integral surfaces. We suppose that

$$u = f(x, y, \lambda) \quad , \quad \lambda \in \mathbb{R} \quad ,$$

describes a one-parameter family of integral surfaces for 1.3.1. Moreover, we assume that its envelope E is given by $u = f(x, y, \lambda(x, y))$ for some function $\lambda : \mathbb{R}^2 \rightarrow \mathbb{R}$.

For each point $(\tilde{x}, \tilde{y}, \tilde{u}) \in E$, we have from Remark 1.2.1 that the tangent plane to the envelope E at $(\tilde{x}, \tilde{y}, \tilde{u})$ is also the tangent plane of the integral surface given by $u = f(x, y, \tilde{\lambda})$ at $(\tilde{x}, \tilde{y}, \tilde{u})$, where $\tilde{\lambda} = \lambda(\tilde{x}, \tilde{y})$.

Moreover, if $h(x, y) = f(x, y, \lambda(x, y))$, then

$$\frac{\partial h}{\partial x}(\tilde{x}, \tilde{y}) = \frac{\partial f}{\partial x}(\tilde{x}, \tilde{y}, \tilde{\lambda}) \quad \text{and} \quad \frac{\partial h}{\partial y}(\tilde{x}, \tilde{y}) = \frac{\partial f}{\partial y}(\tilde{x}, \tilde{y}, \tilde{\lambda})$$

by Remark 1.2.1. Thus,

$$F\left(\tilde{x}, \tilde{y}, h(\tilde{x}, \tilde{y}), \frac{\partial h}{\partial x}(\tilde{x}, \tilde{y}), \frac{\partial h}{\partial y}(\tilde{x}, \tilde{y})\right) = F\left(\tilde{x}, \tilde{y}, f(\tilde{x}, \tilde{y}, \tilde{\lambda}), \frac{\partial f}{\partial x}(\tilde{x}, \tilde{y}, \tilde{\lambda}), \frac{\partial f}{\partial y}(\tilde{x}, \tilde{y}, \tilde{\lambda})\right) = 0 \quad .$$

Since $(\tilde{x}, \tilde{y})$ is arbitrary, we get that the envelope of a family of integral surfaces for (1.3.1) is also an integral surface for (1.3.1).

We now explain how we may find the integral surface of (1.3.1) containing a specific curve Γ as the envelope of a family of integral surfaces of (1.3.1). We first need a method to find a proper family of integral surfaces.

A **complete integral** of (1.3.1) is a two-parameter family of surfaces

$$u = G(x, y, \alpha, \beta) \tag{1.4.1}$$

satisfying (1.3.1) and such that the mapping

$$\begin{aligned} \mathbb{R}^2 &\rightarrow \mathbb{R}^3 \\ (\alpha, \beta) &\mapsto \left(G(x, y, \alpha, \beta), \frac{\partial G}{\partial x}(x, y, \alpha, \beta), \frac{\partial G}{\partial y}(x, y, \alpha, \beta) \right) \end{aligned} \quad (1.4.2)$$

is of rank 2 for all (x, y) . This implies that the parameters α and β are independent (i.e. there is no functional relation between them). There is not any easy method to find complete integrals.

Given a complete integral like (1.4.1) for the partial differential equation (1.3.1), we may define many one-parameter sub-families of integral surfaces. Suppose that

$$u = f(x, y, s) = G(x, y, \alpha(s), \beta(s))$$

for $\alpha, \beta : I \rightarrow \mathbb{R}$ describes a one-parameter sub-family of integral surfaces. We obviously have that $F(x, y, u, p, q) = 0$ if $u = f(x, y, s)$, $p = \frac{\partial f}{\partial x}(x, y, s) = \frac{\partial G}{\partial x}(x, y, \alpha(s), \beta(s))$ and $q = \frac{\partial f}{\partial y}(x, y, s) = \frac{\partial G}{\partial y}(x, y, \alpha(s), \beta(s))$ for all $s \in I$. The envelope of this sub-family of integral surfaces is deduced from

$$u = f(x, y, s) = G(x, y, \alpha(s), \beta(s))$$

and

$$0 = \frac{\partial}{\partial s} f(x, y, s) = \frac{\partial G}{\partial \alpha}(x, y, \alpha(s), \beta(s)) \frac{d\alpha}{ds}(s) + \frac{\partial G}{\partial \beta}(x, y, \alpha(s), \beta(s)) \frac{d\beta}{ds}(s)$$

by solving for s as a function of x and y .

The complete integral can be used to find an integral surface containing a curve $\Gamma = \{x_0(s), y_0(s), u_0(s) : s \in I\}$ for some open interval I . The first step is to extract from the complete integral a one-parameter family of integral surfaces defined by $u = G(x, y, \alpha(s), \beta(s))$ such that $u_0(s) = G(x_0(s), y_0(s), \alpha(s), \beta(s))$ for $s \in I$. The envelope of this one-parameter family of integral surface is the integral surface containing the curve Γ . In other words, we first find a curve

$$\begin{aligned} I &\rightarrow \mathbb{R}^2 \\ s &\mapsto (\alpha(s), \beta(s)) \end{aligned}$$

such that

$$u_0(s) = G(x_0(s), y_0(s), \alpha(s), \beta(s)) \quad (1.4.3)$$

and

$$\frac{du_0}{ds}(s) = \frac{\partial G}{\partial x}(x_0(s), y_0(s), \alpha(s), \beta(s)) \frac{dx_0}{ds}(s)$$

$$+ \frac{\partial G}{\partial y}(x_0(s), y_0(s), \alpha(s), \beta(s)) \frac{dy_0}{ds}(s) \quad (1.4.4)$$

for $s \in I$, and then use

$$u = G(x, y, \alpha(s), \beta(s)) \quad (1.4.5)$$

and

$$\begin{aligned} 0 &= \frac{\partial}{\partial s} G(x, y, \alpha(s), \beta(s)) \\ &= \frac{\partial G}{\partial \alpha}(x, y, \alpha(s), \beta(s)) \frac{d\alpha}{ds}(s) + \frac{\partial G}{\partial \beta}(x, y, \alpha(s), \beta(s)) \frac{d\beta}{ds}(s) \end{aligned} \quad (1.4.6)$$

to find the envelope of the sub-family of surfaces described by $u = G(x, y, \alpha(s), \beta(s))$ that will contains Γ .

To prove that each point $(\tilde{x}, \tilde{y}, \tilde{u}) \in \Gamma$ is on the envelope of the family of integral surfaces given by $u = G(x, y, \alpha(s), \beta(s))$ for $s \in I$, we note that

$$(\tilde{x}, \tilde{y}, \tilde{u}) = (x_0(\tilde{s}), y_0(\tilde{s}), u_0(\tilde{s}))$$

for some $\tilde{s} \in I$. It follows from (1.4.3) that

$$\tilde{u} = u_0(\tilde{s}) = G(x_0(\tilde{s}), y_0(\tilde{s}), \alpha(\tilde{s}), \beta(\tilde{s})) = G(\tilde{x}, \tilde{y}, \alpha(\tilde{s}), \beta(\tilde{s})) .$$

Hence, (1.4.5) is satisfied by $(\tilde{x}, \tilde{y}, \tilde{u})$ and $s = \tilde{s}$.

If we derive (1.4.3) with respect to s , we get

$$\begin{aligned} \frac{du_0}{ds}(s) &= \frac{\partial G}{\partial x}(x_0(s), y_0(s), \alpha(s), \beta(s)) \frac{dx_0}{ds}(s) + \frac{\partial G}{\partial y}(x_0(s), y_0(s), \alpha(s), \beta(s)) \frac{dy_0}{ds}(s) \\ &\quad + \frac{\partial G}{\partial \alpha}(x_0(s), y_0(s), \alpha(s), \beta(s)) \frac{d\alpha}{ds}(s) + \frac{\partial G}{\partial \beta}(x_0(s), y_0(s), \alpha(s), \beta(s)) \frac{d\beta}{ds}(s) \\ &= \frac{du_0}{ds}(s) + \frac{\partial G}{\partial \alpha}(x_0(s), y_0(s), \alpha(s), \beta(s)) \frac{d\alpha}{ds}(s) + \frac{\partial G}{\partial \beta}(x_0(s), y_0(s), \alpha(s), \beta(s)) \frac{d\beta}{ds}(s), \end{aligned}$$

where we have used (1.4.4) to get the second equality. Hence, for $s = \tilde{s}$, we get

$$\begin{aligned} 0 &= \frac{\partial G}{\partial \alpha}(x_0(\tilde{s}), y_0(\tilde{s}), \alpha(\tilde{s}), \beta(\tilde{s})) \frac{d\alpha}{ds}(\tilde{s}) + \frac{\partial G}{\partial \beta}(x_0(\tilde{s}), y_0(\tilde{s}), \alpha(\tilde{s}), \beta(\tilde{s})) \frac{d\beta}{ds}(\tilde{s}) \\ &= \frac{\partial G}{\partial \alpha}(\tilde{x}, \tilde{y}, \alpha(\tilde{s}), \beta(\tilde{s})) \frac{d\alpha}{ds}(\tilde{s}) + \frac{\partial G}{\partial \beta}(\tilde{x}, \tilde{y}, \alpha(\tilde{s}), \beta(\tilde{s})) \frac{d\beta}{ds}(\tilde{s}) = \frac{\partial}{\partial s} G(\tilde{x}, \tilde{y}, \alpha(\tilde{s}), \beta(\tilde{s})) . \end{aligned}$$

Thus, (1.4.6) is satisfied by $(\tilde{x}, \tilde{y}, \tilde{u})$ and $s = \tilde{s}$.

Remark 1.4.1

The one-parameter family of surfaces $u = G(x, y, \alpha(s), \beta(s))$ defined by (1.4.3) and (1.4.4) satisfies

$$F\left(x, y, G(x, y, \alpha(s), \beta(s)), \frac{\partial G}{\partial x}(x, y, \alpha(s), \beta(s)), \frac{\partial G}{\partial y}(x, y, \alpha(s), \beta(s))\right) = 0$$

and

$$\begin{aligned} & \left(\frac{dx_0}{ds}(s), \frac{dy_0}{ds}(s), \frac{du_0}{ds}(s) \right) \\ & \cdot \left(\frac{\partial G}{\partial x}(x_0(s), y_0(s), \alpha(s), \beta(s)), \frac{\partial G}{\partial y}(x_0(s), y_0(s), \alpha(s), \beta(s)), -1 \right) = 0 \end{aligned}$$

for $s \in I$. In particular, if $p_0(s) = \frac{\partial G}{\partial x}(x_0(s), y_0(s), u_0(s), \alpha(s), \beta(s))$ and $q_0(s) = \frac{\partial G}{\partial y}(x_0(s), y_0(s), u_0(s), \alpha(s), \beta(s))$, then

$$F(x_0(s), y_0(s), u_0(s), p_0(s), q_0(s)) = 0$$

and

$$\left(\frac{dx_0}{ds}(s), \frac{dy_0}{ds}(s), \frac{du_0}{ds}(s) \right) \cdot (p_0(s), q_0(s), -1) = 0 .$$

The last equation says that the normal direction $(p_0(s), q_0(s), -1)$ of the tangent plane to the integral surface of (1.3.1) along the curve Γ is perpendicular to the tangential direction $\left(\frac{dx_0}{ds}(s), \frac{dy_0}{ds}(s), \frac{du_0}{ds}(s) \right)$ along Γ .

The possible values for $p_0(s)$ and $q_0(s)$ are therefore given by the intersection of the perpendicular plane to $\left(\frac{dx_0}{ds}(s), \frac{dy_0}{ds}(s), \frac{du_0}{ds}(s) \right)$ with the “cone” produced by the vector $(p, q, -1)$ satisfying

$$F(x_0(s), y_0(s), u_0(s), p, q) = 0$$

as it is shown in Figure 1.6. In Figure 1.6, the cone and the perpendicular plane have been translated from the origin to $(x_0(s), y_0(s), u_0(s))$. ♠

Remark 1.4.2

Given a complete integral $u = G(x, y, \alpha, \beta)$ and a curve Γ defined by

$$\{(x, y, u) = (x_0(s), y_0(s), u_0(s)) : s \in I\} ,$$

where I is an open interval, consider the mapping

$$H : I \times \mathbb{R}^2 \rightarrow \mathbb{R}^2$$

$$\begin{pmatrix} s \\ \alpha \\ \beta \end{pmatrix} \mapsto \begin{pmatrix} u_0(s) - G(x_0(s), y_0(s), \alpha, \beta) \\ \frac{du_0}{ds}(s) - \frac{\partial G}{\partial x}(x_0(s), y_0(s), \alpha, \beta) \frac{dx_0}{ds}(s) + \frac{\partial G}{\partial y}(x_0(s), y_0(s), \alpha, \beta) \frac{dy_0}{ds}(s) \end{pmatrix} .$$

Suppose that $(\tilde{s}, \tilde{\alpha}, \tilde{\beta}) \in I \times \mathbb{R}^2$ is a solution of

$$H \begin{pmatrix} s \\ \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} . \tag{1.4.7}$$

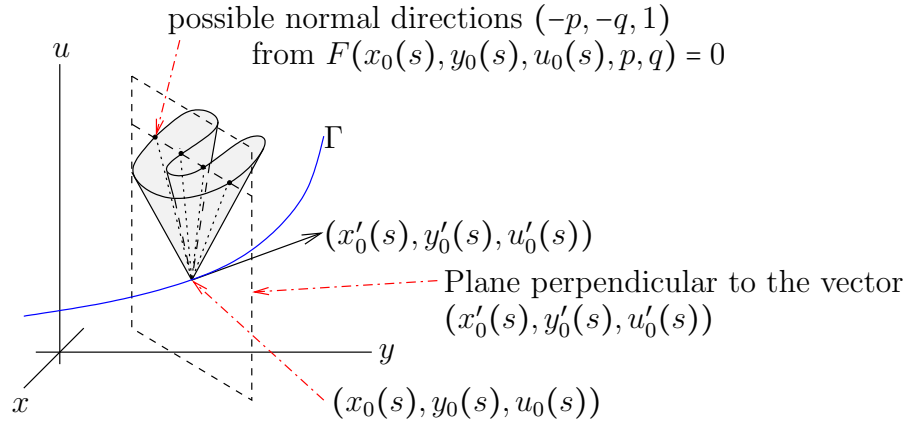


Figure 1.6: The possible initial tangent planes to the integral surfaces for the general Cauchy problem (1.3.1) are associated to the possible normal direction marked by the four thick dots.

The equation in (1.4.7) is equivalent to the union of equations in (1.4.3) and (1.4.4).

We now prove that there exists a differentiable curve $s \mapsto (\alpha(s), \beta(s))$ defined for s in a neighbourhood $J \subset I$ of $\tilde{s}$ such that $(\alpha(\tilde{s}), \beta(\tilde{s})) = (\tilde{\alpha}, \tilde{\beta})$ and

$$H \begin{pmatrix} s \\ \alpha(s) \\ \beta(s) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad (1.4.8)$$

for $s \in J$. Thus, $s \mapsto (\alpha(s), \beta(s))$ satisfies (1.4.3) and (1.4.4) for $s \in J$.

Since the mapping defined in (1.4.2) is of rank 2, we (generally) have that

$$\begin{aligned} \det D_{\alpha, \beta} H(\tilde{s}, \alpha, \beta) \Big|_{(\alpha, \beta) = (\tilde{\alpha}, \tilde{\beta})} &= \det \begin{pmatrix} \frac{\partial G}{\partial \alpha} & \frac{\partial G}{\partial \beta} \\ \frac{\partial^2 G}{\partial x \partial \alpha} \frac{dx_0}{ds} + \frac{\partial^2 G}{\partial y \partial \alpha} \frac{dy_0}{ds} & \frac{\partial^2 G}{\partial x \partial \beta} \frac{dx_0}{ds} + \frac{\partial^2 G}{\partial y \partial \beta} \frac{dy_0}{ds} \end{pmatrix} \\ &= \left(\frac{\partial G}{\partial \alpha} \frac{\partial^2 G}{\partial x \partial \beta} - \frac{\partial G}{\partial \beta} \frac{\partial^2 G}{\partial x \partial \alpha} \right) \frac{dx_0}{ds} + \left(\frac{\partial G}{\partial \alpha} \frac{\partial^2 G}{\partial y \partial \beta} - \frac{\partial G}{\partial \beta} \frac{\partial^2 G}{\partial y \partial \alpha} \right) \frac{dy_0}{ds} \neq 0, \end{aligned}$$

where G and its derivatives are evaluated at $(x, y, \alpha, \beta) = (x_0(\tilde{s}), y_0(\tilde{s}), \tilde{\alpha}, \tilde{\beta})$, and $\frac{dy_0}{ds}$ and $\frac{dx_0}{ds}$ are evaluated at $s = \tilde{s}$. The content of at least one of the two parentheses in the expression above is non-null because the mapping defined in (1.4.2) is of rank 2. It then follows from the implicit function theorem that there exists a unique differentiable curve $s \mapsto (\alpha(s), \beta(s))$, defined in an open neighbourhood J of $\tilde{s}$, such that (1.4.8) is satisfied. and $(\alpha(\tilde{s}), \beta(\tilde{s})) = (\tilde{\alpha}, \tilde{\beta})$ as expected. ♠

Example 1.4.3

Find the complete integral of

$$\frac{\partial u}{\partial x} \frac{\partial u}{\partial y} = u \quad (1.4.9)$$

The function $u(x, y) = 0$ for all x and y satisfies (1.4.9) and the condition $u(x, x) = 0$. Are there non-trivial solutions to (1.4.9) that satisfy $u(x, x) = 0$? Use the complete integral to find a non-trivial solution u of (1.4.9) which satisfies $u(x, x) = 0$.

To find the complete integral, we use the method of separation of variables. If we substitute $u(x, y) = F(x)G(y)$ in (1.4.9), we get $F'(x)G(y)F(x)G'(y) = F(x)G(y)$. If we divide both sides by $F(x)G(y)$, we get $F'(x)G'(y) = 1$. Thus $F'(x) = \frac{1}{G'(y)}$. Since the right hand side is independent of x and the left hand side is independent of y , we must have $F'(x) = \frac{1}{G'(y)} = a$, where a is a constant. We get two first order ordinary differential equation: $F'(x) = a$ and $G'(y) = 1/a$. Their general solutions are $F(x) = ax + b$ and $G(y) = y/a + c$. Hence,

$$u(x, y) = (ax + b) \left(\frac{y}{a} + c \right) = xy + acx + \frac{b}{a}y + bc = xy + \alpha x + \beta y + \alpha\beta$$

for $\alpha = ac$ and $\beta = b/a$. The function

$$u = G(x, y, \alpha, \beta) = xy + \alpha x + \beta y + \alpha\beta$$

is a complete integral because

$$\det \begin{pmatrix} \frac{\partial^2 G}{\partial x \partial \alpha} & \frac{\partial^2 G}{\partial x \partial \beta} \\ \frac{\partial^2 G}{\partial y \partial \alpha} & \frac{\partial^2 G}{\partial y \partial \beta} \end{pmatrix} = \det \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1 \neq 0.$$

Thus, the mapping defined in (1.4.2) is of rank 2.

To find a non-trivial solution u of (1.4.9) which satisfies $u(x, x) = 0$, we first find a family of integral surfaces that touch the curve $\Gamma = \{(x, x, 0) : x \in \mathbb{R}\}$ defined by $x = s$, $y = s$ and $u = 0$ for $s \in \mathbb{R}$. This family is defined by the system of equations

$$\begin{aligned} 0 &= s^2 + \alpha s + \beta s + \alpha\beta \\ 0 &= 2s + \alpha + \beta \end{aligned}$$

obtained from (1.4.5) and (1.4.6). We get $\alpha(s) = -s$ and $\beta(s) = -s$. The one-parameter family of integral surfaces that we are looking for is given by

$$u = G(x, y, -s, -s) = xy - sx - sy + s^2.$$

The envelope of this family of integral surfaces is given by the system of equations

$$\begin{aligned} u &= G(x, y, -s, -s) = xy - sx - sy + s^2 \\ 0 &= \frac{\partial G(x, y, -s, -s)}{\partial s} = -x - y + 2s \end{aligned}$$

obtained from (1.2.1) and (1.2.2). From the second equation, we find $s = (x + y)/2$. If we substitute this value of s in the first equation, we get the required integral surface

$$u = xy - \frac{x(x+y)}{2} - \frac{y(x+y)}{2} + \frac{(x+y)^2}{4} = -\frac{(x-y)^2}{4}$$

contained Γ . ♣

As the following two examples illustrate, the method of separation of variables does not always give a complete integral.

Example 1.4.4

Use the method of separation of variables to find a solution of

$$\frac{\partial u}{\partial x} + \frac{\partial u}{\partial y} = 0 ,$$

and show that it is not a complete integral.

Let $u(x, y) = F(x)G(y)$. We get $F'(x)G(y) + F(x)G'(y) = 0$. If we divide both sides of this equality by $F(x)G(y)$ and isolate the functions of x on one side and those of y on the other side, we get

$$\frac{F'(x)}{F(x)} = -\frac{G'(y)}{G(y)}$$

for all x and y such that $F(x) \neq 0$ and $G(y) \neq 0$. Since the right hand side is independent of x and the left hand side is independent of y , we must have

$$\frac{F'(x)}{F(x)} = -\frac{G'(y)}{G(y)} = \alpha ,$$

where α is a constant. We get two first order ordinary differential equations, $F'(x) - \alpha F(x) = 0$ and $G'(y) + \alpha G(y) = 0$, whose general solutions are $F(x) = \beta_1 e^{\alpha x}$ and $G(y) = \beta_2 e^{-\alpha y}$ respectively. A solution is $u(x, y) = F(x)G(y) = \beta e^{\alpha(x-y)}$, where $\beta = \beta_1 \beta_2$.

However, we have

$$\det \begin{pmatrix} \frac{\partial^2 u}{\partial x \partial \alpha} & \frac{\partial^2 u}{\partial x \partial \beta} \\ \frac{\partial^2 u}{\partial y \partial \alpha} & \frac{\partial^2 u}{\partial y \partial \beta} \end{pmatrix} = \det \begin{pmatrix} \beta(1 + \alpha(x-y))e^{\alpha(x-y)} & \alpha e^{\alpha(x-y)} \\ -\beta(1 + \alpha(x-y))e^{\alpha(x-y)} & -\alpha e^{\alpha(x-y)} \end{pmatrix} = 0$$

for all x and y . Thus, (1.4.2) is not satisfied. ♣

Example 1.4.5

Use the method of separation of variables to find a solution of

$$x \frac{\partial u}{\partial x} - y \frac{\partial u}{\partial y} = 0 ,$$

and show that it is not a complete integral.

Let $u(x, y) = F(x)G(y)$. We get $xF'(x)G(y) - yF(x)G'(y) = 0$. If we divide both sides of this equality by $F(x)G(y)$ and isolate the functions of x on one side and those of y on the other side, we get

$$x \frac{F'(x)}{F(x)} = y \frac{G'(y)}{G(y)}$$

for all x and y such that $F(x) \neq 0$ and $G(y) \neq 0$. Since the right hand side is independent of x and the left hand side is independent of y , we must have

$$x \frac{F'(x)}{F(x)} = y \frac{G'(y)}{G(y)} = \alpha ,$$

where α is a constant. We get two first order ordinary differential equations of the same form: $xF'(x) - \alpha F(x) = 0$ and $yG'(y) - \alpha G(y) = 0$. If we consider the first differential equation, we may write

$$\frac{F'(x)}{F(x)} = \frac{\alpha}{x} .$$

Integrating both sides with respect to x gives $\ln|F(x)| = \alpha \ln|x| + \gamma_1$ for some constant γ_1 . Thus $F(x) = \beta_1|x|^\alpha$, where $\beta_1 = \pm e^{\gamma_1}$. We may also accept $\beta_1 = 0$ to get the trivial solution $u(x, y) = 0$ for all x and y . Similarly, $G(y) = \beta_2|y|^\alpha$. A solution is $u(x, y) = F(x)G(y) = \beta|xy|^\alpha$, where $\beta = \beta_1\beta_2$.

However, we have

$$\det \begin{pmatrix} \frac{\partial^2 u}{\partial x \partial \alpha} & \frac{\partial^2 u}{\partial x \partial \beta} \\ \frac{\partial^2 u}{\partial y \partial \alpha} & \frac{\partial^2 u}{\partial y \partial \beta} \end{pmatrix} = \det \begin{pmatrix} \beta(1 + \alpha \ln(xy))x^{\alpha-1}y^\alpha & \alpha x^{\alpha-1}y^\alpha \\ \beta(1 + \alpha \ln(xy))x^\alpha y^{\alpha-1} & \alpha x^\alpha y^{\alpha-1} \end{pmatrix} = 0$$

for all $x > 0$ and $y > 0$ with similar results for the other choices of x and y . Thus, (1.4.2) is not satisfied. $\clubsuit$

Example 1.4.6

We revisit the eikonal equation of Example 1.3.9 with our knowledge of envelope and complete integral.

The family of integral surfaces for the eikonal equation given by

$$u = G(x, y, \alpha, \beta) = n(x \cos(\alpha) + y \sin(\alpha)) + \beta$$

is a complete integral because

$$\mathbb{R}^2 \rightarrow \mathbb{R}^3$$

$$(\alpha, \beta) \mapsto \left(G(x, y, \alpha, \beta), \frac{\partial G}{\partial x}(x, y, \alpha, \beta), \frac{\partial G}{\partial y}(x, y, \alpha, \beta) \right)$$

is a mapping of rank 2 for all (x, y) . Effectively, the matrix

$$\begin{pmatrix} \frac{\partial G}{\partial \alpha} & \frac{\partial G}{\partial \beta} \\ \frac{\partial^2 G}{\partial x \partial \alpha} & \frac{\partial^2 G}{\partial x \partial \beta} \\ \frac{\partial^2 G}{\partial y \partial \alpha} & \frac{\partial^2 G}{\partial y \partial \beta} \end{pmatrix} = \begin{pmatrix} n(-x \sin(\alpha) + y \cos(\alpha)) & 1 \\ -n \sin(\alpha) & 0 \\ n \cos(\alpha) & 0 \end{pmatrix}$$

is of rank 2 because at least one of the determinants $\det \begin{pmatrix} n(-x \sin(\alpha) + y \cos(\alpha)) & 1 \\ -n \sin(\alpha) & 0 \end{pmatrix}$ and $\det \begin{pmatrix} n(-x \sin(\alpha) + y \cos(\alpha)) & 1 \\ n \cos(\alpha) & 0 \end{pmatrix}$ is non null for all (x, y) .

As in Example 1.3.9, we consider the Cauchy problem with the curve $\Gamma = \{(x_0(s), y_0(s), u_0(s)) : s \in I\}$, where $x_0(s) = \cos(s)$, $y_0(s) = \sin(s)$ and $u_0(s) = 0$ for $s \in I = \mathbb{R}$.

The sub-family of surfaces obtained from the complete integral above with $\alpha = \alpha(s) = s$ and $\beta = \beta(s) = -n$ for all s is the one-parameter family of integral surfaces needed to find the integral surface that contains Γ using the envelope of this family. To verify this, we note that

$$G(x_0(s), y_0(s), \alpha(s), \beta(s)) = n(\cos(s) \cos(s) + \sin(s) \sin(s)) + \beta = 0 = u_0(s)$$

and

$$\begin{aligned} \frac{\partial G}{\partial x}(x_0(s), y_0(s), \alpha(s), \beta(s)) \frac{dx_0}{ds}(s) + \frac{\partial G}{\partial y}(x_0(s), y_0(s), \alpha(s), \beta(s)) \frac{dy_0}{ds}(s) \\ = -n \cos(s) \sin(s) + n \sin(s) \cos(s) = 0 = u'_0(s) \end{aligned}$$

for all s .

The envelope to this family of integral surfaces is given by the equations

$$u = G(x, y, \alpha(s), \beta(s)) = n(x \cos(s) + y \sin(s)) - n \quad (1.4.10)$$

and

$$0 = \frac{\partial}{\partial s} G(x, y, \alpha(s), \beta(s)) = -nx \sin(s) + ny \cos(s) . \quad (1.4.11)$$

(1.4.11) gives $\tan(s) = y/x$. Thus $\cos(s) = x/\sqrt{x^2 + y^2}$ and $\sin(s) = y/\sqrt{x^2 + y^2}$. If we substitute these values of $\cos(s)$ and $\sin(s)$ in (1.4.10), we get

$$u = \frac{nx^2}{\sqrt{x^2 + y^2}} + \frac{ny^2}{\sqrt{x^2 + y^2}} - n = n\sqrt{x^2 + y^2} - n .$$

This is one of the integral surfaces that we found in Example 1.3.9. The other integral surface is given by $\alpha = \alpha(s) = -s$ and $\beta = \beta(s) = -n$ for all s . ♣

1.5 Addendum

The rest of this book is about linear partial differential equations. However, it is worth having a short chapter on nonlinear partial differential equations. This is Chapter 13 at the end of the book. The theory on linear partial differential equations is not easy. The theory on nonlinear partial differential equations is harder.

1.6 Exercises

Suggested exercises:

- In [16]: numbers 1 and 2 in Section 1.6; all the numbers in Section 1.9.
- In [22]: all the numbers in Sections 1.1 and 1.3.
- In [24]: all the numbers in Sections 1.7 and 2.10.
- In [38]: numbers 1 to 4 in Section 1.1; all the numbers in Section 1.2; numbers 4 to 6 in Section 1.5.

Chapter 2

Classification of Partial Differential Equations

2.1 Linear Second Order Partial Differential Equation in the Plane

Let Γ be a curve in $\mathbb{R}^2$, and let $f : \Gamma \rightarrow \mathbb{R}$ and $g : \Gamma \rightarrow \mathbb{R}$ be two analytic functions. We consider the Cauchy Problem of the form

$$\begin{aligned} L(x, y, D) = a_{2,0}(x, y) \frac{\partial^2 u}{\partial x^2}(x, y) + 2a_{1,1}(x, y) \frac{\partial^2 u}{\partial x \partial y}(x, y) + a_{0,2}(x, y) \frac{\partial^2 u}{\partial y^2}(x, y) \\ + a_{1,0}(x, y) \frac{\partial u}{\partial x}(x, y) + a_{0,1}(x, y) \frac{\partial u}{\partial y}(x, y) + a_{0,0}(x, y)u(x, y) + b(x, y) = 0 \end{aligned} \quad (2.1.1)$$

with the conditions

$$u|_{\Gamma} = f \quad \text{and} \quad \frac{\partial u}{\partial \nu} \Big|_{\Gamma} = g, \quad (2.1.2)$$

where $\nu(\mathbf{p})$ for $\mathbf{p} \in \Gamma$ is a normal unit vector to Γ at $\mathbf{p}$. Equation (2.1.1) is a **linear second order partial differential equation**.

We first reduce the Cauchy Problem (2.1.1) and (2.1.2) to one where Γ is the x axis. Let $\mathbf{p}$ be a point on Γ , In a neighbourhood of $\mathbf{p}$, we may define two orthogonal families of curves $\eta(x, y) = h$ and $\xi(x, y) = k$ such that Γ is the curve $\xi(x, y) = 0$ (Figure 2.1). We may also assume that $\mathbf{p}$ is the intersection of the curves $\xi(x, y) = 0$ and $\eta(x, y) = 0$.

The condition for two families of curves to be transverse in a neighbourhood of $\mathbf{p}$ is that

$$\frac{\partial(\eta, \xi)}{\partial(x, y)} = \begin{pmatrix} \frac{\partial \eta}{\partial x} & \frac{\partial \eta}{\partial y} \\ \frac{\partial \xi}{\partial x} & \frac{\partial \xi}{\partial y} \end{pmatrix} \neq 0 \quad (2.1.3)$$

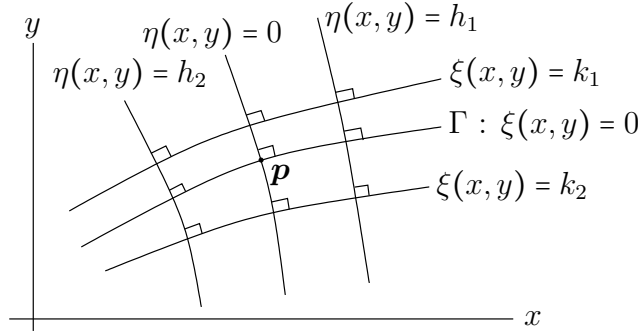


Figure 2.1: The orthogonal families of curves $\eta(x, y) = h$ and $\xi(x, y) = k$ provide a change of coordinates to reduce a linear second order partial differential equation to a standard form

in the neighbourhood of $\mathbf{p}$. Effectively, for $\mathbf{q}$ near $\mathbf{p}$,

$$\begin{aligned} \left. \frac{\partial(\eta, \xi)}{\partial(x, y)} \right|_{\mathbf{q}} = 0 &\Leftrightarrow \left(\frac{\partial\eta}{\partial x}(\mathbf{q}), \frac{\partial\eta}{\partial y}(\mathbf{q}) \right) \cdot \left(-\frac{\partial\xi}{\partial y}(\mathbf{q}), \frac{\partial\xi}{\partial x}(\mathbf{q}) \right) = 0 \\ &\Leftrightarrow \left(-\frac{\partial\xi}{\partial y}(\mathbf{q}), \frac{\partial\xi}{\partial x}(\mathbf{q}) \right) \text{ is orthogonal to } \left(\frac{\partial\eta}{\partial x}(\mathbf{q}), \frac{\partial\eta}{\partial y}(\mathbf{q}) \right) \\ &\Leftrightarrow \left(\frac{\partial\eta}{\partial x}(\mathbf{q}), \frac{\partial\eta}{\partial y}(\mathbf{q}) \right) \text{ is parallel to } \left(\frac{\partial\xi}{\partial x}(\mathbf{q}), \frac{\partial\xi}{\partial y}(\mathbf{q}) \right) \end{aligned}$$

because $\left(-\frac{\partial\xi}{\partial y}(\mathbf{q}), \frac{\partial\xi}{\partial x}(\mathbf{q}) \right)$ is also orthogonal to $\left(\frac{\partial\xi}{\partial x}(\mathbf{q}), \frac{\partial\xi}{\partial y}(\mathbf{q}) \right)$, This shows that the curves $\xi(x, y) = k$ and $\eta(x, y) = h$ have the same perpendicular line at $\mathbf{q}$, therefore the two curves are tangent at $\mathbf{q}$.

Because of (2.1.3), we may use the Inverse Function Theorem to express x and y in a neighbourhood of $\mathbf{p}$ as functions of ξ and η in a neighbourhood of the origin. We will use this change of variables to reduce (2.1.1).

If we substitute

$$\begin{aligned} u(\eta, \xi) &= u(x, y), \quad \frac{\partial u}{\partial x} = \frac{\partial u}{\partial \xi} \frac{\partial \xi}{\partial x} + \frac{\partial u}{\partial \eta} \frac{\partial \eta}{\partial x}, \quad \frac{\partial u}{\partial y} = \frac{\partial u}{\partial \xi} \frac{\partial \xi}{\partial y} + \frac{\partial u}{\partial \eta} \frac{\partial \eta}{\partial y}, \\ \frac{\partial^2 u}{\partial x^2} &= \frac{\partial^2 u}{\partial \xi^2} \left(\frac{\partial \xi}{\partial x} \right)^2 + 2 \frac{\partial^2 u}{\partial \eta \partial \xi} \frac{\partial \eta}{\partial x} \frac{\partial \xi}{\partial x} + \frac{\partial^2 u}{\partial \eta^2} \left(\frac{\partial \eta}{\partial x} \right)^2 + \frac{\partial u}{\partial \xi} \frac{\partial^2 \xi}{\partial x^2} + \frac{\partial u}{\partial \eta} \frac{\partial^2 \eta}{\partial x^2}, \\ \frac{\partial^2 u}{\partial y^2} &= \frac{\partial^2 u}{\partial \xi^2} \left(\frac{\partial \xi}{\partial y} \right)^2 + 2 \frac{\partial^2 u}{\partial \eta \partial \xi} \frac{\partial \eta}{\partial y} \frac{\partial \xi}{\partial y} + \frac{\partial^2 u}{\partial \eta^2} \left(\frac{\partial \eta}{\partial y} \right)^2 + \frac{\partial u}{\partial \xi} \frac{\partial^2 \xi}{\partial y^2} + \frac{\partial u}{\partial \eta} \frac{\partial^2 \eta}{\partial y^2} \end{aligned}$$

and

$$\frac{\partial^2 u}{\partial x \partial y} = \frac{\partial^2 u}{\partial \xi^2} \frac{\partial \xi}{\partial x} \frac{\partial \xi}{\partial y} + \frac{\partial^2 u}{\partial \eta \partial \xi} \frac{\partial \eta}{\partial y} \frac{\partial \xi}{\partial x} + \frac{\partial^2 u}{\partial \eta \partial \xi} \frac{\partial \eta}{\partial x} \frac{\partial \xi}{\partial y} + \frac{\partial^2 u}{\partial \eta^2} \frac{\partial \eta}{\partial x} \frac{\partial \eta}{\partial y} + \frac{\partial u}{\partial \xi} \frac{\partial^2 \xi}{\partial y \partial x} + \frac{\partial u}{\partial \eta} \frac{\partial^2 \eta}{\partial y \partial x}$$

in (2.1.1), we get

$$A_{2,0} \frac{\partial^2 u}{\partial \eta^2} + 2A_{1,1} \frac{\partial^2 u}{\partial \eta \partial \xi} + A_{0,2} \frac{\partial^2 u}{\partial \xi^2} + A_{1,0} \frac{\partial u}{\partial \eta} + A_{0,1} \frac{\partial u}{\partial \xi} + A_{0,0} u + B = 0, \quad (2.1.4)$$

where

$$\begin{aligned} A_{2,0}(\eta, \xi) &= a_{2,0} \left(\frac{\partial \eta}{\partial x} \right)^2 + 2a_{1,1} \frac{\partial \eta}{\partial x} \frac{\partial \eta}{\partial y} + a_{0,2} \left(\frac{\partial \eta}{\partial y} \right)^2, \\ A_{1,1}(\eta, \xi) &= a_{2,0} \frac{\partial \eta}{\partial x} \frac{\partial \xi}{\partial x} + a_{1,1} \left(\frac{\partial \eta}{\partial y} \frac{\partial \xi}{\partial x} + \frac{\partial \eta}{\partial x} \frac{\partial \xi}{\partial y} \right) + a_{0,2} \frac{\partial \eta}{\partial y} \frac{\partial \xi}{\partial y}, \\ A_{0,2}(\eta, \xi) &= a_{2,0} \left(\frac{\partial \xi}{\partial x} \right)^2 + 2a_{1,1} \frac{\partial \xi}{\partial x} \frac{\partial \xi}{\partial y} + a_{0,2} \left(\frac{\partial \xi}{\partial y} \right)^2, \\ A_{1,0}(\eta, \xi) &= a_{2,0} \frac{\partial^2 \eta}{\partial x^2} + 2a_{1,1} \frac{\partial^2 \eta}{\partial x \partial y} + a_{0,2} \frac{\partial^2 \eta}{\partial y^2} + a_{1,0} \frac{\partial \eta}{\partial x} + a_{0,1} \frac{\partial \eta}{\partial y}, \\ A_{0,1}(\eta, \xi) &= a_{2,0} \frac{\partial^2 \xi}{\partial x^2} + 2a_{1,1} \frac{\partial^2 \xi}{\partial x \partial y} + a_{0,2} \frac{\partial^2 \xi}{\partial y^2} + a_{1,0} \frac{\partial \xi}{\partial x} + a_{0,1} \frac{\partial \xi}{\partial y}, \\ A_{0,0}(\eta, \xi) &= a_{0,0} \quad \text{and} \quad B(\eta, \xi) = b(x, y) \end{aligned}$$

with the right hand side of each expression evaluated at $(x, y) = (x(\eta, \xi), y(\eta, \xi))$.

The initial condition given in (2.1.2) becomes

$$u(\eta, 0) = f(x(\eta, 0), y(\eta, 0)) \quad \text{and} \quad \frac{\partial u}{\partial \xi}(\eta, 0) = g(x(\eta, 0), y(\eta, 0)). \quad (2.1.5)$$

As we will see in Chapter 3, to find the series expansion of the solution u of (2.1.4) and (2.1.5) in a neighbourhood of the origin in the η, ξ plane (so near $\mathbf{p}$ in the x, y plane), we need to compute $\frac{\partial^j u}{\partial \xi^j}$ for $j \geq 2$ at the origin. For that, we assume that (2.1.4) can be solved for $\frac{\partial^2 u}{\partial \xi^2}$ at $(\eta, \xi) = (0, 0)$. This is possible only if

$$A_{0,2}(0, 0) = a_{2,0}(\mathbf{p}) \left(\frac{\partial \xi}{\partial x}(\mathbf{p}) \right)^2 + 2a_{1,1}(\mathbf{p}) \frac{\partial \xi}{\partial x}(\mathbf{p}) \frac{\partial \xi}{\partial y}(\mathbf{p}) + a_{0,2}(\mathbf{p}) \left(\frac{\partial \xi}{\partial y}(\mathbf{p}) \right)^2 \neq 0. \quad (2.1.6)$$

This condition is fundamental and is the subject of the next section.

Before moving on to the next section, we note that

$$A_{1,1}^2 - A_{2,0} A_{0,2} = (a_{1,1}^2 - a_{2,0} a_{0,2}) \left(\frac{\partial \eta}{\partial x} \frac{\partial \xi}{\partial y} - \frac{\partial \eta}{\partial y} \frac{\partial \xi}{\partial x} \right)^2 \quad (2.1.7)$$

with the left hand side evaluated at (η, ξ) and the right hand side evaluated at $(x, y) = (x(\eta, \xi), y(\eta, \xi))$. This identity will be used later for the classification of the linear second order partial differential equations. The reader should verify this identity.

2.2 Characteristic curves

As we will see later, not all initial curves Γ will yield a well posed Cauchy problem formed of (2.1.1) and (2.1.2); namely, one with a unique solution.

Definition 2.2.1

A **characteristic curve** for the Cauchy problem (2.1.1) and (2.1.2) is a curve Γ such that it is not possible from (2.1.1) and (2.1.2) to determine all the second order (mixed or not) derivatives of u (in the neighbourhood of Γ).

The previous definition is a little bit vague and does not provide precise conditions to determine if a curve is a characteristic curve or not. We now find such conditions. Suppose that

$$\Gamma = \{(x_0(s), y_0(s)) : s \in I\} \quad (2.2.1)$$

for two differentiable functions $x_0 : I \rightarrow \mathbb{R}$ and $y_0 : I \rightarrow \mathbb{R}$ defined on an open interval I . Then, the derivative of the initial condition $u(x_0(s), y_0(s)) = f(x_0(s), y_0(s))$ yields

$$\frac{d}{ds} f(x_0(s), y_0(s)) = \frac{\partial u}{\partial x}(x_0(s), y_0(s)) x'_0(s) + \frac{\partial u}{\partial y}(x_0(s), y_0(s)) y'_0(s) . \quad (2.2.2)$$

Moreover,

$$\begin{aligned} g(x_0(s), y_0(s)) &= \frac{\partial u}{\partial \nu}(x_0(s), y_0(s)) = \nabla u(x_0(s), y_0(s)) \cdot \nu(x_0(s), y_0(s)) \\ &= \frac{\partial u}{\partial x}(x_0(s), y_0(s)) \nu_1(x_0(s), y_0(s)) + \frac{\partial u}{\partial y}(x_0(s), y_0(s)) \nu_2(x_0(s), y_0(s)) . \end{aligned} \quad (2.2.3)$$

The system of equations (2.2.2) and (2.2.3) determine $\frac{\partial u}{\partial x}$ and $\frac{\partial u}{\partial y}$ uniquely along Γ because

$$\begin{aligned} \det \begin{pmatrix} x'_0(s) & y'_0(s) \\ \nu_1(x_0(s), y_0(s)) & \nu_2(x_0(s), y_0(s)) \end{pmatrix} \\ = (-y'_0(s), x'_0(s)) \cdot (\nu_1(x_0(s), y_0(s)), \nu_2(x_0(s), y_0(s))) = \epsilon \|(y'_0(s), x'_0(s))\| \neq 0 , \end{aligned}$$

where $\epsilon = 1$ or -1 according to the directions of $(-y'_0(s), x'_0(s))$ and $\nu(x_0(s), y_0(s))$. We have that $(-y'_0(s), x'_0(s))$ is perpendicular to the curve Γ at $(x_0(s), y_0(s))$ and so parallel to the unit vector $\nu(x_0(s), y_0(s))$.

We may therefore assume that, instead of $\frac{\partial u}{\partial \nu} \Big|_{\Gamma} = g$, we are given

$$\frac{\partial u}{\partial x} \Big|_{\Gamma} = g_1 \quad \text{and} \quad \frac{\partial u}{\partial y} \Big|_{\Gamma} = g_2 . \quad (2.2.4)$$

For consistency, we have from (2.2.2) that

$$\frac{d}{ds} f(x_0(s), y_0(s)) = g_1((x_0(s), y_0(s))) x'_0(s) + g_2(x_0(s), y_0(s)) y'_0(s) .$$

The derivatives $\frac{\partial^2 u}{\partial x^2}$, $\frac{\partial^2 u}{\partial x \partial y}$ and $\frac{\partial^2 u}{\partial y^2}$ along Γ are the solutions (if there are solutions) of the linear system of equation formed of (2.1.1) and the derivatives with respect to s of the equations in (2.2.4); namely,

$$\begin{aligned} & a_{2,0}(x_0(s), y_0(s)) \frac{\partial^2 u}{\partial x^2}(x_0(s), y_0(s)) + 2a_{1,1}(x_0(s), y_0(s)) \frac{\partial^2 u}{\partial x \partial y}(x_0(s), y_0(s)) \\ & + a_{0,2}(x_0(s), y_0(s)) \frac{\partial^2 u}{\partial y^2}(x_0(s), y_0(s)) = -a_{1,0}(x_0(s), y_0(s)) \frac{\partial u}{\partial x}(x_0(s), y_0(s)) \\ & - a_{0,1}(x_0(s), y_0(s)) \frac{\partial u}{\partial y}(x_0(s), y_0(s)) - a_{0,0}(x_0(s), y_0(s)) u(x_0(s), y_0(s)) - b(x_0(s), y_0(s)) , \\ & \frac{\partial^2 u}{\partial x^2}(x_0(s), y_0(s)) x'_0(s) + \frac{\partial^2 u}{\partial x \partial y}(x_0(s), y_0(s)) y'_0(s) = \frac{d}{ds} g_1(s) \end{aligned}$$

and

$$\frac{\partial^2 u}{\partial x \partial y}(x_0(s), y_0(s)) x'_0(s) + \frac{\partial^2 u}{\partial y^2}(x_0(s), y_0(s)) y'_0(s) = \frac{d}{ds} g_2(s) .$$

This system has a (unique) solution if and only if

$$\begin{aligned} & \det \begin{pmatrix} a_{2,0}(x_0(s), y_0(s)) & 2a_{1,1}(x_0(s), y_0(s)) & a_{0,2}(x_0(s), y_0(s)) \\ x'_0(s) & y'_0(s) & 0 \\ 0 & x'_0(s) & y'_0(s) \end{pmatrix} \\ & = a_{2,0}(x_0(s), y_0(s)) (y'_0(s))^2 - 2a_{1,1}(x_0(s), y_0(s)) x'_0(s) y'_0(s) \\ & + a_{0,2}(x_0(s), y_0(s)) (x'_0(s))^2 \neq 0 . \end{aligned} \quad (2.2.5)$$

This motivates the following definition.

Definition 2.2.2

The **principal symbol** of (2.1.1) is the quadratic form

$$\begin{aligned} Q(x, y, \zeta_1, \zeta_2) &= a_{2,0}(x, y) \zeta_1^2 + 2a_{1,1}(x, y) \zeta_1 \zeta_2 + a_{0,2}(x, y) \zeta_2^2 \\ &= \begin{pmatrix} \zeta_1 & \zeta_2 \end{pmatrix} \begin{pmatrix} a_{2,0}(x, y) & a_{1,1}(x, y) \\ a_{1,1}(x, y) & a_{0,2}(x, y) \end{pmatrix} \begin{pmatrix} \zeta_1 \\ \zeta_2 \end{pmatrix} . \end{aligned}$$

for (x, y) and (ζ_1, ζ_2) in $\mathbb{R}^2$.

We have proved the following proposition.

Proposition 2.2.3

A curve Γ given by (2.2.1) is a characteristic curve if and only if $Q(x_0(s), y_0(s), -y'_0(s), x'_0(s)) = 0$ for all $s \in I$.

Since $(-y'_0(s), x'_0(s))$ is orthogonal to the curve Γ at $(x_0(s), y_0(s))$. The condition above can be restated as follows. A curve Γ is a characteristic curve if and only if at each point (x, y) of Γ we have that $Q(x, y, \zeta_1, \zeta_2) = 0$ for all orthogonal vectors ζ to the curve Γ at (x, y) .

If Γ is given implicitly as $\mu(x, y) = c$ for some constant c , the condition above becomes the following. A curve Γ is a characteristic curve if and only if

$$Q\left(x, y, \frac{\partial \mu}{\partial x}(x, y), \frac{\partial \mu}{\partial y}(x, y)\right) = 0$$

for all $(x, y) \in \Gamma$.

In Section 2.1, we have found that (2.1.6) was necessary and sufficient to reduce the Cauchy problem (2.1.1) and (2.1.2) to one where the curve Γ is replaced by an axis and where we can compute all the partial derivatives of u . The condition in (2.1.6) can be expressed with the principal symbol for (2.1.1); namely, it is the condition

$$Q\left(p_1, p_2, \frac{\partial \xi}{\partial x}(p_1, p_2), \frac{\partial \xi}{\partial y}(p_1, p_2)\right) \neq 0 .$$

This simply says that Γ is not a characteristic curve near $\mathbf{p} = (p_1, p_2)$.

If the curve Γ for the conditions (2.1.2) of (2.1.1) is a characteristic curve, the Cauchy Problem will not be well posed. We may end up with no possible solution or many solutions.

2.3 Classification of Linear Second Order Partial Differential Equation in the Plane

The second order linear partial differential equations can be classify (at each point) into three categories. The existence of characteristic curves differs from one category to the other.

Definition 2.3.1

Let

$$C(x, y) = \det \begin{pmatrix} a_{2,0}(x, y) & a_{1,1}(x, y) \\ a_{1,1}(x, y) & a_{0,2}(x, y) \end{pmatrix} = a_{2,0}(x, y)a_{0,2}(x, y) - a_{1,1}^2(x, y) .$$

1. The partial differential equation in (2.1.1) is **elliptic** at (x, y) if $C(x, y) > 0$.
2. The partial differential equation in (2.1.1) is **hyperbolic** at (x, y) if $C(x, y) < 0$.
3. The partial differential equation in (2.1.1) is **parabolic** at (x, y) if $C(x, y) = 0$.

Since the principal symbol depends on x and y , a partial differential equation may be elliptic in one region of the plane, parabolic in a second region of the plane, and hyperbolic in a third region of the plane.

The two families of curves $\xi(x, y) = k$ and $\eta(x, y) = h$, used in Section 2.1 to reduce the linear second order partial differential equation in (2.1.1) to a simpler form, do not have

to be orthogonal. Moreover, Γ does not have to be a member of one of the families. The only requirement is that the curves of one family intersect the curves of the other family transversely. Obviously, we lose control of the transformations on Γ and the conditions on Γ . In particular, the condition $\left. \frac{\partial u}{\partial \nu} \right|_{\Gamma} = g$ may not be reduced to the simple expression $\frac{\partial u}{\partial \xi}(\eta, 0) = g(x(\eta, 0), y(\eta, 0))$. The derivative with respect to ξ may no longer correspond to a directional derivative perpendicular to Γ .

From (2.1.7), we have that the change of coordinates provided by the transversal families of curves $\xi(x, y) = k$ and $\eta(x, y) = h$ in Section 2.1 does not change the category of the partial differential equation.

To study the properties of the categories of partial differential equation, we will only use families of curves that satisfy the transversal condition when they intersect. In some cases, we will even chose families of characteristic curves.

2.3.1 Elliptic Equations

If (2.1.1) is elliptic on an open set $U \subset \mathbb{R}^2$, then we have

$$\det \begin{pmatrix} a_{2,0}(x, y) & a_{1,1}(x, y) \\ a_{1,1}(x, y) & a_{0,2}(x, y) \end{pmatrix} > 0$$

for $(x, y) \in U$. Thus, the principal symbol is either **positive definite** for all $(x, y) \in U$ or **negative definite** for all $(x, y) \in U$; namely, $Q(x, y, \zeta_1, \zeta_2) > 0$ for all $(\zeta_1, \zeta_2) \in \mathbb{R}^2 \setminus \{(0, 0)\}$ or $Q(x, y, \zeta_1, \zeta_2) < 0$ for all $(\zeta_1, \zeta_2) \in \mathbb{R}^2 \setminus \{(0, 0)\}$ respectively. There is no point other than the origin where the principal symbol is null. Therefore, there are not characteristic curves and, in particular, any curve Γ can be used for the conditions (2.1.2).

We can reduce (2.1.1) to a simpler form. Consider the polynomial

$$P(x, y, \phi) = a_{2,0}(x, y)\phi^2(x, y) - 2a_{1,1}(x, y)\phi(x, y) + a_{0,2}(x, y). \quad (2.3.1)$$

This polynomial is suggested by the condition (2.2.5) used to motivate the principal symbol. There is no real root for this polynomial but there are two distinct complex roots:

$$\phi_{\pm}(x, y) = \frac{a_{1,1}(x, y) \pm i\sqrt{a_{2,0}(x, y)a_{0,2}(x, y) - a_{1,1}^2(x, y)}}{a_{2,0}(x, y)}.$$

If we solve $\frac{dy}{dx} = \phi_+(x, y)$, we get a family of curves $\psi(x, y) = c \in \mathbb{C}$. Since

$$\frac{dy}{dx} = \phi_+(x, y) = -\frac{\partial \psi}{\partial x} \bigg/ \frac{\partial \psi}{\partial y},$$

we get from (2.3.1) that

$$a_{2,0}(x, y) \left(\frac{\partial \psi}{\partial x} \right)^2 + 2a_{1,1}(x, y) \frac{\partial \psi}{\partial x} \frac{\partial \psi}{\partial y} + a_{0,2}(x, y) \left(\frac{\partial \psi}{\partial y} \right)^2 = 0. \quad (2.3.2)$$

Let $\eta(x, y) = \operatorname{Re} \psi(x, y) = \frac{\psi(x, y) + \overline{\psi(x, y)}}{2}$ and $\xi(x, y) = \operatorname{Im} \psi(x, y) = \frac{\psi(x, y) - \overline{\psi(x, y)}}{2i}$. We consider the families of curves

$$\eta(x, y) = h \quad \text{and} \quad \xi(x, y) = k ,$$

where h and k are real constants. Since

$$\frac{\partial \psi}{\partial x} \bigg/ \frac{\partial \psi}{\partial y} = -\phi_+(x, y) \neq -\phi_-(x, y) = \overline{\left(\frac{\partial \psi}{\partial x} \bigg/ \frac{\partial \psi}{\partial y} \right)} = \frac{\partial \bar{\psi}}{\partial x} \bigg/ \frac{\partial \bar{\psi}}{\partial y}$$

because $\operatorname{Im} \phi_{\pm} \neq 0$, we have

$$\begin{aligned} \frac{\partial(\eta, \xi)}{\partial(x, y)} &= \frac{\partial \eta}{\partial x} \frac{\partial \xi}{\partial y} - \frac{\partial \eta}{\partial y} \frac{\partial \xi}{\partial x} = \frac{1}{4i} \left(\frac{\partial \psi}{\partial x} + \overline{\frac{\partial \psi}{\partial x}} \right) \left(\frac{\partial \psi}{\partial y} - \overline{\frac{\partial \psi}{\partial y}} \right) - \frac{1}{4i} \left(\frac{\partial \psi}{\partial y} + \overline{\frac{\partial \psi}{\partial y}} \right) \left(\frac{\partial \psi}{\partial x} - \overline{\frac{\partial \psi}{\partial x}} \right) \\ &= -\frac{1}{2i} \left(\frac{\partial \psi}{\partial x} \overline{\frac{\partial \psi}{\partial y}} - \frac{\partial \psi}{\partial y} \overline{\frac{\partial \psi}{\partial x}} \right) \neq 0 \end{aligned}$$

and so (2.1.3) is satisfied.

We get from (2.3.2) that

$$\begin{aligned} 0 &= a_{2,0}(x, y) \left(\frac{\partial \eta}{\partial x} + i \frac{\partial \xi}{\partial x} \right)^2 + 2a_{1,1}(x, y) \left(\frac{\partial \eta}{\partial x} + i \frac{\partial \xi}{\partial x} \right) \left(\frac{\partial \eta}{\partial y} + i \frac{\partial \xi}{\partial y} \right) + a_{0,2}(x, y) \left(\frac{\partial \eta}{\partial y} + i \frac{\partial \xi}{\partial y} \right)^2 \\ &= \left(a_{2,0}(x, y) \left(\frac{\partial \eta}{\partial x} \right)^2 + 2a_{1,1}(x, y) \frac{\partial \eta}{\partial x} \frac{\partial \eta}{\partial y} + a_{0,2}(x, y) \left(\frac{\partial \eta}{\partial y} \right)^2 \right) \\ &\quad - \left(a_{2,0}(x, y) \left(\frac{\partial \xi}{\partial x} \right)^2 + 2a_{1,1}(x, y) \frac{\partial \xi}{\partial x} \frac{\partial \xi}{\partial y} + a_{0,2}(x, y) \left(\frac{\partial \xi}{\partial y} \right)^2 \right) \\ &\quad + 2i \left(a_{2,0}(x, y) \frac{\partial \eta}{\partial x} \frac{\partial \xi}{\partial x} + a_{1,1}(x, y) \left(\frac{\partial \eta}{\partial y} \frac{\partial \xi}{\partial x} + \frac{\partial \eta}{\partial x} \frac{\partial \xi}{\partial y} \right) + a_{0,2}(x, y) \frac{\partial \eta}{\partial y} \frac{\partial \xi}{\partial y} \right) . \end{aligned}$$

The real and imaginary parts of this expression yield

$$\begin{aligned} &a_{2,0}(x, y) \left(\frac{\partial \eta}{\partial x} \right)^2 + 2a_{1,1}(x, y) \frac{\partial \eta}{\partial x} \frac{\partial \eta}{\partial y} + a_{0,2}(x, y) \left(\frac{\partial \eta}{\partial y} \right)^2 \\ &= a_{2,0}(x, y) \left(\frac{\partial \xi}{\partial x} \right)^2 + 2a_{1,1}(x, y) \frac{\partial \xi}{\partial x} \frac{\partial \xi}{\partial y} + a_{0,2}(x, y) \left(\frac{\partial \xi}{\partial y} \right)^2 \end{aligned}$$

and

$$a_{2,0}(x, y) \frac{\partial \eta}{\partial x} \frac{\partial \xi}{\partial x} + a_{1,1}(x, y) \left(\frac{\partial \eta}{\partial y} \frac{\partial \xi}{\partial x} + \frac{\partial \eta}{\partial x} \frac{\partial \xi}{\partial y} \right) + a_{0,2}(x, y) \frac{\partial \eta}{\partial y} \frac{\partial \xi}{\partial y} = 0 .$$

With this change of coordinates, (2.1.1) becomes (2.1.4) with $A_{2,0} = A_{0,2} \neq 0$ (since Q is positively defined) and $A_{1,1} = 0$. If we divide by $A_{2,0}$, we get an equation of the form

$$\frac{\partial^2 u}{\partial \eta^2} + \frac{\partial^2 u}{\partial \xi^2} + \alpha_{1,0}(\eta, \xi) \frac{\partial u}{\partial \eta} + \alpha_{0,1}(\eta, \xi) \frac{\partial u}{\partial \xi} + \alpha_{0,0}(\eta, \xi) u + \beta(\eta, \xi) = 0 , \quad (2.3.3)$$

where $\alpha_{1,0} = A_{1,0}/A_{2,0}$, etc. (2.3.3) is the **standard form** of an elliptic equation.

Example 2.3.2

The Laplace Equation

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$$

is an example of an elliptic equation on all $\mathbb{R}^2$. ♣

2.3.2 Hyperbolic Equations

Suppose that (2.1.1) is hyperbolic on an open set of $\mathbb{R}^2$, then we have

$$\det \begin{pmatrix} a_{2,0}(x, y) & a_{1,1}(x, y) \\ a_{1,1}(x, y) & a_{0,2}(x, y) \end{pmatrix} < 0$$

for (x, y) in this set. Thus, the principal symbol represents a saddle point in (ζ_1, ζ_2) . There are two lines in $\mathbb{R}^2$ along which the principal symbol is null. Therefore, there are characteristic curves and not all curves Γ can be used for the conditions (2.1.2).

We can reduce (2.1.1) to a simpler form. Consider again the polynomial

$$P(x, y, \phi) = a_{2,0}(x, y)\phi^2(x, y) - 2a_{1,1}(x, y)\phi(x, y) + a_{0,2}(x, y) . \quad (2.3.4)$$

The two real roots of this polynomial are

$$\phi_{\pm}(x, y) = \frac{a_{1,1}(x, y) \pm \sqrt{a_{1,1}^2(x, y) - a_{2,0}(x, y)a_{0,2}(x, y)}}{a_{2,0}(x, y)} .$$

Consider the two families of curves $\eta(x, y) = h$ and $\xi(x, y) = k$ where the first family of curve is given by the implicit solution of $\frac{dy}{dx} = \phi_+(x, y)$ and the second family of curves is given by the implicit solution of $\frac{dy}{dx} = \phi_-(x, y)$.

The relation in (2.1.3) is satisfied because

$$\frac{\partial \eta}{\partial x} \bigg/ \frac{\partial \eta}{\partial y} = -\phi_+(x, y) \neq -\phi_-(x, y) = \frac{\partial \xi}{\partial x} \bigg/ \frac{\partial \xi}{\partial y} ,$$

If we substitute $\phi_{\pm}$ in (2.3.4), we get

$$a_{2,0}(x, y) \left(\frac{\partial \eta}{\partial x} \right)^2 + 2a_{1,1}(x, y) \frac{\partial \eta}{\partial x} \frac{\partial \eta}{\partial y} + a_{0,2}(x, y) \left(\frac{\partial \eta}{\partial y} \right)^2 = 0$$

and

$$a_{2,0}(x, y) \left(\frac{\partial \xi}{\partial x} \right)^2 + 2a_{1,1}(x, y) \frac{\partial \xi}{\partial x} \frac{\partial \xi}{\partial y} + a_{0,2}(x, y) \left(\frac{\partial \xi}{\partial y} \right)^2 = 0 .$$

These two equations are

$$Q\left(x, y, \frac{\partial \eta}{\partial x}(x, y), \frac{\partial \eta}{\partial y}(x, y)\right) = 0 \quad \text{and} \quad Q\left(x, y, \frac{\partial \xi}{\partial x}(x, y), \frac{\partial \xi}{\partial y}(x, y)\right) = 0$$

respectively. Hence, the curves of the two families are characteristic curves.

With the change of coordinates $(x, y) = (x(\eta, \xi), y(\eta, \xi))$, equation (2.1.1) is reduced to (2.1.4) with $A_{2,0} = A_{0,2} = 0$ and $A_{1,1} \neq 0$. If we divide by $A_{1,1}$, we get an equation of the form

$$\frac{\partial^2 u}{\partial \eta \partial \xi} + \alpha_{1,0}(\eta, \xi) \frac{\partial u}{\partial \eta} + \alpha_{0,1}(\eta, \xi) \frac{\partial u}{\partial \xi} + \alpha_{0,0}(\eta, \xi) u + \beta(\eta, \xi) = 0, \quad (2.3.5)$$

where $\alpha_{1,0} = A_{1,0}/A_{1,1}$, etc. (2.3.5) is the **standard form** of an hyperbolic equation.

Example 2.3.3

The wave equation

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}$$

is an example of an hyperbolic equation on all $\mathbb{R}^2$. We will study it extensively later. ♣

Example 2.3.4

Consider the partial differential equation

$$\frac{\partial^2 u}{\partial x \partial y} + 2 \frac{\partial u}{\partial x} = 0. \quad (2.3.6)$$

This is an hyperbolic equation already in its standard form. With $p = \frac{\partial u}{\partial x}$, (2.3.6) becomes

$$\frac{\partial p}{\partial y} + 2p = 0.$$

For x fixed, this is a separable ordinary differential equation whose solution is $p = A(x)e^{-2y}$, where A is an arbitrary function. We now consider

$$\frac{\partial u}{\partial x} = A(x)e^{-2y}.$$

We can integrate both sides with respect to x to get

$$u = \underbrace{\left(\int A(x) dx \right)}_{\equiv B(x)} e^{-2y} + C(y) = B(x)e^{-2y} + C(y),$$

where C is an arbitrary functions. ♣

Example 2.3.5

The partial differential equation

$$\frac{\partial^2 u}{\partial x \partial y} - \frac{\partial^2 u}{\partial y^2} = 0 \quad (2.3.7)$$

is of the form (2.1.1) with $a_{2,0}(x, y) = 0$, $a_{1,1}(x, y) = 1/2$ and $a_{0,2}(x, y) = -1$. Hence

$$\det \begin{pmatrix} a_{2,0}(x, y) & a_{1,1}(x, y) \\ a_{1,1}(x, y) & a_{0,2}(x, y) \end{pmatrix} = \det \begin{pmatrix} 0 & 1/2 \\ 1/2 & -1 \end{pmatrix} = -1/4 < 0$$

and (2.3.7) is hyperbolic. To get the families of curves used to reduce this hyperbolic partial differential equation to its standard form, we need to interchange the role of x and y in the previous discussion. We use

$$\phi_{\pm}(x, y) = \frac{a_{1,1}(x, y) \pm \sqrt{a_{1,1}^2(x, y) - a_{2,0}(x, y)a_{0,2}(x, y)}}{a_{0,2}(x, y)} = -1/2 \mp 1/2 .$$

We get a family of curves $\eta(x, y) = h$ from the implicit solution of $\frac{dx}{dy} = -1$ and another family of curves $\xi(x, y) = k$ from the implicit solution of $\frac{dx}{dy} = 0$. We respectively get $x = -y + \eta$ and $x = \xi$. Hence $\eta(x, y) = x + y$ and $\xi(x, y) = x$.

Since $\frac{\partial \eta}{\partial x} = 1$, $\frac{\partial \eta}{\partial y} = 1$, $\frac{\partial \xi}{\partial x} = 1$ and $\frac{\partial \xi}{\partial y} = 0$, we get

$$\frac{\partial u}{\partial y} = \frac{\partial u}{\partial \eta} \frac{\partial \eta}{\partial y} + \frac{\partial u}{\partial \xi} \frac{\partial \xi}{\partial y} = \frac{\partial u}{\partial \eta} \quad , \quad \frac{\partial^2 u}{\partial y^2} = \frac{\partial}{\partial y} \left(\frac{\partial u}{\partial \eta} \right) = \frac{\partial^2 u}{\partial \eta^2} \frac{\partial \eta}{\partial y} + \frac{\partial^2 u}{\partial \xi \partial \eta} \frac{\partial \xi}{\partial y} = \frac{\partial^2 u}{\partial \eta^2}$$

and

$$\frac{\partial^2 u}{\partial x \partial y} = \frac{\partial}{\partial x} \left(\frac{\partial u}{\partial \eta} \right) = \frac{\partial^2 u}{\partial \eta^2} \frac{\partial \eta}{\partial x} + \frac{\partial^2 u}{\partial \xi \partial \eta} \frac{\partial \xi}{\partial x} = \frac{\partial^2 u}{\partial \eta^2} + \frac{\partial^2 u}{\partial \xi \partial \eta} .$$

If we substitute these expressions for the derivatives in (2.3.7), we get $\frac{\partial^2 u}{\partial \xi \partial \eta} = 0$. We integrate with respect to ξ to get $\frac{\partial u}{\partial \eta} = H(\eta)$ for some function H , and with respect to η to get

$$u = \underbrace{\int H(\eta) d\eta}_{\equiv F(\eta)} + G(\xi) = F(\eta) + G(\xi)$$

for some function G . The final solution is

$$u(x, y) = u(\xi, \eta) = F(\eta) + G(\xi) = F(x + y) + G(x) .$$

♣

Example 2.3.6

The partial differential equation

$$x \frac{\partial^2 u}{\partial x^2} - y \frac{\partial^2 u}{\partial x \partial y} + \frac{\partial u}{\partial x} = 0 \tag{2.3.8}$$

is of the form (2.1.1) with $a_{2,0}(x, y) = x$, $a_{1,1}(x, y) = -y/2$ and $a_{0,2}(x, y) = 0$. Hence

$$\det \begin{pmatrix} a_{2,0}(x, y) & a_{1,1}(x, y) \\ a_{1,1}(x, y) & a_{0,2}(x, y) \end{pmatrix} = \det \begin{pmatrix} x & -y/2 \\ -y/2 & 0 \end{pmatrix} = -y^2/4 .$$

The partial differential equation in (2.3.8) is hyperbolic for $y \neq 0$ and parabolic for $y = 0$. We will consider only the region $y > 0$. We have

$$\phi_{\pm}(x, y) = \frac{a_{1,1}(x, y) \pm \sqrt{a_{1,1}^2(x, y) - a_{2,0}(x, y)a_{0,2}(x, y)}}{a_{2,0}(x, y)} = (-1/2 \pm 1/2) \frac{y}{x} .$$

We get a family of curves $\eta(x, y) = h$ from the implicit solution of $\frac{dy}{dx} = -\frac{y}{x}$ and another family of curves $\xi(x, y) = k$ from the implicit solution of $\frac{dy}{dx} = 0$. From $\frac{dy}{dx} = -\frac{y}{x}$, we get $\eta = xy$. We get $\xi = y$ from $\frac{dy}{dx} = 0$. Since $\frac{\partial \eta}{\partial x} = y = \xi$, $\frac{\partial \eta}{\partial y} = x = \frac{\eta}{\xi}$, $\frac{\partial \xi}{\partial x} = 0$ and $\frac{\partial \xi}{\partial y} = 1$, we get

$$\begin{aligned} \frac{\partial u}{\partial x} &= \frac{\partial u}{\partial \eta} \frac{\partial \eta}{\partial x} + \frac{\partial u}{\partial \xi} \frac{\partial \xi}{\partial x} = \xi \frac{\partial u}{\partial \eta} , \\ \frac{\partial^2 u}{\partial x^2} &= \frac{\partial}{\partial x} \left(\frac{\partial u}{\partial x} \right) = \frac{\partial}{\partial \eta} \left(\xi \frac{\partial u}{\partial \eta} \right) \frac{\partial \eta}{\partial x} + \frac{\partial}{\partial \xi} \left(\xi \frac{\partial u}{\partial \eta} \right) \frac{\partial \xi}{\partial x} = \xi^2 \frac{\partial^2 u}{\partial \eta^2} \end{aligned}$$

and

$$\frac{\partial^2 u}{\partial x \partial y} = \frac{\partial}{\partial y} \left(\frac{\partial u}{\partial x} \right) = \frac{\partial}{\partial \eta} \left(\xi \frac{\partial u}{\partial \eta} \right) \frac{\partial \eta}{\partial y} + \frac{\partial}{\partial \xi} \left(\xi \frac{\partial u}{\partial \eta} \right) \frac{\partial \xi}{\partial y} = \eta \frac{\partial^2 u}{\partial \eta^2} + \frac{\partial u}{\partial \eta} + \xi \frac{\partial^2 u}{\partial \xi \partial \eta} .$$

If we substitute these expressions for the derivatives in (2.3.8), we get

$$0 = \frac{\eta}{\xi} \left(\xi^2 \frac{\partial^2 u}{\partial \eta^2} \right) - \xi \left(\eta \frac{\partial^2 u}{\partial \eta^2} + \frac{\partial u}{\partial \eta} + \xi \frac{\partial^2 u}{\partial \xi \partial \eta} \right) + \xi \frac{\partial u}{\partial \eta} = -\xi^2 \frac{\partial^2 u}{\partial \xi \partial \eta} .$$

Since $\xi = y \neq 0$, we get $\frac{\partial^2 u}{\partial \xi \partial \eta} = 0$. As we have seen before, its solution is $u(\xi, \eta) = F(\eta) + G(\xi)$ for some functions F and G . Therefore, the final solution is

$$u(x, y) = u(\xi, \eta) = F(\eta) + G(\xi) = F(xy) + G(y) .$$

♣

Example 2.3.7

Consider the partial differential equation

$$x \frac{\partial^2 u}{\partial x^2} + 2x^2 \frac{\partial^2 u}{\partial x \partial y} - \frac{\partial u}{\partial x} - 1 = 0 \quad (2.3.9)$$

It is a partial differential equation of the form (2.1.1) with $a_{2,0}(x, y) = x$, $a_{1,1}(x, y) = x^2$ and $a_{0,2}(x, y) = 0$. Hence

$$\det \begin{pmatrix} a_{2,0}(x, y) & a_{1,1}(x, y) \\ a_{1,1}(x, y) & a_{0,2}(x, y) \end{pmatrix} = \det \begin{pmatrix} x & x^2 \\ x^2 & 0 \end{pmatrix} = -x^4$$

and (2.3.9) is parabolic on the y axis and hyperbolic everywhere else. The families of curves used to reduce this hyperbolic partial differential equation to its standard form will work for $x < 0$ and $x > 0$. Since

$$\phi_{\pm}(x, y) = \frac{a_{1,1}(x, y) \pm \sqrt{a_{1,1}^2(x, y) - a_{2,0}(x, y)a_{0,2}(x, y)}}{a_{2,0}(x, y)} = \frac{x^2 \pm x^2}{x} = x \pm x ,$$

we get a family of curves $\eta(x, y) = h$ from the implicit solution of $\frac{dy}{dx} = 2x$ and another family of curves $\xi(x, y) = k$ from the implicit solution of $\frac{dy}{dx} = 0$. The solution of the first equation is $y = x^2 + \eta$ and the solution of the second is $y = \xi$. Hence, $\eta(x, y) = y - x^2$ and $\xi(x, y) = y$.

To complete the example, we assumed that $x > 0$ and thus $x = \sqrt{\xi - \eta}$. The solution below does not change if we assume that $x < 0$ and $x = -\sqrt{\xi - \eta}$.

Since $\frac{\partial \eta}{\partial x} = -2x = -2\sqrt{\xi - \eta}$, $\frac{\partial \eta}{\partial y} = 1$, $\frac{\partial \xi}{\partial x} = 0$ and $\frac{\partial \xi}{\partial y} = 1$, we get

$$\begin{aligned} \frac{\partial u}{\partial x} &= \frac{\partial u}{\partial \eta} \frac{\partial \eta}{\partial x} + \frac{\partial u}{\partial \xi} \frac{\partial \xi}{\partial x} = -2\sqrt{\xi - \eta} \frac{\partial u}{\partial \eta} , \\ \frac{\partial^2 u}{\partial x^2} &= \frac{\partial}{\partial x} \left(\frac{\partial u}{\partial x} \right) = \frac{\partial}{\partial \eta} \left(-2\sqrt{\xi - \eta} \frac{\partial u}{\partial \eta} \right) \frac{\partial \eta}{\partial x} + \frac{\partial}{\partial \xi} \left(-2\sqrt{\xi - \eta} \frac{\partial u}{\partial \eta} \right) \frac{\partial \xi}{\partial x} \\ &= \left(\frac{1}{\sqrt{\xi - \eta}} \frac{\partial u}{\partial \eta} - 2\sqrt{\xi - \eta} \frac{\partial^2 u}{\partial \eta^2} \right) (-2\sqrt{\xi - \eta}) = -2 \frac{\partial u}{\partial \eta} + 4(\xi - \eta) \frac{\partial^2 u}{\partial \eta^2} \end{aligned}$$

and

$$\begin{aligned} \frac{\partial^2 u}{\partial x \partial y} &= \frac{\partial}{\partial y} \left(\frac{\partial u}{\partial x} \right) = \frac{\partial}{\partial \eta} \left(-2\sqrt{\xi - \eta} \frac{\partial u}{\partial \eta} \right) \frac{\partial \eta}{\partial y} + \frac{\partial}{\partial \xi} \left(-2\sqrt{\xi - \eta} \frac{\partial u}{\partial \eta} \right) \frac{\partial \xi}{\partial y} \\ &= \left(\frac{1}{\sqrt{\xi - \eta}} \frac{\partial u}{\partial \eta} - 2\sqrt{\xi - \eta} \frac{\partial^2 u}{\partial \eta^2} \right) + \left(-\frac{1}{\sqrt{\xi - \eta}} \frac{\partial u}{\partial \eta} - 2\sqrt{\xi - \eta} \frac{\partial^2 u}{\partial \eta \partial \xi} \right) \\ &= -2\sqrt{\xi - \eta} \left(\frac{\partial^2 u}{\partial \eta^2} + \frac{\partial^2 u}{\partial \eta \partial \xi} \right) . \end{aligned}$$

If we substitute these expressions for the derivatives in (2.3.9), we get

$$\begin{aligned} 0 &= \sqrt{\xi - \eta} \left(-2 \frac{\partial u}{\partial \eta} + 4(\xi - \eta) \frac{\partial^2 u}{\partial \eta^2} \right) + 2(\xi - \eta) \left(-2\sqrt{\xi - \eta} \left(\frac{\partial^2 u}{\partial \eta^2} + \frac{\partial^2 u}{\partial \eta \partial \xi} \right) \right) \\ &\quad + 2\sqrt{\xi - \eta} \frac{\partial u}{\partial \eta} - 1 = -4(\xi - \eta)^{3/2} \frac{\partial^2 u}{\partial \eta \partial \xi} - 1 . \end{aligned}$$

Thus,

$$\frac{\partial^2 u}{\partial \xi \partial \eta} = -\frac{1}{4} (\xi - \eta)^{-3/2} .$$

An integration with respect to ξ followed by and integration with respect to η yields

$$u(\eta, \xi) = -\sqrt{\xi - \eta} + P(\xi) + Q(\eta)$$

for arbitrary functions P and Q . Finally,

$$u(x, y) = -x + P(y) + Q(y - x^2) .$$

♣

2.3.3 Parabolic Equations

Suppose that (2.1.1) is parabolic on an open set of $\mathbb{R}^2$, then we have

$$\det \begin{pmatrix} a_{2,0}(x, y) & a_{1,1}(x, y) \\ a_{1,1}(x, y) & a_{0,2}(x, y) \end{pmatrix} = 0 \quad (2.3.10)$$

for (x, y) in this set. Hence, the principal symbol is degenerated. There are characteristic curves and not all curves Γ can be used for the conditions (2.1.2).

Without loss of generality, we may assume that $a_{2,0}(x, y) \neq 0$ and $a_{0,2}(x, y) \neq 0$. If $a_{2,0}(x, y) = 0$ or $a_{0,2}(x, y) = 0$, then (2.3.10) gives $a_{1,1}(x, y) = 0$ and we already have a standard form for (2.1.1). If $a_{2,0}(x, y) \neq 0$ and $a_{0,2}(x, y) \neq 0$, it follows from (2.3.10) that $a_{1,1}(x, y) \neq 0$.

We can reduce (2.1.1) to a simpler form. Consider again the polynomial

$$P(x, y, \phi) = a_{2,0}(x, y)\phi^2(x, y) - 2a_{1,1}(x, y)\phi(x, y) + a_{0,2}(x, y) . \quad (2.3.11)$$

There is only one non-zero root of this polynomial, it is

$$\phi(x, y) = \frac{a_{1,1}(x, y)}{a_{2,0}(x, y)} .$$

Consider a family of curve $\eta(x, y) = h$ given by the implicit solution of $\frac{dy}{dx} = \phi(x, y)$. Choose any other family of curves $\xi(x, y) = k$ such that (2.1.3) is satisfied. As we have shown before, it is always possible to find such a family at least locally near a point $\mathbf{p}$ of the curve Γ .

Since

$$\frac{dy}{dx} = \phi(x, y) = -\frac{\partial \eta}{\partial x} \bigg/ \frac{\partial \eta}{\partial y} ,$$

we get after substituting ϕ in (2.3.11) that

$$a_{2,0}(x, y) \left(\frac{\partial \eta}{\partial x} \right)^2 + 2a_{1,1}(x, y) \frac{\partial \eta}{\partial x} \frac{\partial \eta}{\partial y} + a_{0,2}(x, y) \left(\frac{\partial \eta}{\partial y} \right)^2 = 0 . \quad (2.3.12)$$

This equation is $Q\left(x, y, \frac{\partial \eta}{\partial x}(x, y), \frac{\partial \eta}{\partial y}(x, y)\right) = 0$. Hence, the curves of this family are characteristic curves.

It follows (2.3.12) that $A_{2,0} = 0$ in (2.1.7). Since the right hand side of (2.1.7) is null by hypothesis, we must also have that $A_{1,1} = 0$. Therefore, with the change of coordinates $(x, y) = (x(\eta, \xi), y(\eta, \xi))$, equation (2.1.1) is reduced to (2.1.4) with $A_{2,0} = A_{1,1} = 0$ and $A_{0,2} \neq 0$. If we divide by $A_{0,2}$, we get an equation of the form

$$\frac{\partial^2 u}{\partial \xi^2} + \alpha_{1,0}(\eta, \xi) \frac{\partial u}{\partial \eta} + \alpha_{0,1}(\eta, \xi) \frac{\partial u}{\partial \xi} + \alpha_{0,0}(\eta, \xi)u + \beta(\eta, \xi) = 0, \quad (2.3.13)$$

where $\alpha_{1,0} = A_{1,0}/A_{0,2}$, etc. (2.3.13) is the **standard form** of a parabolic equation.

Example 2.3.8

The heat equation

$$\frac{\partial u}{\partial t} = c^2 \frac{\partial^2 u}{\partial x^2}$$

is an example of a parabolic equation on all $\mathbb{R}^2$. ♣

2.4 More on the Classification of Partial Differential Equations

There are several direction to generalize the classification of partial differential equations. We present some of them below.

2.4.1 Classification of Second Order Semi-Linear Partial Differential Equations in the Plane

We consider the Cauchy problem

$$\begin{aligned} L(x, y, u, D) = & a_{2,0}(x, y, u) \frac{\partial^2 u}{\partial x^2}(x, y) + 2a_{1,1}(x, y, u) \frac{\partial^2 u}{\partial x \partial y}(x, y) + a_{0,2}(x, y, u) \frac{\partial^2 u}{\partial y^2}(x, y) \\ & + a_{1,0}(x, y, u) \frac{\partial u}{\partial x}(x, y) + a_{0,1}(x, y, u) \frac{\partial u}{\partial y}(x, y) + a_{0,0}(x, y, u) = 0 \end{aligned} \quad (2.4.1)$$

with the conditions

$$u|_{\Gamma} = f \quad \text{and} \quad \frac{\partial u}{\partial \nu} \Big|_{\Gamma} = g, \quad (2.4.2)$$

where $\nu(x, y)$ for $(x, y) \in \Gamma$ is a normal unit vector to Γ at (x, y) . Equation (2.4.1) is a **second order semi-linear Partial Differential Equation**. The coefficients $a_{i,j}$ may depend on x, y as well as on u .

The classification of the linear second order partial differential equation given in the previous section is still valid for the semi-linear second order partial differential equation. The only difference is that the **principal symbol** is now defined by

$$Q(x, y, u, \zeta_1, \zeta_2) = a_{2,0}(x, y, u) \zeta_1^2 + 2a_{1,1}(x, y, u) \zeta_1 \zeta_2 + a_{0,2}(x, y, u) \zeta_2^2$$

$$= \begin{pmatrix} \zeta_1 & \zeta_2 \end{pmatrix} \begin{pmatrix} a_{2,0}(x, y, u) & a_{1,1}(x, y, u) \\ a_{1,1}(x, y, u) & a_{0,2}(x, y, u) \end{pmatrix} \begin{pmatrix} \zeta_1 \\ \zeta_2 \end{pmatrix}.$$

Now, the principal symbol depends of x , y and u . Unfortunately, there is not any nice method to reduce a second order semi-linear partial differential equation to a standard form.

2.4.2 System of First Order Partial Differential Equations

Like for ordinary differential equations, we may reduce (a system of) partial differential equations of order greater than one to a system of partial differential equations of order one.

Example 2.4.1

The partial differential equation

$$x \frac{\partial^2 u}{\partial x^2} + 2x^2 \frac{\partial^2 u}{\partial x \partial y} - \frac{\partial u}{\partial x} - 1 = 0$$

of Example 2.3.7 can be reduced to a system of first order partial differential equations.

Let $v_1 = \frac{\partial u}{\partial x}$ and $v_2 = \frac{\partial u}{\partial y}$. Then,

$$\frac{\partial v_1}{\partial y} = \frac{\partial^2 u}{\partial x \partial y} = \frac{\partial v_2}{\partial x}$$

and

$$x \frac{\partial v_1}{\partial x} + 2x^2 \frac{\partial v_1}{\partial y} - v_1 - 1 = 0.$$

We get the system

$$\begin{pmatrix} 0 & -1 \\ x & 0 \end{pmatrix} \frac{\partial v}{\partial x} + \begin{pmatrix} 1 & 0 \\ 2x^2 & 0 \end{pmatrix} \frac{\partial v}{\partial y} = \begin{pmatrix} 0 \\ v_1 + 1 \end{pmatrix}, \quad v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}.$$

Once we have found v_1 and v_2 , then u is given by integrating v_1 with respect of x or v_2 with respect to y . ♣

Let $\mathcal{L}(\mathbb{R}^n)$ be the space of $n \times n$ matrices. A **system of linear first order Partial Differential Equations** is a system of the form

$$A(x, y) \frac{\partial v}{\partial x}(x, y) + B(x, y) \frac{\partial v}{\partial y}(x, y) = c(x, y, v(x, y)), \quad (2.4.3)$$

where $A : \mathbb{R}^2 \rightarrow \mathcal{L}(\mathbb{R}^n)$, $B : \mathbb{R}^2 \rightarrow \mathcal{L}(\mathbb{R}^n)$ and $c : \mathbb{R}^{2+n} \rightarrow \mathbb{R}^n$ are continuous functions. If we combine the condition

$$v|_{\Gamma} = f \quad (2.4.4)$$

for some curve Γ in the plane and some function $f : \Gamma \rightarrow \mathbb{R}^n$ with (2.4.3), then we get the general form of the **Cauchy Problem** for systems of first order partial differential equation.

As for single partial differential equation, we need to define characteristic curves.

Definition 2.4.2

A **characteristic curve** for the Cauchy problem (2.4.3) and (2.4.4) is a curve Γ such that (2.4.3) and (2.4.4) are not enough to determine all the first order partial derivatives of v in a neighbourhood of Γ .

As for single partial differential equation, we need to give a precise condition to determine if a curve is or is not a characteristic curve. We limit the discussion to $n = 2$. Suppose that

$$\Gamma = \{(x_0(s), y_0(s)) : s \in I\} , \quad (2.4.5)$$

where $x_0 : I \rightarrow \mathbb{R}$ and $y_0 : I \rightarrow \mathbb{R}$ are differentiable functions, and I is an open interval. The condition on Γ that v must satisfy is of the form

$$v(x_0(s), y_0(s)) = f(s) \quad (2.4.6)$$

for $s \in I$.

There are two differential (systems of) equations that must be satisfied on Γ .

$$\begin{aligned} A(x_0(s), y_0(s)) \frac{\partial v}{\partial x}(x_0(s), y_0(s)) + B(x_0(s), y_0(s)) \frac{\partial v}{\partial y}(x_0(s), y_0(s)) \\ = c(x_0(s), y_0(s), v(x_0(s), y_0(s))) \end{aligned} \quad (2.4.7)$$

and

$$x'_0(s) \frac{\partial v}{\partial x}(x_0(s), y_0(s)) + y'_0(s) \frac{\partial v}{\partial y}(x_0(s), y_0(s)) = \frac{d}{ds} f(s) . \quad (2.4.8)$$

The second equation being the derivative of (2.4.6) with respect to s .

(2.4.7) multiplied by $y'_0(s)$ minus (2.4.8) multiplied by $B(x_0(s), y_0(s))$ (from the left) yields

$$\begin{aligned} (y'_0(s)A(x_0(s), y_0(s)) - x'_0(s)B(x_0(s), y_0(s))) \frac{\partial v}{\partial x}(x_0(s), y_0(s)) \\ = \left(y'_0(s)c(x_0(s), y_0(s), v(x_0(s), y_0(s))) - B(x_0(s), y_0(s)) \frac{d}{ds} f(s) \right) \end{aligned}$$

and (2.4.7) multiplied by $x'_0(s)$ minus (2.4.8) multiplied by $A(x_0(s), y_0(s))$ (from the left) yields

$$\begin{aligned} (x'_0(s)B(x_0(s), y_0(s)) - y'_0(s)A(x_0(s), y_0(s))) \frac{\partial v}{\partial y}(x_0(s), y_0(s)) \\ = \left(x'_0(s)c(x_0(s), y_0(s), v(x_0(s), y_0(s))) - A(x_0(s), y_0(s)) \frac{d}{ds} f(s) \right) . \end{aligned}$$

The first system can be solved for $\frac{\partial v}{\partial x}(x_0(s), y_0(s))$ and the second for $\frac{\partial v}{\partial y}(x_0(s), y_0(s))$ if the matrix

$$(x'_0(s)B(x_0(s), y_0(s)) - y'_0(s)A(x_0(s), y_0(s))) \quad (2.4.9)$$

is invertible.

Definition 2.4.3

The **principal symbol** for (2.4.3) is defined by $Q(x, y, \xi_1, \xi_2) = \xi_1 A(x, y) + \xi_2 B(x, y)$ for (x, y) and ξ_1, ξ_2 in $\mathbb{R}^2$.

The following proposition is a consequence of (2.4.9).

Proposition 2.4.4

A curve Γ given by (2.4.5) is a characteristic curve for the Cauchy Problem (2.4.3) and (2.4.4) if and only if $Q(x_0(s), y_0(s), -y'_0(s), x'_0(s))$ is singular for all $s \in I$.

Since $(-y'_0(s), x'_0(s))$ is perpendicular to the curve Γ at $(x_0(s), y_0(s))$. The condition above can be restated as follows. A curve Γ is a characteristic curve if and only if at each point (x, y) of Γ we have that $Q(x, y, \zeta_1, \zeta_2)$ is singular for all orthogonal vectors ζ to the curve Γ at (x, y) .

If Γ is given implicitly as $\mu(x, y) = c$ for some constant c , the condition above becomes the following. A curve Γ is a characteristic curve if and only if

$$Q\left(x, y, \frac{\partial \mu}{\partial x}(x, y), \frac{\partial \mu}{\partial y}(x, y)\right)$$

is singular for all $(x, y) \in \Gamma$.

2.4.3 Classification of Linear Partial Differential Equations in $\mathbb{R}^n$

Though it is possible to reduce a linear partial differential equation of order greater than one to a system of linear first order partial differential equations, it is not necessarily advantageous to do so. Moreover, since the reduced linear system of first order partial differential equations may not have the same numbers of equations and unknown, it is not trivial to define the principal symbol for such a system (see [28]). For this reason, we present only the classification of general linear partial differential equations of the form

$$L(\mathbf{x}, D)u = \sum_{|\alpha| \leq m} a_\alpha(\mathbf{x}) D^\alpha u = f, \quad (2.4.10)$$

where $u : \mathbb{R}^n \rightarrow \mathbb{R}$ is at least of class C^m , $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is continuous and $a_\alpha : \mathbb{R}^n \rightarrow \mathbb{R}$ is of class C^∞ for all multi-indices involved in the sum. We will relax the conditions on f and a_α starting in Chapter 12.

Definition 2.4.5

A subset S of $\mathbb{R}^n$ is an **hyper-surface** if S is an orientable manifold of dimension $n-1$. We will say that S is an hyper-surface of class C^k (resp. an analytic hyper-surface) if S is a manifold of class C^k (resp. an analytic manifold).

If S be a hyper-surface, then we may express S locally as the solution of an equation $\phi(\mathbf{x}) = 0$ for some function $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$. If S is of class C^k (resp. analytic), we may assume

that ϕ is of class C^k (resp. analytic). We will assume from now on that the hyper-surfaces that we consider are sufficiently smooth.

As for second order partial differential equation in the plane, we have the following definitions.

Definition 2.4.6

Consider the Cauchy problem given by (2.4.10) with conditions on u and its (directional) derivatives of order α with $|\alpha| < m$ on a hyper-surface S . Then S is a **characteristic surface** for this Cauchy problem if (2.4.10) and the conditions on u and its (directional) derivatives are not enough to determine all the partial derivatives (mixed or not) of order m of u in the neighbourhood of S .

Definition 2.4.7

The **principal symbol** (also called the **characteristic form**) associated to the linear partial differential operator $L(\mathbf{x}, D)$ defined in (2.4.10) is

$$Q(\mathbf{x}, \xi) = \sum_{|\alpha|=m} a_\alpha(\mathbf{x}) \xi^\alpha \quad (2.4.11)$$

for $\mathbf{x}$ and ξ in $\mathbb{R}^n$.

We have the following result.

Proposition 2.4.8

An hyper-surface S is a characteristic surface if and only if $Q(\mathbf{x}, \nu(\mathbf{x})) = 0$ for all $\mathbf{x} \in S$, where $\nu(\mathbf{x})$ is any orthogonal vector to S at $\mathbf{x}$.

Before giving a classification of the linear partial differential equations of order greater than 2, we expand the classification of the linear partial differential equations of order 2 given in Section 2.3 to $\mathbb{R}^n$ with $n > 2$.

Consider the linear second order partial differential equation

$$L(\mathbf{x}, D)u = \sum_{i,j=1}^n a_{i,j}(\mathbf{x}) \frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^n b_i(\mathbf{x}) \frac{\partial u}{\partial x_i} + c(\mathbf{x})u = 0, \quad (2.4.12)$$

where $u : \mathbb{R}^n \rightarrow \mathbb{R}$, $a_{i,j} \in C^2(\mathbb{R}^n)$, $b_i \in C^1(\mathbb{R}^n)$ and $c \in C(\mathbb{R}^n)$. Its principal symbol takes the form

$$Q(\mathbf{x}, \xi) = \sum_{i,j=1}^n a_{i,j}(\mathbf{x}) \xi_i \xi_j = \xi^\top A(\mathbf{x}) \xi, \quad (2.4.13)$$

where $A(\mathbf{x})$ is the $n \times n$ matrix with entries $a_{i,j}(\mathbf{x})$.

Definition 2.4.9

Consider the linear partial differential operator $L(\mathbf{x}, D)$ in (2.4.12) with its principal symbol Q given in (2.4.13).

1. $L(\mathbf{x}, D)$ is **elliptic** at $\mathbf{x}$ if all eigenvalues of $A(\mathbf{x})$ have the same sign.
2. $L(\mathbf{x}, D)$ is **hyperbolic** at $\mathbf{x}$ if all eigenvalues of $A(\mathbf{x})$ are non-null and all but one have the same sign.
3. $L(\mathbf{x}, D)$ is **ultrahyperbolic** at $\mathbf{x}$ if all eigenvalues of $A(\mathbf{x})$ are non-null and there are more than one of each sign.
4. $L(\mathbf{x}, D)$ is **parabolic** at $\mathbf{x}$ if $A(\mathbf{x})$ is singular; namely, $A(\mathbf{x})$ has at least one zero eigenvalue.

In $\mathbb{R}^2$, this definition is equivalent to Definition 2.3.1. Recall that all the eigenvalues of a symmetric matrix are real.

We will not attempt to provide a full classification of partial differential equations of order greater than 2. In fact, we will consider only elliptic partial differential equations and prove one nice result about them.

Definition 2.4.10

Let $L(\mathbf{x}, D)$ be the linear partial differential operator defined in (2.4.10) and Q be the associated principal symbol defined in (2.4.11). $L(\mathbf{x}, D)$ is **elliptic** at $\mathbf{x}$ if $Q(\mathbf{x}, \boldsymbol{\xi}) \neq 0$ for all $\boldsymbol{\xi} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$.

In the case where $m = 2$ in (2.4.10), the definition of elliptic partial differential equation given in Definition 2.4.9 is equivalent to the definition given in Definition 2.4.10.

The following proposition shows that the order of an elliptic linear partial differential operator must be even.

Proposition 2.4.11

Let $L(\mathbf{x}, D)$ be an elliptic linear partial differential operator of order m at $\mathbf{x} \in \mathbb{R}^n$, $n > 1$.

1. If $a_\alpha : \mathbb{R}^n \rightarrow \mathbb{R}$ for all multi-indices α such that $|\alpha| = m$, then m is even. Moreover, the associated principal symbol $Q(\mathbf{x}, \boldsymbol{\xi})$ is either positive for all $\boldsymbol{\xi} \neq \mathbf{0}$ or negative for all $\boldsymbol{\xi} \neq \mathbf{0}$.
2. If $a_\alpha : \mathbb{R}^n \rightarrow \mathbb{C}$ for some multi-indices α such that $|\alpha| = m$ and $n > 2$, then m is even.

Proof.

i) Suppose that $a_\alpha : \mathbb{R}^n \rightarrow \mathbb{R}$ for all multi-indices α such that $|\alpha| = m$.

Moreover, suppose that there exist ξ_1 and ξ_2 in $\mathbb{R}^n$ such that $Q(\mathbf{x}, \xi_1) < 0 < Q(\mathbf{x}, \xi_2)$. Since Q is continuous, we may assume that the line $\{(1-s)\xi_1 + s\xi_2 : s \in [0, 1]\}$ does not contain the origin. If the origin is on this line, we replace ξ_1 by another point ξ'_1 near ξ_1 where we still have $Q(\mathbf{x}, \xi'_1) < 0$. Consider the function

$$\begin{aligned} g : [0, 1] &\rightarrow \mathbb{R} \\ s &\mapsto Q(\mathbf{x}, (1-s)\xi_1 + s\xi_2) \end{aligned}$$

Since Q is continuous, g is a continuous function such that $g(0) < 0 < g(1)$. By the Intermediate Value Theorem, there exists $s_0 \in]0, 1[$ such that $g(s_0) = 0$. Thus $Q(\mathbf{x}, (1-s_0)\xi_1 + s_0\xi_2) = 0$ with $(1-s_0)\xi_1 + s_0\xi_2 \neq \mathbf{0}$. This is a contradiction that $L(\mathbf{x}, D)$ is elliptic.

Therefore, for $\xi \in \mathbb{R}^n \setminus \{\mathbf{0}\}$, we have that $Q(\mathbf{x}, -\xi) = (-1)^m Q(\mathbf{x}, \xi)$ and $Q(\mathbf{x}, \xi)$ are of the same sign. This is possible only if m is even.

ii) Suppose that $a_\alpha : \mathbb{R}^n \rightarrow \mathbb{C}$ for some multi-indices α such that $|\alpha| = m$ and $n > 2$.

For each $\mathbf{y} \in \mathbb{R}^{n-1}$, let

$$p_{\mathbf{y}}(z) = Q(\mathbf{x}, (\mathbf{y}, z)) \quad , \quad z \in \mathbb{C} .$$

Let $n_+(\mathbf{y})$ be the number of roots (counted with multiplicity) of $p_{\mathbf{y}}$ with a positive imaginary part, and let $n_-(\mathbf{y})$ be the number of roots (counted with multiplicity) of $p_{\mathbf{y}}$ with a negative imaginary part.

Since $p_{\mathbf{y}}(z) \neq 0$ for all $z \in \mathbb{R}$ and $\mathbf{y} \in \mathbb{R}^{n-1} \setminus \{\mathbf{0}\}$ because $L(\mathbf{x}, D)$ is elliptic, we have $n_+(\mathbf{y}) + n_-(\mathbf{y}) = m$ for $\mathbf{y} \in \mathbb{R}^{n-1} \setminus \{\mathbf{0}\}$.

Moreover, it follows from Rouché's theorem (see [5]) that n_+ and n_- are locally constant. So, they are constant on the largest connected component of $\mathbb{R}^{n-1}$ that does not contain the origin (the only value of $\mathbf{y}$ for which $p_{\mathbf{y}}$ may have real roots). Since $\mathbb{R}^{n-1} \setminus \{\mathbf{0}\}$ is connected for $n > 2$, n_+ and n_- are constant on $\mathbb{R}^{n-1} \setminus \{\mathbf{0}\}$.

Finally, $n_+(-\mathbf{y}) = n_-(\mathbf{y})$ for $\mathbb{R}^{n-1} \setminus \{\mathbf{0}\}$ because $p_{-\mathbf{y}}(-z) = (-1)^m p_{\mathbf{y}}(z)$. Since n_+ and n_- are constant on $\mathbb{R}^{n-1} \setminus \{\mathbf{0}\}$, we get $n_-(\mathbf{y}) = n_+(-\mathbf{y}) = n_+(\mathbf{y})$ for $\mathbf{y} \in \mathbb{R}^{n-1} \setminus \{\mathbf{0}\}$.

Hence, $m = n_+(\mathbf{y}) + n_-(\mathbf{y}) = 2n_+(\mathbf{y})$ and it follows that m is even. ■

To be able to use variational methods to solve elliptic partial differential equations in Chapter 12, elliptic operator in general is not strong enough. We need to request a little more from our elliptic operators. Though we may not use this full generality later, we assume that the coefficients a_α in the linear partial differential operator $L(\mathbf{x}, D)$ defined in (2.4.10) are complex valued functions.

Since we are interested in elliptic linear partial differential equations, it follows from Proposition 2.4.11 that we may assume that the order m of the linear partial differential operator $L(\mathbf{x}, D)$ defined in (2.4.10) is even.

Definition 2.4.12

Let $L(\mathbf{x}, D)$ be the linear partial differential operator of even order defined in (2.4.10) and Q be the associated principal symbol defined in (2.4.11). The operator $L(\mathbf{x}, D)$ is **strongly or uniformly elliptic** on an open subset $\Omega \subset \mathbb{R}^n$ if there exists a constant

$C > 0$ such that

$$(-1)^{m/2} \operatorname{Re} Q(\mathbf{x}, \boldsymbol{\xi}) \geq C \|\boldsymbol{\xi}\|^m \quad (2.4.14)$$

for all $\boldsymbol{\xi} \in \mathbb{R}^n$ and $\mathbf{x} \in \Omega$.

Note that original partial differential equation may have to be multiplied by -1 to get (2.4.14). We assume that this is always done if necessary.

Remark 2.4.13

If the linear partial differential operator $L(\mathbf{x}, D)$ is elliptic at $\mathbf{x}$, then

$$0 < C_{\mathbf{x}} = \inf_{\|\boldsymbol{\xi}\|_2=1} |Q(\mathbf{x}, \boldsymbol{\xi})| < \infty$$

because $\boldsymbol{\xi} \rightarrow |Q(\mathbf{x}, \boldsymbol{\xi})|$ is continuous on the compact set $B = \{\boldsymbol{\xi} \in \mathbb{R}^n : \|\boldsymbol{\xi}\| = 1\}$ and so reaches its minimum at a point of B . If furthermore we assume that the linear partial differential equation is posed on a bounded open set Ω with coefficients continuous on $\overline{\Omega}$, then we may find $C > 0$ such that $C \leq C_{\mathbf{x}}$ for all $\mathbf{x} \in \Omega$ because $\mathbf{x} \rightarrow C_{\mathbf{x}}$ is continuous on the compact set $\overline{\Omega}$. We therefore have that

$$\inf_{\|\boldsymbol{\xi}\|_2=1} |Q(\mathbf{x}, \boldsymbol{\xi})| \geq C$$

for all $\mathbf{x} \in \Omega$. It follows that

$$|Q(\mathbf{x}, \boldsymbol{\xi})| \geq C \|\boldsymbol{\xi}\|^m$$

for all $\boldsymbol{\xi} \in \mathbb{R}^n$ and $\mathbf{x} \in \Omega$. If we also assume that all the coefficients a_{α} in $L(\mathbf{x}, D)$ are functions from $\mathbb{R}^n$ to $\mathbb{R}$, then $\operatorname{Re} Q(\mathbf{x}, \boldsymbol{\xi}) = Q(\mathbf{x}, \boldsymbol{\xi})$ and thus $L(\mathbf{x}, D)$ is strongly elliptic on Ω because we can always multiply the a_{α} by 1 or -1 to get the right sign for the definition of strongly elliptic. ♠

Remark 2.4.14

The definition of a strongly elliptic linear differential operator $L(\mathbf{x}, D)$ as defined in (2.4.10) is sometime stated as follows. A linear differential operator $L(\mathbf{x}, D)$ is strongly elliptic on a open set $\Omega \subset \mathbb{R}^n$ if there exist a complex valued function $\gamma \in C^{\infty}(\overline{\Omega})$ and a constant $C > 0$ such that $|\gamma(\mathbf{x})| = 1$ for all $\mathbf{x} \in \overline{\Omega}$ and

$$\operatorname{Re} (\gamma(\mathbf{x}) Q(\mathbf{x}, \boldsymbol{\xi})) \geq C \|\boldsymbol{\xi}\|^m \quad (2.4.15)$$

for all $\boldsymbol{\xi} \in \mathbb{R}^n$ and $\mathbf{x} \in \overline{\Omega}$. This is not a more general definition than Definition 2.4.12 because we can always multiply the linear partial differential equation by $(-1)^{m/2}/\gamma$ to obtain a linear partial differential equation equivalent to the original one. The coefficients a_{α} of the original linear partial differential equation are replaced by the coefficients $(-1)^{m/2}a_{\alpha}/\gamma$ in the new linear partial differential equation and thus (2.4.15) is reduced to (2.4.14). ♠

2.5 Exercises

Suggested exercises:

- In [22]: numbers 1 to 5 in Section 2.2.
- In [24]: all the numbers in Sections 3.6.
- In [38]: all the numbers in Section 1.6.

Chapter 3

Cauchy Problem

3.1 Cauchy-Kovalevski's Theorem

Let S be an hyper-surface in $\mathbb{R}^n$ and $g_j : S \rightarrow \mathbb{R}$ for $0 \leq j \leq m-1$ be continuous functions. The **Cauchy Problem** is to find a function $u : \mathbb{R}^n \rightarrow \mathbb{R}$ such that

$$G(\mathbf{x}, \{D^\alpha u(\mathbf{x}) : \alpha \in \mathbb{N}^n, |\alpha| \leq m\}) = 0 \quad (3.1.1)$$

for $\mathbf{x} \in \mathbb{R}^n$,

$$u|_S = g_0, \quad \frac{\partial u}{\partial \nu}|_S = g_1, \quad \dots, \quad \frac{\partial^{m-2} u}{\partial \nu^{m-2}}|_S = g_{m-2} \text{ and } \frac{\partial^{m-1} u}{\partial \nu^{m-1}}|_S = g_{m-1}, \quad (3.1.2)$$

where $\nu : S \rightarrow \mathbb{R}^n$ is a differentiable function such that $\nu(\mathbf{x})$ is a unit normal vector to the hyper-surface S at $\mathbf{x} \in S$. Recall that

$$\frac{\partial u}{\partial \nu}(\mathbf{x}) = \lim_{s \rightarrow 0} \frac{u(\mathbf{x} + s\nu(\mathbf{x})) - u(\mathbf{x})}{s} = \nabla u(\mathbf{x}) \cdot \nu(\mathbf{x})$$

is the directional derivative of u at $\mathbf{x}$ in the direction $\nu(\mathbf{x})$.

As we have seen for the one-dimensional partial differential equations, not all hyper-surfaces S yields a Cauchy problem that has a unique solution. We will give later conditions on S that ensure that a Cauchy Problem is well posed.

Using a local changed of coordinates near a point $\mathbf{p}$, we may assume that the hyper-surface S is defined by

$$S = \{\mathbf{x} \in \mathbb{R}^n : x_n = 0\}.$$

Hence the Cauchy Problem becomes

$$F(\mathbf{x}, \{D^\alpha u(\mathbf{x}) : \alpha \in \mathbb{N}^n, |\alpha| \leq m\}) = 0 \quad (3.1.3)$$

for $\mathbf{x} \in \mathbb{R}^n$,

$$u|_S = f_0, \quad \frac{\partial u}{\partial x_n}|_S = f_1, \quad \dots, \quad \frac{\partial^{m-2} u}{\partial x_n^{m-2}}|_S = f_{m-2} \text{ and } \frac{\partial^{m-1} u}{\partial x_n^{m-1}}|_S = f_{m-1}.$$

Let us assume that $F, f_0, f_1, \dots, f_{m-1}$ are analytic functions. Moreover, let us assume that u is also analytic. From $u(x_1, x_2, \dots, x_{n-1}, 0) = f_0(x_1, x_2, \dots, x_{n-1})$, we can compute

$$D^\alpha u(x_1, x_2, \dots, x_{n-1}, 0) = D^\alpha f_0(x_1, x_2, \dots, x_{n-1})$$

for all $\alpha \in \mathbb{N}^n$ with $\alpha_n = 0$. From $\frac{\partial u}{\partial x_n}(x_1, x_2, \dots, x_{n-1}, 0) = f_1(x_1, x_2, \dots, x_{n-1})$, we can compute

$$D^\alpha u(x_1, x_2, \dots, x_{n-1}, 0) = D^\alpha f_1(x_1, x_2, \dots, x_{n-1})$$

for all $\alpha \in \mathbb{N}^n$ with $\alpha_n = 1$. Proceeding like this for the derivative of u up to order $m-1$, we can compute $D^\alpha u(x_1, x_2, \dots, x_{n-1}, 0)$ for all $\alpha \in \mathbb{N}^n$ with $\alpha_n < m$.

The problem is to compute $\frac{\partial^j u}{\partial x_n^j}(x_1, x_2, \dots, x_{n-1}, 0)$ for $j \geq m$. For this, we assume that (3.1.3) can be solved for $\frac{\partial^m u}{\partial x_n^m}$ for $\mathbf{x}$ near the hyper-surface S . Namely, we assume that we can write

$$\frac{\partial^m u}{\partial x_n^m} = H(\mathbf{x}, \{D^\alpha u(\mathbf{x}) : \alpha \in \mathbb{N}^n, |\alpha| \leq m, \alpha \neq m\mathbf{e}_n\}) \quad (3.1.4)$$

for $\mathbf{x} \in \mathbb{R}^n$ near S , where, as usual, $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ is the canonical basis of $\mathbb{R}^n$. Not all hyper-surfaces S will meet this requirement. This is the basis for the classification of partial differential equations given in Chapter 2.

Using (3.1.4), we may compute derivatives of u of order greater or equal to m with respect to x_n . Thus, all the mixed derivatives of u at a point $\mathbf{a} = (a_1, a_2, \dots, a_{n-1}, 0)$ can be computed. Since we assume that u is analytic, we can write for $\mathbf{x}$ near $\mathbf{p}$ that

$$u(\mathbf{x}) = \sum_{\alpha \in \mathbb{N}^n} \frac{1}{\alpha!} D^\alpha u(\mathbf{a}) (\mathbf{x} - \mathbf{p})^\alpha,$$

where

$$\alpha! = \alpha_1! \alpha_2! \dots \alpha_n!$$

and

$$(\mathbf{x} - \mathbf{p})^\alpha = (x_1 - p_1)^{\alpha_1} (x_2 - p_2)^{\alpha_2} \dots (x_n - p_n)^{\alpha_n}.$$

This is the essence of the following theorem that we will not prove.

Theorem 3.1.1 (Cauchy-Kovalevski)

Let $S \subset \mathbb{R}^n$ be a analytic hyper-surface. Assume that G is an analytic function in a neighbourhood of S , and $g_0 : S \rightarrow \mathbb{R}, g_1 : S \rightarrow \mathbb{R}, \dots, g_{m-1} : S \rightarrow \mathbb{R}$ are analytic functions on S . If S is a non-characteristic hyper-surface (See Definition 2.4.6), then there exists a unique analytic solution u of (3.1.1) and (3.1.2) in a neighbourhood of S .

A proof of this theorem can be found in [16, 39] for instance.

The uniqueness in Cauchy-Kovalevski theorem is among the analytic functions. It does not say anything about less smooth solutions. The next theorem roughly states that, for

linear partial differential equations, the uniqueness is true among all functions. This theorem will also not be proved. A proof can be found in [35] for instance.

Theorem 3.1.2 (Holmgren)

Consider the m^{th} order linear partial differential equation

$$L(\mathbf{x}, D)u = \sum_{|\alpha| \leq m} a_\alpha(\mathbf{x}) D^\alpha u = 0$$

on $\mathbb{R}^n$, where the functions $a_\alpha : \mathbb{R}^n \rightarrow \mathbb{R}$ are real analytic. Suppose that S is an analytic non-characteristic hyper-surface. Then, the unique solution of class $C^m(\mathbb{R}^n)$ which satisfy the Cauchy conditions $D^\alpha u = 0$ on S for all multi-indices α with $|\alpha| < m$ is $u = 0$ on R .

[16] proves a little bit more general version of Holmgren's theorem. The proof of Holmgren's theorem is not easy. Before stating this version, we need a little definition.

Definition 3.1.3

Let $\Omega \subset \mathbb{R}^n$ be an open set. Suppose that

1. there exists a real analytic bijection $F : B \times]a, b[\rightarrow \Omega$, where B is an open ball in $\mathbb{R}^{n-1}$ and $a < b$,
2. the Jacobian of F is different from 0 everywhere on $B \times]a, b[$, and
3. $\Omega = \bigcup_{\lambda \in]a, b[} S_\lambda$, where $S_\lambda = \{\mathbf{x} : \mathbf{x} = F(\mathbf{y}, \lambda) \text{ for } \mathbf{y} \in B\}$.

Then the family of hyper-surfaces $\{S_\lambda\}_{\lambda \in]a, b[}$ is an **analytic field** for Ω .

Theorem 3.1.4 (Holmgren)

Consider the m^{th} order linear partial differential equation

$$L(\mathbf{x}, D)u = \sum_{|\alpha| \leq m} a_\alpha(\mathbf{x}) D^\alpha u = 0$$

on an open set Ω of $\mathbb{R}^n$, where the functions $a_\alpha : \Omega \rightarrow \mathbb{R}$ are real analytic. Suppose that $\{S_\lambda\}_{\lambda \in]a, b[}$ is an analytic field for Ω .

Let $R = \{\mathbf{x} : \mathbf{x} \in \Omega \text{ with } x_n \geq 0\}$, $Z = \{\mathbf{x} : \mathbf{x} \in \Omega \text{ with } x_n = 0\}$ and $\Omega_\mu = \bigcup_{a < \lambda \leq \mu} S_\lambda$.

If $Z \neq \emptyset$, Z and the S_λ are not characteristic hyper-surfaces, and $\Omega_\mu \cap R$ is a closed subset of Ω (with the induced topology on Ω) for all $\mu \in]a, b[$, then the unique solution of class $C^m(\overline{R})$ which satisfy the Cauchy conditions $D^\alpha u = 0$ on Z for all multi-indices α with $|\alpha| < m$ is $u = 0$ on R .

If $R \subset \mathbb{R}^n$, recall that $f \in C^m(\overline{R})$ means that $f \in C(\overline{R}) \cap C^m(R)$ and $D^\alpha f$ can be extended

continuously to $\overline{R}$ for all multi-indices α with $|\alpha| \leq m$.

The proof of Theorem 3.1.2 consists in fact in constructing a specific analytic field.

3.2 Lewy Example

The next theorem provide an example of a partial differential equation that cannot have a classical solution if analyticity is dropped from the assumption on the partial differential equation. This theorem is due to Hans Lewy. For this theorem, we consider complex valued functions.

Theorem 3.2.1 (Lewy)

Let

$$L(x, y, D) = \frac{\partial}{\partial x} + i \frac{\partial}{\partial y} - 2i(x + iy) \frac{\partial}{\partial t}$$

and $f : \mathbb{R} \rightarrow \mathbb{R}$ be a continuous function. If $u : \mathbb{R}^3 \rightarrow \mathbb{C}$ is a solution of $L(x, y, D)u = f$ of class C^1 in the neighbourhood of the origin, then f is analytic at $t = 0$.

Proof.

Let

$$v(r, \theta, t) = -e^{i\theta} \sqrt{r} u(\sqrt{r} \cos(\theta), \sqrt{r} \sin(\theta), t)$$

for $(r, \theta, t) \in \Omega_1 =]0, k[\times \mathbb{R} \times]-k, k[$, where k is small enough to have $L(x, y, D)u = f$ in Ω_1 . Then $v \in C^1(\Omega_1)$ and

$$L(x, y, D)u = -2 \frac{\partial v}{\partial r} - \frac{i}{r} \frac{\partial v}{\partial \theta} + 2i \frac{\partial v}{\partial t} = f$$

on Ω_1 if we use the substitution $(x, y) = (\sqrt{r} \cos(\theta), \sqrt{r} \sin(\theta))$. Note that $L(x, y, D)$ in the r, θ coordinates given by $(x, y) = (\sqrt{r} \cos(\theta), \sqrt{r} \sin(\theta))$ is

$$L(r, \theta, D) = 2\sqrt{r} e^{i\theta} \frac{\partial}{\partial r} + \frac{i}{\sqrt{r}} e^{i\theta} \frac{\partial}{\partial \theta} - 2i\sqrt{r} e^{i\theta} \frac{\partial}{\partial t}.$$

Let

$$V(t, r) = \int_0^{2\pi} v(r, \theta, t) d\theta$$

for $(t, r) \in \Omega_2 =]-k, k[\times]0, k[$. Then $V \in C^1(\Omega_2)$ and

$$\begin{aligned} \frac{\partial V}{\partial t}(t, r) + i \frac{\partial V}{\partial r}(t, r) &= \int_0^{2\pi} \left(i \frac{\partial v}{\partial r}(r, \theta, t) - \frac{1}{2r} \frac{\partial v}{\partial \theta}(r, \theta, t) + \frac{\partial v}{\partial t}(r, \theta, t) \right) d\theta \\ &= \frac{1}{2i} \int_0^{2\pi} f(t) d\theta = -\pi i f(t), \end{aligned}$$

where we have use the fact that $\int_0^{2\pi} \frac{\partial v}{\partial \theta}(r, \theta, t) d\theta = v(r, \theta, t) \Big|_0^{2\pi} = 0$ because v is 2π -periodic.

Let $F(t) = \int_0^t f(s) ds$ and $W(t, r) = V(t, r) + \pi i F(t)$ for $(t, r) \in \Omega_2$. The function W satisfies the Cauchy-Riemann Equation

$$\frac{\partial W}{\partial t} + i \frac{\partial W}{\partial r} = 0$$

on Ω_2 . Thus W is an holomorphic function of $w = t + ri$ for $(t, r) \in \Omega_2$.

The continuity of u near the origin implies the continuity of v in Ω_1 . Moreover, $v(0, \theta, t) = 0$ for $(\theta, t) \in \mathbb{R} \times]-k, k[$. Hence, V is continuous on Ω_2 and $V(t, 0) = 0$ for $t \in]-k, k[$. Thus, W is continuous on Ω_2 and $W(t, 0) = \pi i F(t)$ for $t \in]-k, k[$; namely, $\operatorname{Re} W(t, 0) = 0$ for $t \in]-k, k[$.

By Schwarz reflection principle applied to iW , the function

$$\tilde{W}(t, r) = \begin{cases} iW(t, r) & \text{if } r \geq 0 \\ -i\overline{W(t, -r)} & \text{if } r < 0 \end{cases}$$

defines an holomorphic extension on $] -k, k[\times] -k, k[$ of iW . Thus $\tilde{W}(t, 0) = -\pi F(t)$ is real analytic at $t = 0$. It follows that $f = \frac{dF}{dt}$ is real analytic at $t = 0$. ■

Remark 3.2.2

Given $(x_0, y_0, t_0) \in \mathbb{R}^3$, we define a translation in $\mathbb{R}^3$ by

$$\begin{aligned} T : \mathbb{R}^3 &\rightarrow \mathbb{R}^3 \\ (x, y, t) &\mapsto (x + x_0, y + y_0, t + t_0 + 2y_0x - 2x_0y) \end{aligned}$$

If $L(x, y, D)$ is the Lewy operator defined in Theorem 3.2.1, we have $L(x, y, D)(u \circ T) = (L(x, y, D)u) \circ T$ for any function $u \in C^1(\mathbb{R}^3)$. Thus, solving $L(x, y, D)u = f(t)$ near (x_0, y_0, t_0) is equivalent to solving $L(s, y, D)(u \circ T) = f(t + t_0 + 2y_0x - 2x_0y)$ near the origin. It follows from Lewy's Theorem that there is no solution of class C^1 of $L(x, y, D)u = f(t)$ near (x_0, y_0, t_0) unless f is analytic at $t = -t_0 - 2y_0x + 2x_0y$.

H. Lewy has also shown that there are functions $f \in C^\infty(\mathbb{R}^3)$ such that $L(x, y, D)u = f$ has no solution u of class $C^{1,\mu}(\mathbb{R}^n)$ for $\mu > 0$ in a neighborhood of the origin and by extension anywhere in $\mathbb{R}^3$. The spaces $C^{k,\mu}(\mathbb{R}^n)$ of Hölder continuous functions are defined later in Definition 11.7.8. A proof of this result can be found in [35]. ♠

A generalization of Theorem 3.2.1 is given later in Theorem 11.9.2.

3.3 Exercises

Suggested exercises:

- In [22]: 1, 2, 4 and 7 in Section 2.1.

Chapter 4

Review of Functional Analysis

The following material is normally presented in a functional analysis course. We assume that the reader has already taken an introductory course in topology and one in measure theory. Good references for the material of this section are [3, 27, 30, 31]. We particularly suggest [3]. It contains a lot of exercises with detailed solutions to many of them.

We assume that the reader has learned the basic concepts of topological, metric, linear and normed spaces. In particular, we assume that the reader is familiar know what is a **Banach space**; namely, of complete normed space. This is the starting point for this chapter.

4.1 Hilbert Spaces

Definition 4.1.1

An **inner product space** H over the complex numbers is a vector space over $\mathbb{C}$ equipped with a mapping $\langle \cdot, \cdot \rangle : H \times H \rightarrow \mathbb{C}$ such that

1. $\langle x, y \rangle \geq 0$ for all $x, y \in H$.
2. $\langle x, x \rangle = 0$ if and only if $x = 0 \in H$.
3. $\langle x, y \rangle = \overline{\langle y, x \rangle}$ for all $x, y \in H$.
4. $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$ for all $x, y, z \in H$.
5. $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$ for all $x, y \in H$ and $\alpha \in \mathbb{C}$.

The mapping $\langle \cdot, \cdot \rangle : H \times H \rightarrow \mathbb{C}$ is called a **scalar product** or an **inner product** on H .

Hilbert spaces are metric spaces. We can define a norm and a distance in a natural way.

Theorem 4.1.2

If H is an inner product space over the complex numbers, then $\|x\| = \sqrt{\langle x, x \rangle}$ for $x \in H$ defines a norm on H . This norm is called the **norm Induced** by the scalar product.

As expected for a norm, the norm induced by a scalar product has the following properties.

Theorem 4.1.3

If H is an inner product space over the complex numbers, then

1. $|\langle x, y \rangle| \leq \|x\| \|y\|$ for all $x, y \in H$. This is called the **Schwarz inequality**.
2. $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in H$. This is called the **triangle inequality**.

Since an inner product space H is a metric space with the metric $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$ for all $\mathbf{x}, \mathbf{y} \in H$, we can define the convergence of sequences $\{\mathbf{x}_j\}_{j=0}^{\infty}$ in H exactly as it is done in any metric space. The metric d defines a topology on H .

Definition 4.1.4

Let H is an inner product space over the complex numbers with the scalar product $\langle \cdot, \cdot \rangle : H \times H \rightarrow \mathbb{C}$. If H is complete with respect to the norm induce by this scalar product (i.e. every Cauchy sequence in H is converging to an element of H), then H is called an **Hilbert space** over the complex numbers.

Since the inner product defines a metric on an inner product space, we may define the continuity of a function $f : H_1 \rightarrow H_2$ from an inner product space H_1 to an inner product space H_2 as it is normally done for functions between two metric spaces.

Proposition 4.1.5

Let H be an Hilbert space over the complex numbers and let $y \in H$, then

1. $x \mapsto \langle x, y \rangle$ is a continuous linear mapping.
2. $x \mapsto \langle y, x \rangle$ is a continuous conjugate-linear mapping.
3. $x \mapsto \|x\|$ is a continuous mapping.

In addition to open or close subsets of an Hilbert space, there is another type of subsets of an Hilbert space that will play an important role later on when solving partial differential equations.

Definition 4.1.6

A subset V of a vector space H is **convex** if $\lambda x + (1 - \lambda)y \in V$ for all $\lambda \in [0, 1]$ and all $x, y \in V$.

Theorem 4.1.7

If V is a closed, convex and non-empty subset of a Hilbert space H , then there exists a unique $y \in V$ such that $\|y\| = \inf_{x \in V} \|x\|$.

In the previous theorem, V may not be compact.

Definition 4.1.8

A continuous linear function from a Hilbert space H to $\mathbb{C}$ is called a **continuous linear functional** on H . The space of all continuous linear functionals on H is denoted H^* . This space is called the **dual space** of H .

Theorem 4.1.9 (Riesz Representation Theorem)

Let H be a Hilbert space and let $g : H \rightarrow \mathbb{C}$ be a continuous linear functional, then there exists a unique $y \in H$ such that $g(x) = \langle x, y \rangle$ for all $x \in H$.

Since an inner product is defined on a Hilbert space, it is possible to define the angle between elements of a Hilbert space; in particular, it is possible to define when two elements are orthogonal to each other.

Definition 4.1.10

Let H be a Hilbert space. Two elements $x_1, x_2 \in H$ are **orthogonal** or **perpendicular** if $\langle x_1, x_2 \rangle = 0$. A set $S = \{x_\alpha\}_{\alpha \in A} \subset H$, where A is some index set, is **orthogonal** if $\langle x_\alpha, x_\beta \rangle = 0$ for $\alpha \neq \beta$ and $\langle x_\alpha, x_\beta \rangle \neq 0$ for $\alpha = \beta$. The set S is **orthonormal** if

$$\langle x_\alpha, x_\beta \rangle = \begin{cases} 0 & \text{if } \alpha \neq \beta \\ 1 & \text{if } \alpha = \beta \end{cases}$$

Theorem 4.1.11

Let H be a Hilbert space and let $\{x_j\}_{j=1}^J$ be a finite orthonormal subset of H . Given $x \in H$, we have

$$\left\| x - \sum_{j=1}^J \langle x, x_j \rangle x_j \right\| \leq \left\| x - \sum_{j=1}^J \lambda_j x_j \right\|$$

for all $\lambda_j \in \mathbb{C}$, and

$$\left\| x - \sum_{j=1}^J \langle x, x_j \rangle x_j \right\| = \left\| x - \sum_{j=1}^J \lambda_j x_j \right\|$$

implies that $\lambda_j = \langle x, x_j \rangle$ for all j .

Corollary 4.1.12

Let H be an Hilbert space and $S = \{x_\alpha\}_{\alpha \in A}$, where A is some index set, be an orthonormal subset of H . Then

$$\sum_{\alpha \in A} |\langle x, x_\alpha \rangle|^2 \leq \|x\|^2 \quad (4.1.1)$$

for all $x \in H$.

The sum in (4.1.1) is in fact the square of the norm of the function

$$\begin{aligned} F_x : A &\rightarrow \mathbb{C} \\ \alpha &\mapsto \langle x, x_\alpha \rangle \end{aligned}$$

in the space $\ell^2(A)$ of square integrable functions with respect to the counting measure on A . In other words, (4.1.1) is

$$\int_A |F_x(\alpha)|^2 d\nu(\alpha) \leq \|x\|^2,$$

where ν is the counting measure on A . The previous corollary states that $F_x \in \ell^2(A)$ for all $x \in H$.

Since $F_x \in \ell^2(A)$, it follows that the set of all α such that $F_x(\alpha) \neq 0$ is at most countable. More precisely, for each positive integer n , the set of all α such that $|F_x(\alpha)| > 1/n$ is finite or empty.

Theorem 4.1.13

Let H be an Hilbert space and $S = \{x_\alpha\}_{\alpha \in A}$, where A is some index set, be an orthonormal subset of H . The following statements are equivalent

1. S is a maximal orthonormal subset of H .
2. The set of all finite linear combinations of elements of S is dense in H .
3. For every $x \in H$, we have $\|x\|^2 = \sum_{\alpha \in A} |\langle x, x_\alpha \rangle|^2$.
4. For every $x, y \in H$, we have $\langle x, y \rangle = \sum_{\alpha \in A} \langle x, x_\alpha \rangle \langle y, x_\alpha \rangle$.

Definition 4.1.14

Let H be an Hilbert space. A **complete orthonormal set** or an **orthonormal basis** for H is an orthonormal set $S = \{x_\alpha\}_{\alpha \in A} \subset H$, where A is some index set, such that S satisfies any of the equivalent statements in Theorem 4.1.13.

A **complete orthogonal set** or a **orthogonal basis** for H is an orthogonal set $S = \{x_\alpha\}_{\alpha \in A}$, where A is some index set, such that $\{\|x_\alpha\|^{-1}x_\alpha\}_{\alpha \in A}$ is an orthonormal basis.

A consequence of the previous definition is that a complete orthogonal set S is an orthogonal set such that all finite linear combinations of elements of S is dense in H .

If $S = \{x_\alpha\}_{\alpha \in A}$, where A is some index set, is an orthonormal basis of an Hilbert space H , it follows from Theorem 4.1.13 that

$$\begin{aligned} T : H &\rightarrow \ell^2(A) \\ x &\mapsto F_x \end{aligned} \tag{4.1.2}$$

is a one-to-one and onto isometry, and moreover $\langle x, y \rangle = \int_A F_x(\alpha) \overline{F_y(\alpha)} d\nu(\alpha)$, where ν is the counting measure on A . A consequence of this isomorphism is that orthonormal bases in H are isomorphic (i.e. have the same cardinality).

4.2 General L^2 Spaces

Suppose that μ is a measure on a measurable space Ω . Let $L^2(\Omega) \equiv L^2(\Omega, \mu)$ be the space of measurable functions $f : \Omega \rightarrow \mathbb{C}$ such that $\int_\Omega |f|^2 d\mu$ is finite. We can define a scalar product on $L^2(\Omega)$ by

$$\langle f, g \rangle = \int_\Omega f \overline{g} d\mu \quad , \quad f, g \in L^2(\Omega) .$$

The associated L^2 -norm is

$$\|f\|_2 = \sqrt{\langle f, f \rangle} = \left(\int_\Omega |f|^2 d\mu \right)^{1/2} , \quad f \in L^2(\Omega) .$$

Equipped with this norm, $L^2(\Omega)$ is an Hilbert space.

Remark 4.2.1

We use the term “function” for an element $f \in L^2(\Omega)$. The reader should remember that in fact f represents the equivalence class of functions which are equal to f almost everywhere. ♠

It follows from Theorem 4.1.13, that an orthonormal set of functions $\{v_\alpha\}_{\alpha \in A}$, where A is some index set, is complete if $\langle f, v_\alpha \rangle = 0$ for all $\alpha \in A$ implies that $f = 0$ almost everywhere on Ω .

Let $S = \{v_\alpha\}_{\alpha \in A}$, where A is some index set, be an orthonormal basis of $L^2(\Omega)$. The **(generalized) Fourier series** of a function $f \in L^2(\Omega)$ with respect to the orthonormal basis S is

$$\sum_{\alpha \in A} a_\alpha v_\alpha ,$$

where

$$a_\alpha = \langle f, v_\alpha \rangle = \int_\Omega f \overline{v_\alpha} d\mu \quad , \quad \alpha \in A .$$

Recall that the set of all α such that $a_\alpha \neq 0$ is at most countable. So, the sum is an ordinary sum. More precisely, take any ordering $\{\alpha_j\}_{j=0}^\infty$ of the set of index $\alpha \in A$ such that $a_\alpha \neq 0$.

Because of the isometry defined in (4.1.2), we have

$$\left\| f - \sum_{j=0}^J a_{\alpha_j} v_{\alpha_j} \right\|_2 = \left(\int_{j \in B_J} |F_f(j)|^2 d\nu \right)^{1/2} \rightarrow 0 \quad \text{as } J \rightarrow \infty, \quad (4.2.1)$$

where $F_f : \mathbb{N} \rightarrow \mathbb{C}$ is defined by $F_f(j) = \langle f, v_{\alpha_j} \rangle$ for all $j \in \mathbb{N}$, $B_J = \{J+1, J+2, \dots\}$ and ν is the counting measure on $\mathbb{N}$. The convergence to 0 of the integral above is due to the fact that $F_f \in \ell^2(\mathbb{N})$. In other words, (4.2.1) states that

$$\left(\int_{\Omega} \left| f - \sum_{j=1}^J a_{\alpha_j} v_{\alpha_j} \right|^2 d\mu \right)^{1/2} = \left(\sum_{j=J+1}^{\infty} |a_{\alpha_j}|^2 \right)^{1/2} \rightarrow 0 \quad \text{as } J \rightarrow \infty$$

whatever the ordering $\{\alpha_j\}_{j=0}^{\infty}$ of the set of indices $\alpha \in A$ such that $a_{\alpha} \neq 0$. We write

$$f = \sum_{\alpha \in A} a_{\alpha} v_{\alpha}$$

whenever 4.2.1 is satisfied. This does not mean that we have pointwise convergence. We address this issue in Section 4.2.3 below.

4.2.1 L^2 Spaces on the Real Line

An important special case of the general theory of L^2 spaces is given by $\Omega = [a, b]$ equipped with the measure μ defined on $[a, b]$ by

$$\mu(I) = \int_I p(x) dx$$

for all Lebesgue measurable subsets I of $[a, b]$, where $p : [a, b] \rightarrow \mathbb{R}$ is a piecewise continuous function on the interval $[a, b]$ such that $p(x) > 0$ for almost all $x \in [a, b]$. We could also have used an open interval $]a, b[$, bounded or not bounded. The results are the same.

Recall that a function f is **piecewise continuous** on an interval I if f is continuous on the interval I except at a finite number of points and $f(c^+) = \lim_{x \rightarrow c^+} f(x)$ and $f(c^-) = \lim_{x \rightarrow c^-} f(x)$ for all $c \in I$; in particular at the points where f is discontinuous.

$L^2[a, b] \equiv L^2([a, b], \mu)$ is the space of measurable functions $f : [a, b] \rightarrow \mathbb{R}$ such that $\int_a^b f^2 d\mu = \int_a^b f^2(x)p(x) dx$ is finite. The scalar product on $L^2[a, b]$ is

$$\langle f, g \rangle = \int_a^b f(x)g(x)p(x) dx \quad , \quad f, g \in L^2[a, b]. \quad (4.2.2)$$

The associated L^2 -norm is

$$\|f\|_2 = \sqrt{\langle f, f \rangle} = \left(\int_a^b f^2(x)p(x) dx \right)^{1/2} \quad , \quad f \in L^2[a, b]. \quad (4.2.3)$$

The $L^2[a, b]$ spaces have countable orthonormal bases. Suppose that $\{v_n : n \in \mathbb{N}\} \subset L^2[a, b]$ is an orthonormal base for $L^2[a, b]$, the Fourier series of a function $f \in L^2[a, b]$ with respect to this bases is

$$\sum_{n=0}^{\infty} a_n v_n , \quad (4.2.4)$$

where

$$a_n = \langle f, v_n \rangle = \int_a^b f(x) v_n(x) p(x) dx \quad , \quad n \in \mathbb{N} .$$

Sometime, we only have a complete orthogonal set of functions $\{v_n : n \in \mathbb{N}\} \subset L^2[a, b]$. In this case, the Fourier series of $f \in L^2[a, b]$ with respect to this set of functions is

$$\sum_{n=0}^{\infty} a_n v_n , \quad (4.2.5)$$

where

$$a_n = \frac{\langle f, v_n \rangle}{\|v_n\|^2} = \left(\int_a^b (v_n(x))^2 p(x) dx \right)^{-1} \int_a^b f(x) v_n(x) p(x) dx \quad , \quad n \in \mathbb{N} .$$

The convergence of the series in (4.2.4) and (4.2.5) implies that

$$\left\| f - \sum_{n=0}^N a_n v_n \right\|_2 = \left(\int_a^b \left(f(x) - \sum_{n=0}^N a_n v_n(x) \right)^2 p(x) dx \right)^{1/2} \rightarrow 0 \quad \text{as } N \rightarrow \infty .$$

We say that $\sum_{n=0}^N a_n v_n$ converges in L^2 to f as $N \rightarrow \infty$, and we write

$$f = \sum_{n=0}^{\infty} a_n v_n .$$

Recall that this equality is not pointwise but it is an equality in $L^2[a, b]$.

4.2.2 L^2 Spaces on the Unit Circle

An important class of L^2 spaces is the space of measurable complex valued functions $L^2(T) \equiv L^2(T, \nu)$, where T is the unit circle in $\mathbb{C}$ and the Lebesgue measure ν on T is obtained from the arc length divided by 2π . In particular, $\nu(T) = 1$.

Let P be the space of 2π -periodic complex valued functions on $\mathbb{R}$ such that

$$\left(\int_{-\pi}^{\pi} |f|^2 d\mu \right)^{1/2} < \infty ,$$

where μ is the Lebesgue measure on $[-\pi, \pi]$ divided by 2π . If the norm on P is defined by

$$\|f\|_2 = \left(\int_{-\pi}^{\pi} |f|^2 d\mu \right)^{1/2} ,$$

we have an isometry between $L^2(T, \nu)$ and P defined by

$$\begin{aligned} Q : L^2(T, \nu) &\rightarrow P \\ f &\mapsto f(e^{xi}) \end{aligned}$$

We have

$$\left(\int_T |f(z)|^2 d\nu(z) \right)^{1/2} = \left(\int_{-\pi}^{\pi} |f(e^{xi})|^2 d\mu(x) \right)^{1/2},$$

Every measurable set $B \subset T$ is of the form $B = e^{Ai} \equiv \{e^{xi} : x \in A\}$ where $A \subset [-\pi, \pi]$ is a measurable set, and vice-versa. We also have that $\nu(e^{Ai}) = \mu(A)$.

The following definition is justified by the isometry above.

Definition 4.2.2

P is denoted $L^2(T) = L^2(T, \nu)$.

The context will determine which of the two spaces we may be working with.

The set $\{e^{nxi} : n \in \mathbb{Z}\}$ is a complete orthonormal set in $L^2(T)$. The **classical Fourier expansion** of a complex valued function $f \in L^2(T)$ is

$$\sum_{n \in \mathbb{Z}} c_n e^{nxi},$$

where

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-nxi} dx, \quad n \in \mathbb{Z}. \quad (4.2.6)$$

4.2.3 Pointwise Convergence of the Classical Fourier Series

In this section, we prove the following theorem.

Theorem 4.2.3

Let $f : \mathbb{R} \rightarrow \mathbb{C}$ be a 2π -periodic, piecewise differentiable function. Then

$$\frac{f(x^+) + f(x^-)}{2} = \lim_{N \rightarrow \infty} \sum_{n=-N}^N c_n e^{nxi},$$

where the c_n are defined in (4.2.6), $f(x^+) = \lim_{s \rightarrow x^+} f(s)$ and $f(x^-) = \lim_{s \rightarrow x^-} f(s)$.

Less restrictive conditions are given in [29] for the pointwise convergence of the Fourier series. Recall that f is **piecewise differentiable** on $[a, b]$ if f and f' are piecewise continuous on $[a, b]$.

Before proving this theorem, we will need a very famous Lemma.

Lemma 4.2.4 (Riemann-Lebesgue)

If $f : [a, b] \rightarrow \mathbb{R}$ is piecewise differentiable, then

$$\lim_{R \rightarrow \infty} \int_a^b f(y) \sin(Ry) dy = 0 .$$

Proof.

Let $a < y_1 < y_2 < \dots < y_{m-1} < b$ be the points where f is not differentiable. Let $y_0 = a$ and $y_m = b$. Using integration by parts, we have

$$\begin{aligned} \int_a^b f(y) \sin(Ry) dy &= \sum_{j=0}^{m-1} \int_{y_j}^{y_{j+1}} f(y) \sin(Ry) dy \\ &= \frac{1}{R} \sum_{j=0}^{m-1} \left(-f(y_{j+1}^+) \cos(Ry_{j+1}) + f(y_j^-) \cos(Ry_j) + \int_{y_j}^{y_{j+1}} f'(y) \cos(Ry) dy \right) . \end{aligned}$$

Let $M_1 = \sup_{y \in [a, b]} |f(y)|$ and $M_2 = \sup_{y \in [a, b]} |f'(y)|$. M_1 and M_2 are finite because f is piecewise differentiable and thus also piecewise continuous. We have

$$\begin{aligned} \left| \int_a^b f(y) \sin(Ry) dy \right| &\leq \frac{1}{R} \sum_{j=0}^{m-1} \left(|f(y_{j+1}^+)| + |f(y_j^-)| + \int_{y_j}^{y_{j+1}} |f'(y)| dy \right) \\ &\leq \frac{1}{R} (2mM_1 + M_2(b-a)) \rightarrow 0 \quad \text{as } R \rightarrow \infty . \quad \blacksquare \end{aligned}$$

Let

$$S_N(x) = \sum_{n=-N}^N c_n e^{nxi}$$

for N a positive integer and $x \in \mathbb{R}$. We get from (4.2.6) that

$$S_N(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \sum_{n=-N}^N e^{n(x-t)i} dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) D_N(x-t) dt ,$$

where

$$D_N(y) = \sum_{n=-N}^N e^{nyi} .$$

D_N is called the **Dirichlet kernel**. From

$$(1 - e^{yi}) \sum_{n=-N}^N e^{nyi} = \sum_{n=-N}^N (e^{nyi} - e^{(n+1)yi}) = e^{-Nyi} - e^{(N+1)yi} ,$$

we get

$$D_N(y) = \frac{e^{-Nyi} - e^{(N+1)yi}}{1 - e^{yi}} = \frac{e^{-(N+1/2)yi} - e^{(N+1/2)yi}}{e^{-(y/2)i} - e^{(y/2)i}} = \frac{\sin((N+1/2)y)}{\sin(y/2)}$$

for $y \neq 2n\pi$ with $n \in \mathbb{Z}$.

The function $D_N(y)$ is even and $\int_{-\pi}^{\pi} D_N(y) dy = 2\pi$.

Proof of Theorem 4.2.3.

Given $x \in [-\pi, \pi]$.

$$\begin{aligned} S_N(x) - \frac{f(x^+) - f(x^-)}{2} &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) D_N(x-t) dt - \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{f(x^+) + f(x^-)}{2} D_N(t) dt \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x-t) D_N(t) dt - \frac{1}{2\pi} \int_0^{\pi} (f(x^+) + f(x^-)) D_N(t) dt , \end{aligned}$$

where we have used the substitution $t \rightarrow x-t$ in the first integral and the fact that f and D_N are 2π -periodic. For the second integral, we have use the fact that D_N is an even function. Thus

$$\begin{aligned} S_N(x) - \frac{f(x^+) - f(x^-)}{2} &= \frac{1}{2\pi} \int_0^{\pi} f(x-t) D_N(t) dt - \frac{1}{2\pi} \int_0^{\pi} f(x^-) D_N(t) dt \\ &\quad + \frac{1}{2\pi} \int_{-\pi}^0 f(x-t) D_N(t) dt - \frac{1}{2\pi} \int_0^{\pi} f(x^+) D_N(t) dt \\ &= \frac{1}{2\pi} \int_0^{\pi} f(x-t) D_N(t) dt - \frac{1}{2\pi} \int_0^{\pi} f(x^-) D_N(t) dt \\ &\quad + \frac{1}{2\pi} \int_0^{\pi} f(x+t) D_N(t) dt - \frac{1}{2\pi} \int_0^{\pi} f(x^+) D_N(t) dt \\ &= \frac{1}{2\pi} \int_0^{\pi} (f(x-t) - f(x^-)) D_N(t) dt + \frac{1}{2\pi} \int_0^{\pi} (f(x+t) - f(x^+)) D_N(t) dt . \end{aligned}$$

Let

$$I_1(N) = \frac{1}{2\pi} \int_0^{\pi} (f(x-t) - f(x^-)) D_N(t) dt$$

and

$$I_2(N) = \frac{1}{2\pi} \int_0^{\pi} (f(x+t) - f(x^+)) D_N(t) dt .$$

i) We prove that $I_2(N) \rightarrow 0$ as $N \rightarrow \infty$. Namely, given $\epsilon > 0$, we find $N_0 > 0$ such that $|I_2(N)| < \epsilon$ for $N > N_0$.

Given $0 < \delta < \pi$, we have that

$$I_2(N) = I_{2,1}(N, \delta) + I_{2,2}(N, \delta) ,$$

where

$$I_{2,1}(M, \delta) = \frac{1}{2\pi} \int_0^{\delta} (f(x+t) - f(x^+)) D_N(t) dt$$

and

$$I_{2,2}(N, \delta) = \frac{1}{2\pi} \int_{\delta}^{\pi} (f(x+t) - f(x^+)) D_N(t) dt .$$

Choose $\delta_0 \in]0, \pi[$ such that f' is continuous on $]x, x + \delta_0[$. Let $M = \sup_{t \in]x, x + \delta_0[} |f'(t)|$. By the Mean Value Theorem, we have that $|f(x+t) - f(x^+)| \leq Mt$ for $0 < t < \delta_0$. We also have $0 \leq t \leq \pi \sin(t/2)$ for $0 \leq t \leq \pi$. Hence, for $0 < \delta < \min(\delta_0, \epsilon/M)$, we get

$$|I_{2,1}(N, \delta)| \leq \frac{1}{2\pi} \int_0^\delta Mt |D_N(t)| dt \leq \frac{M}{2} \int_0^\delta \frac{t}{\pi \sin(t/2)} dt \leq \frac{M\delta}{2} < \frac{\epsilon}{2}$$

We fix δ between 0 and $\min(\delta_0, \epsilon/M)$. From Riemann-Lebesgue's Lemma, we can find $N_0 > 0$ such that

$$|I_{2,2}(N, \delta)| = \left| \frac{1}{2\pi} \int_\delta^\pi \left(\frac{f(x+t) - f(x^+)}{\sin(t/2)} \right) \sin((N+1/2)t) dt \right| < \frac{\epsilon}{2}$$

for $N > N_0$. Hence, for $N > N_0$, we have $|I_2(N)| < \epsilon$.

We had to split the integral $I_2(N)$ because $(f(x+t) - f(x^+))/\sin(t/2)$ or its derivative may not be bounded at the origin. So, $(f(x+t) - f(x^+))/\sin(t/2)$ may not be piecewise differentiable on $[0, \pi]$.

ii) A proof similar to the one given in (i) yields $I_1(N) \rightarrow 0$ as $N \rightarrow \infty$. ■

4.3 Classical Fourier Series of real Valued Functions

We consider the space $L^2[-L, L] = L^2([-L, L], m)$, where m is the Lebesgue measure on the interval $[-L, L]$. A complete orthogonal set for this space is given by the trigonometric functions

$$\left\{ \cos\left(\frac{n\pi x}{L}\right) : n \geq 0 \right\} \cup \left\{ \sin\left(\frac{n\pi x}{L}\right) : n \geq 1 \right\}.$$

To be more precise, we have

$$\begin{aligned} \int_{-L}^L \sin\left(\frac{i\pi x}{L}\right) \sin\left(\frac{j\pi x}{L}\right) dx &= \begin{cases} 0 & \text{if } i \neq j \text{ and } i, j > 0 \\ L & \text{if } i = j > 0 \end{cases} \\ \int_{-L}^L \cos\left(\frac{i\pi x}{L}\right) \cos\left(\frac{j\pi x}{L}\right) dx &= \begin{cases} 0 & \text{if } i \neq j \text{ and } i, j \geq 0 \\ L & \text{if } i = j > 0 \\ 2L & \text{if } i = j = 0 \end{cases} \end{aligned}$$

and

$$\int_{-L}^L \sin\left(\frac{i\pi x}{L}\right) \cos\left(\frac{j\pi x}{L}\right) dx = 0 \quad \text{if } i, j \geq 0.$$

From (4.2.5), we get the **classical Fourier expansion** of a function $f \in L^2[-L, L]$.

$$a_0 + \sum_{n=1}^{\infty} \left(a_n \cos\left(\frac{n\pi x}{L}\right) + b_n \sin\left(\frac{n\pi x}{L}\right) \right),$$

where $a_0 = \frac{1}{2L} \int_{-L}^L f(x) dx$, $a_n = \frac{1}{L} \int_{-L}^L f(x) \cos\left(\frac{n\pi x}{L}\right) dx$ for $n \geq 1$, and $b_n = \frac{1}{L} \int_{-L}^L f(x) \sin\left(\frac{n\pi x}{L}\right) dx$ for $n \geq 1$.

If $f : \mathbb{R} \rightarrow \mathbb{R}$ is periodic of period $T = 2L$ and $f|_{[-L, L]} \in L^2[-L, L]$, the coefficients of the Fourier expansion are also given by $a_0 = \frac{1}{2L} \int_{\alpha}^{\alpha+2L} f(x) dx$, $a_n = \frac{1}{L} \int_{\alpha}^{\alpha+2L} f(x) \cos\left(\frac{n\pi x}{L}\right) dx$ for $n \geq 1$ and $b_n = \frac{1}{L} \int_{\alpha}^{\alpha+2L} f(x) \sin\left(\frac{n\pi x}{L}\right) dx$ for $n \geq 1$, where α may be any real number.

Example 4.3.1

Find the Fourier series of the periodic function f of period 2π defined by

$$f(x) = \begin{cases} 0 & \text{if } -\pi \leq x \leq 0 \\ 1 & \text{if } 0 < x < \pi \end{cases}$$

We have

$$a_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) dx = \frac{1}{2\pi} \int_0^{\pi} dx = \frac{1}{2},$$

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(nx) dx = \frac{1}{\pi} \int_0^{\pi} \cos(nx) dx = \frac{\sin(nx)}{n\pi} \Big|_0^{\pi} = 0$$

and

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(nx) dx = \frac{1}{\pi} \int_0^{\pi} \sin(nx) dx = -\frac{\cos(nx)}{n\pi} \Big|_0^{\pi} = \begin{cases} 2/(n\pi) & \text{if } n \text{ is odd} \\ 0 & \text{if } n \text{ is even} \end{cases}$$

Hence,

$$f = \frac{1}{2} + \sum_{m=1}^{\infty} \frac{2}{(2m-1)\pi} \sin((2m-1)x).$$

As predicted by Theorem 4.2.3, we have that the value of the series is $1/2$ when $x = n\pi$ with $n \in \mathbb{Z}$. ♣

Example 4.3.2

Find the Fourier series of the periodic function f of period 6 defined by

$$f(x) = \begin{cases} 1 & \text{if } -3 \leq x \leq 0 \\ x & \text{if } 0 < x < 3 \end{cases}$$

We have

$$a_0 = \frac{1}{6} \int_{-3}^3 f(x) dx = \frac{1}{6} \int_{-3}^0 dx + \frac{1}{6} \int_0^3 x dx = \frac{5}{4},$$

$$a_n = \frac{1}{3} \int_{-3}^3 f(x) \cos\left(\frac{n\pi x}{3}\right) dx = \frac{1}{3} \int_{-3}^0 \cos\left(\frac{n\pi x}{3}\right) dx + \frac{1}{3} \int_0^3 x \cos\left(\frac{n\pi x}{3}\right) dx$$

$$\begin{aligned}
&= \frac{1}{n\pi} \sin\left(\frac{n\pi x}{3}\right) \Big|_{-3}^0 + \left(\frac{x}{n\pi} \sin\left(\frac{n\pi x}{3}\right) + \frac{3}{n^2\pi^2} \cos\left(\frac{n\pi x}{3}\right) \right) \Big|_0^3 \\
&= \frac{3}{n^2\pi^2} (\cos(n\pi) - 1) = \begin{cases} 0 & \text{if } n \text{ is even} \\ -6/(n^2\pi^2) & \text{if } n \text{ is odd} \end{cases}
\end{aligned}$$

and

$$\begin{aligned}
b_n &= \frac{1}{3} \int_{-3}^3 f(x) \sin\left(\frac{n\pi x}{3}\right) dx = \frac{1}{3} \int_{-3}^0 \sin\left(\frac{n\pi x}{3}\right) dx + \frac{1}{3} \int_0^3 x \sin\left(\frac{n\pi x}{3}\right) dx \\
&= -\frac{1}{n\pi} \cos\left(\frac{n\pi x}{3}\right) \Big|_{-3}^0 + \left(-\frac{x}{n\pi} \cos\left(\frac{n\pi x}{3}\right) + \frac{3}{n^2\pi^2} \sin\left(\frac{n\pi x}{3}\right) \right) \Big|_0^3 \\
&= \left(-\frac{1}{n\pi} + \frac{1}{n\pi} \cos(n\pi) \right) - \frac{3}{n\pi} \cos(n\pi) = \begin{cases} 1/(n\pi) & \text{if } n \text{ is odd} \\ -3/(n\pi) & \text{if } n \text{ is even} \end{cases}
\end{aligned}$$

Hence,

$$\begin{aligned}
f &= \frac{5}{4} - \sum_{m=1}^{\infty} \frac{6}{(2m-1)^2\pi^2} \cos\left(\frac{(2m-1)\pi x}{3}\right) - \sum_{m=1}^{\infty} \frac{3}{2m\pi} \sin\left(\frac{2m\pi x}{3}\right) \\
&\quad + \sum_{m=1}^{\infty} \frac{1}{(2m-1)\pi} \sin\left(\frac{(2m-1)\pi x}{3}\right).
\end{aligned}$$

From Theorem 4.2.3, we have that

$$\frac{1}{2} = \frac{5}{4} - \sum_{m=1}^{\infty} \frac{6}{(2m-1)^2\pi^2}.$$

This is the value of the series at $x = 6n$ with $n \in \mathbb{Z}$. Moreover,

$$2 = \frac{5}{4} + \sum_{m=1}^{\infty} \frac{6}{(2m-1)^2\pi^2}.$$

This is the value of the series at $x = 3 + 6n$ with $n \in \mathbb{Z}$. ♣

4.3.1 Odd and Even Functions

We may use the properties that $\int_{-L}^L g(t) dt = 0$ if g is an odd function and $\int_{-L}^L g(t) dt = 2 \int_0^L g(t) dt$ if g is an even function to obtain simpler formulae for the Fourier series of odd and even functions.

If $f \in L^2[-L, L]$ is an even function, its Fourier expansion is

$$a_0 + \sum_{n=1}^{\infty} a_n \cos\left(\frac{n\pi x}{L}\right), \tag{4.3.1}$$

where $a_0 = \frac{1}{L} \int_0^L f(x) dx$ and $a_n = \frac{2}{L} \int_0^L f(x) \cos\left(\frac{n\pi x}{L}\right) dx$ for $n \geq 1$.

If $f \in L^2[-L, L]$ is an odd function, its Fourier expansion is

$$\sum_{n=1}^{\infty} b_n \sin\left(\frac{n\pi x}{L}\right), \quad (4.3.2)$$

where $b_n = \frac{2}{L} \int_0^L f(x) \sin\left(\frac{n\pi x}{L}\right) dx$ for $n \geq 1$.

The **Fourier sine series** of a function $f : [0, L] \rightarrow \mathbb{R}$ is the Fourier series of the odd function $F : [-L, L] \rightarrow \mathbb{R}$ defined by

$$F(x) = \begin{cases} f(x) & \text{if } 0 \leq x \leq L \\ -f(-x) & \text{if } -L < x < 0 \end{cases}$$

if $F \in L^2[-L, L]$. The **Fourier cosine series** of a function $f : [0, L] \rightarrow \mathbb{R}$ is the Fourier series of the even function $F : [-L, L] \rightarrow \mathbb{R}$ defined by

$$F(x) = \begin{cases} f(x) & \text{if } 0 \leq x \leq L \\ f(-x) & \text{if } -L \leq x < 0 \end{cases}$$

if $F \in L^2[-L, L]$.

The Fourier series in (4.3.1) and (4.3.1) converge to $f = F|_{[0, L]}$ in $L^2[0, L]$ because they converge to F in $L^2[-L, L]$.

Example 4.3.3

Find the Fourier series of the periodic function f of period π defined by $f(x) = |\sin(x)|$ whose graph is given in Figure 4.1.

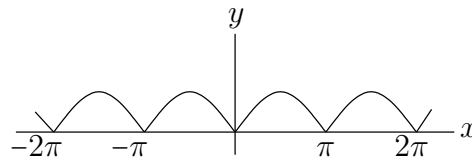


Figure 4.1: Graph of the π -period function define by $f(x) = |\sin(x)|$

We note that f is an even function. Thus, the coefficients b_n of the Fourier series are all null. We have $L = \pi/2$ in the formulae for the a_n . Hence

$$a_0 = \frac{2}{\pi} \int_0^{\pi/2} f(x) dx = \frac{2}{\pi} \int_0^{\pi/2} \sin(x) dx = -\frac{2}{\pi} \cos(x) \Big|_0^{\pi/2} = \frac{2}{\pi}$$

and

$$\begin{aligned}
 a_n &= \frac{4}{\pi} \int_0^{\pi/2} f(x) \cos(2nx) \, dx = \frac{4}{\pi} \int_0^{\pi/2} \sin(x) \cos(2nx) \, dx \\
 &= \frac{2}{\pi} \int_0^{\pi/2} (\sin((2n+1)x) - \sin((2n-1)x)) \, dx \\
 &= \frac{2}{\pi} \left(-\frac{\cos((2n+1)x)}{2n+1} + \frac{\cos((2n-1)x)}{2n-1} \right) \Big|_0^{\pi/2} = \frac{4}{(1-4n^2)\pi} .
 \end{aligned}$$

We get

$$f = \frac{2}{\pi} + \sum_{n=1}^{\infty} \frac{4}{(1-4n^2)\pi} \cos(2nx) .$$

♣

Example 4.3.4

Find the Fourier cosine series of the function $f(x) = e^x$ for $0 \leq x < 1$.

In other words, we want the Fourier series of the even function F of period 2 defined by

$$F(x) = \begin{cases} e^x & \text{if } 0 \leq x \leq 1 \\ e^{-x} & \text{if } -1 \leq x < 0 \end{cases}$$

Since F is an even function, the coefficients b_n of the Fourier series are all null. We have $L = 1$ in the formulae for the a_n . Hence

$$a_0 = \int_0^1 f(x) \, dx = \int_0^1 e^x \, dx = e - 1$$

and

$$\begin{aligned}
 a_n &= 2 \int_0^1 f(x) \cos(n\pi x) \, dx = 2 \int_0^1 e^x \cos(n\pi x) \, dx \\
 &= \frac{2(n\pi)^2}{(n\pi)^2 + 1} \left(e^x \frac{\sin(n\pi x)}{n\pi} + e^x \frac{\cos(n\pi x)}{(n\pi)^2} \right) \Big|_0^1 = \frac{2}{(n\pi)^2 + 1} (e \cos(n\pi) - 1) \\
 &= \begin{cases} \frac{2(e-1)}{(n\pi)^2 + 1} & \text{if } n \text{ is even} \\ \frac{-2(e+1)}{(n\pi)^2 + 1} & \text{if } n \text{ is odd} \end{cases}
 \end{aligned}$$

We get

$$e^x = e - 1 + \sum_{m=1}^{\infty} \frac{2(e-1)}{(2m\pi)^2 + 1} \cos(2m\pi x) - \sum_{m=1}^{\infty} \frac{2(e+1)}{((2m-1)\pi)^2 + 1} \cos((2m-1)\pi x)$$

for $0 \leq x < 1$.

♣

4.3.2 More Special Cases

By expanding the definition of a function $f : [0, L] \rightarrow \mathbb{R}$ to an interval larger than $[0, L]$, it is possible to obtain different Fourier series of f on $[0, L]$ in terms of the sine and cosine functions.

For instance, if we expand a function $f \in L^2[0, L]$ to an even function $F[-2L, 2L] \rightarrow \mathbb{R}$ as follows

$$F(x) = \begin{cases} -f(2L - x) & \text{if } L < x \leq 2L \\ f(x) & \text{if } 0 \leq x \leq L \\ f(-x) & \text{if } -L \leq x < 0 \\ -f(2L + x) & \text{if } -2L \leq x < -L \end{cases}$$

then $F \in L^2[-2L, 2L]$ and the Fourier cosine series of F on $[0, 2L]$ restricted to $[0, L]$ gives the following series representation of f on $[0, L]$.

$$\sum_{n=1}^{\infty} a_n \cos\left(\frac{(2n-1)\pi x}{2L}\right) \quad (4.3.3)$$

for $x \in [0, L]$, where $a_n = \frac{2}{L} \int_0^L f(x) \cos\left(\frac{(2n-1)\pi x}{2L}\right) dx$ for $n \geq 1$. To obtain this formula, we need to note that

$$\cos\left(\frac{m\pi x}{2L}\right) = \begin{cases} \cos\left(\frac{m\pi(2L-x)}{2L}\right) & \text{if } m \text{ even} \\ -\cos\left(\frac{m\pi(2L-x)}{2L}\right) & \text{if } m \text{ odd} \end{cases}$$

for $L \leq x \leq 2L$. Hence,

$$F(x) \cos\left(\frac{m\pi x}{2L}\right) = F(x) \cos\left(\frac{m\pi(2L-x)}{2L}\right) = -F(2L-x) \cos\left(\frac{m\pi(2L-x)}{2L}\right)$$

for m even and $L \leq x \leq 2L$, and

$$F(x) \cos\left(\frac{m\pi x}{2L}\right) = -F(x) \cos\left(\frac{m\pi(2L-x)}{2L}\right) = F(2L-x) \cos\left(\frac{m\pi(2L-x)}{2L}\right)$$

for m odd and $L \leq x \leq 2L$. We then get that

$$a_m = \frac{2}{2L} \int_0^{2L} f(x) \cos\left(\frac{m\pi x}{2L}\right) dx = \begin{cases} \frac{2}{L} \int_0^L f(x) \cos\left(\frac{m\pi x}{2L}\right) dx & \text{if } m \text{ odd} \\ 0 & \text{if } m \text{ even} \end{cases}$$

If we expand a function $f \in L^2[0, L]$ to an odd function $F[-2L, 2L] \rightarrow \mathbb{R}$ as follows

$$F(x) = \begin{cases} f(2L - x) & \text{if } L < x \leq 2L \\ f(x) & \text{if } 0 \leq x \leq L \\ -f(-x) & \text{if } -L \leq x < 0 \\ -f(2L + x) & \text{if } -2L \leq x < -L \end{cases}$$

then $F \in L^2[-2L, 2L]$ and the Fourier sine series of F on $[0, 2L]$ restricted to $[0, L]$ gives the following series representation of f on $[0, L]$.

$$\sum_{n=1}^{\infty} b_n \sin\left(\frac{(2n-1)\pi x}{2L}\right) \quad (4.3.4)$$

for $x \in [0, L]$, where $b_n = \frac{2}{L} \int_0^L f(x) \sin\left(\frac{(2n-1)\pi x}{2L}\right) dx$ for $n \geq 1$. One may proceed as we did above to obtain this results.

The Fourier series in (4.3.3) and (4.3.4) converge to $f = F|_{[0,L]}$ in $L^2[0, L]$ because they converge to F in $L^2[-2L, 2L]$.

4.3.3 Fourier Series of Functions of Two Variables

If, in the general definition of L^2 spaces, we have $\Omega = [a, b] \times [b, c]$ and μ is the Lebesgue measure on Ω , then we get the space $L^2([a, b] \times [b, c]) = L^2([a, b] \times [b, c], \mu)$ of L^2 integrable functions on the rectangle $[a, b] \times [b, c]$. We invite the reader to write the definitions of scalar product, norm, orthogonal set, complete set, ... in the special case of $L^2([a, b] \times [c, d])$.

We can extend in a natural way the concept of classical Fourier series to functions in $L^2([a, b] \times [b, c])$. For instance, the Fourier sine series for a function $f \in L^2([0, a] \times [0, b])$ is

$$\sum_{m=1}^{\infty} \sum_{n=1}^{\infty} c_{m,n} \sin\left(\frac{m\pi x}{a}\right) \sin\left(\frac{n\pi y}{b}\right),$$

where $c_{m,n} = \frac{4}{ab} \int_0^b \int_0^a f(x, y) \sin\left(\frac{m\pi x}{a}\right) \sin\left(\frac{n\pi y}{b}\right) dx dy$ for $n, m \geq 1$.

We leave it to the reader to provide other examples of classical Fourier series in $L^2([a, b] \times [c, d])$.

4.4 Operator Theory

4.4.1 Bounded Operators

We start by reviewing some definitions about mapping between Banach spaces. In this section, X and Y are two Banach spaces over the complex numbers.

Definition 4.4.1

A **linear operator** or **linear mapping** from X to Y is a function $A : X \rightarrow Y$ such that $A(\lambda_1 x_1 + \lambda_2 x_2) = \lambda_1 A(x_1) + \lambda_2 A(x_2)$ for all $\lambda_1, \lambda_2 \in \mathbb{C}$ and $x_1, x_2 \in X$. If $Y = \mathbb{C}$, then the linear operator A is called a **linear functional**.

The **range** of A is the linear space $\mathcal{R}(A) = \{y \in Y : y = A(x) \text{ for some } x \in X\}$. The **kernel** of A is the linear space $\mathcal{N}(A) = \{x \in X : A(x) = 0\}$.

Definition 4.4.2

A linear operator $A : X \rightarrow Y$ is **bounded** if there exist a constant $C \geq 0$ such that $\|A(x)\| \leq C\|x\|$ for all $x \in X$. The space of bounded linear operators from X to Y is denoted $\mathcal{L}(X, Y)$. If $Y = X$, The space of bounded linear operators on X is simply denoted $\mathcal{L}(X)$.

The following results are fundamental results about bounded linear operators between Banach spaces.

Proposition 4.4.3

Let

$$\|A\| = \sup_{\substack{x \in X \\ \|x\|=1}} \|A(x)\|, \quad A \in \mathcal{L}(X, Y).$$

1. $\|\cdot\|$ defines a norm on $\mathcal{L}(X, Y)$ called the **induce norm from X and Y** .
2. Equipped with the norm $\|\cdot\|$, $\mathcal{L}(X, Y)$ is a Banach space.

Theorem 4.4.4 (Open Mapping Theorem)

If $A \in \mathcal{L}(X, Y)$ is onto, then the image of an open set is an open set.

Theorem 4.4.5 (Bounded Inverse Theorem)

If $A \in \mathcal{L}(X, Y)$ is one-to-one and onto, then $A^{-1} \in \mathcal{L}(Y, X)$.

Definition 4.4.6

The **graph** of a linear operator $A : X \rightarrow Y$ is the subspace

$$\mathcal{G}(A) = \{(x, A(x)) : x \in X\} \subset X \times Y.$$

A linear operator $A : X \rightarrow Y$ is **closed** if its graph is closed in $X \times Y$. In other words, a linear operator $A : X \rightarrow Y$ is closed if, for any converging sequence $\{(x_n, A(x_n))\}_{n=0}^{\infty} \subset \mathcal{G}(A)$, we have that $(x, y) = \lim_{n \rightarrow \infty} (x_n, A(x_n)) \in \mathcal{G}(A)$; namely, $y = A(x)$.

Theorem 4.4.7 (Closed Graph Theorem)

If $A : X \rightarrow Y$ is a closed linear operator, then $A \in \mathcal{L}(X, Y)$.

Definition 4.4.8

The **resolvent set** $\rho(A)$ of $A \in \mathcal{L}(X)$ is the set of all $\lambda \in \mathbb{C}$ such that $A - \lambda \text{Id} : X \rightarrow X$ is one-to-one and onto. The **spectrum** of $A \in \mathcal{L}(X)$ is the set $\sigma(A) = \mathbb{C} \setminus \rho(A)$.

From the Open Mapping Theorem, we have that $(A - \lambda \text{Id})^{-1} \in \mathcal{L}(X)$ if λ is in the resolvent set of A .

Definition 4.4.9

If λ is in the resolvent set of A , then $R_\lambda(A) = (A - \lambda \text{Id})^{-1}$ is called the **resolvent** of A .

$\lambda \in \sigma(A)$ if $A - \lambda \text{Id}$ is not one-to-one, or $A - \lambda \text{Id}$ is not onto, or both.

Definition 4.4.10

If $A - \lambda \text{Id}$ is not one-to-one, then λ is called an **eigenvalue** of A . We have that $\mathcal{N}(A - \lambda \text{Id}) \neq \{0\}$ and the elements of $\mathcal{N}(A - \lambda \text{Id}) \setminus \{0\}$ are called the **eigenvectors** associated to λ . $\mathcal{N}(A - \lambda \text{Id})$ is called the **eigenspace** associated to λ .

If x is an eigenvector associated to λ , then $Ax = \lambda x$.

Proposition 4.4.11

If $A \in \mathcal{L}(X)$ and $\lambda \in \sigma(A)$, then $|\lambda| \leq \|A\|$.

Definition 4.4.12

The **spectral radius** of $A \in \mathcal{L}(X)$ is defined as $r_\sigma(A) = \sup_{\lambda \in \sigma(A)} |\lambda|$.

From the previous proposition, $r_\sigma(A) \leq \|A\|$.

We list below some of the important properties of the spectrum of a bounded linear operator.

Proposition 4.4.13

If $A \in \mathcal{L}(X)$, then $\sigma(A) \subset \mathbb{C}$ is compact.

Proposition 4.4.14

If $A \in \mathcal{L}(X)$ and $X \neq \{0\}$, then $\sigma(A) \neq \emptyset$.

The following definition generalizes the definition of dual space given in Definition 4.1.8.

Definition 4.4.15

Let X be a Banach space. The **dual space** X^* of X is the set of all bounded linear functionals on X .

Example 4.4.16

If H is an Hilbert space, we have from the Riesz Representation Theorem, Theorem 4.1.9, that there exists a mapping $T : H^* \rightarrow H$ such that $u(x) = \langle x, T(u) \rangle$ for all $x \in H$. It follows

from Schwarz inequality that $|u(x)| \leq \|x\| \|T(u)\|$ for all $x \in H$. Thus, $\|u\| \leq \|T(u)\|$ where the norm on the left is the norm of the linear mapping $u : H \rightarrow \mathbb{R}$ and the norm on the right is the norm of $T(u) \in H$. In fact, we have that $\|u\| = \|T(u)\|$ because $u(\|T(u)\|^{-1}T(u)) = \|T(u)\|$. Hence, $T : H^* \rightarrow H$ is an isometry from H^* into H . ♣

Definition 4.4.17

The **adjoint** of $A \in \mathcal{L}(X, Y)$ is the linear mapping $A^* : Y^* \rightarrow X^*$ defined by

$$A^*(w)(x) = w(A(x)) \quad , \quad x \in X \text{ and } w \in Y^* .$$

Theorem 4.4.18

If $A \in \mathcal{L}(X, Y)$, then $A^* \in \mathcal{L}(Y^*, X^*)$ and $\|A\| = \|A^*\|$

Obviously, in the previous theorem,

$$\|A^*\| = \sup_{\substack{w \in Y^* \\ \|w\|=1}} \|A^*(w)\| \quad , \quad A^* \in \mathcal{L}(Y^*, X^*) ,$$

and

$$\|w\| = \sup_{\substack{y \in Y \\ \|y\|=1}} |w(y)| \quad , \quad w \in Y^* ,$$

with a similar definition for the norm on X^* .

Definition 4.4.19

A linear operator $A \in \mathcal{L}(X, Y)$ is **compact** if $\overline{A(V)}$ is compact for every bounded set $V \subset X$. The set of all compact linear operators from X to Y is denoted $\mathcal{K}(X, Y)$ and the set of all compact linear operators from X to itself is denoted $\mathcal{K}(X)$.

A little review of some topological concepts are useful for future applications.

Definition 4.4.20

A set $U \subset X$ is **precompact** if $\overline{U}$ is compact. A set $U \subset X$ is **sequentially compact** if every sequence in U has a subsequence converging in X .

Inside a Banach space, these two concepts are equivalent.

Proposition 4.4.21

$\mathcal{K}(X, Y)$ is a closed subspace of $\mathcal{L}(X, Y)$ with respect to the operator norm induced by the norms on X and Y .

Proposition 4.4.22

Let X , Y and Z be three Banach spaces.

1. If $A \in \mathcal{K}(X, Y)$ and $B \in \mathcal{L}(Y, Z)$, then $B \circ A \in \mathcal{K}(X, Z)$.
2. If $A \in \mathcal{L}(X, Y)$ and $B \in \mathcal{K}(Y, Z)$, then $B \circ A \in \mathcal{K}(X, Z)$.

Proof.

We prove (2). Let $U \subset X$ be a bounded set, then $A(U) \subset Y$ is bounded because A is bounded¹. Thus, $\overline{B(A(U))}$ is a compact set since B is a compact operator and $A(U)$ is a bounded set. ■

Theorem 4.4.23

Let A be a linear operator from X to Y .

1. If $A \in \mathcal{L}(X, Y)$ and $\dim \mathcal{R}(A) < \infty$, then $A \in \mathcal{K}(X, Y)$.
2. If $A \in \mathcal{K}(X, Y)$ and $\mathcal{R}(A)$ is closed, then $\dim \mathcal{R}(A) < \infty$.
3. $A \in \mathcal{K}(X, Y)$ if and only if $A^* \in \mathcal{K}(Y^*, X^*)$.

Surprisingly, the next theorem will be quite useful for the study of differential equations. For instance, it will be used to describe the eigenvalues for the Sturm-Liouville problem and the Laplace operator.

Theorem 4.4.24

Consider $A \in \mathcal{K}(X)$.

1. If $\lambda \neq 0$, then $\dim \mathcal{N}(A - \lambda \text{Id}) < \infty$.
2. If $\lambda \neq 0$, then $\mathcal{R}(A - \lambda \text{Id})$ is closed.
3. If $\lambda \neq 0$, then $\dim \mathcal{N}(A - \lambda \text{Id}) = \dim \mathcal{N}(A^* - \lambda \text{Id})$.
4. If $r > 0$ and E is the set of eigenvalues λ of A such that $|\lambda| > r$, then E is finite and $\mathcal{R}(A - \lambda \text{Id}) \neq X$ for all $\lambda \in E$.
5. If $\lambda \in \sigma(A)$ and $\lambda \neq 0$, then λ is an eigenvalue.
6. $\sigma(A)$ is compact and at most countable. If $\sigma(A)$ is countable, then 0 is the only limit point of $\sigma(A)$.

¹If $B_r(\mathbf{0}) \supset U$, then $B_{\|A\|_r}(\mathbf{0}) \supset A(U)$.

4.4.2 Unbounded Operators

As in the previous section, X and Y are two Banach spaces over the complex numbers. We consider linear operators from a subspace Z of X to Y where Z may not be equal to X . We generalize many of the definitions and results that we have presented in the previous section to these new linear operators.

Definition 4.4.25

A **linear operator** or **linear mapping** from X to Y is a pair $(\mathcal{D}(A), A)$, where $\mathcal{D}(A)$ is a subspace of X and $A : \mathcal{D}(A) \rightarrow Y$ is such that $A(\lambda_1 x_1 + \lambda_2 x_2) = \lambda_1 A(x_1) + \lambda_2 A(x_2)$ for all $\lambda_1, \lambda_2 \in \mathbb{C}$ and $x_1, x_2 \in \mathcal{D}(A)$. $\mathcal{D}(A)$ is called the **domain** of A . The **range** of A is the linear space $\mathcal{R}(A) = \{y \in Y : y = A(x) \text{ for some } x \in \mathcal{D}(A)\}$. The **kernel** of A is the linear space $\mathcal{N}(A) = \{x \in \mathcal{D}(A) : A(x) = 0\}$.

Definition 4.4.26

A linear operator $(\mathcal{D}(A), A)$ is **bounded** if there exists a constant $C \geq 0$ such that $\|A(x)\| \leq C\|x\|$ for all $x \in \mathcal{D}(X)$.

The following result is well know for linear operators between finite dimensional vector spaces.

Proposition 4.4.27

If $(\mathcal{D}(A), A)$ is a linear operator such that $\dim \mathcal{D}(A) < \infty$, then $(\mathcal{D}(A), A)$ is bounded.

Definition 4.4.28

A linear operator $(\mathcal{D}(B), B)$ from X to Y is an **extension** of a linear operator $(\mathcal{D}(A), A)$ from X to Y if $\mathcal{D}(A) \subset \mathcal{D}(B)$ and $A(x) = B(x)$ for all $x \in \mathcal{D}(A)$.

Theorem 4.4.29

Given a linear operator $(\mathcal{D}(A), A)$ such that $\mathcal{D}(A)$ is a closed subspace of X and

$$M = \sup_{\substack{x \in \mathcal{D}(A) \\ \|x\|=1}} \|A(x)\| < \infty, \quad (4.4.1)$$

then there exists a bounded linear operator $B \in \mathcal{L}(X, Y)$ such that (X, B) is an extension of $(\mathcal{D}(A), A)$ and $\|B\| = M$. M defined by (4.4.1) is the norm of A acting on the subspace $\mathcal{D}(A)$.

Definition 4.4.30

The **graph** of a linear operator $(\mathcal{D}(A), A)$ is the subspace

$$\mathcal{G}(A) = \{(x, A(x)) : x \in \mathcal{D}(A)\} \subset X \times Y .$$

Definition 4.4.31

A linear operator $(\mathcal{D}(A), A)$ is **closed** if its graph is closed in $X \times Y$. In other words, a linear operator $(\mathcal{D}(A), A)$ is closed if, for any converging sequence $\{(x_n, A(x_n))\}_{n=0}^{\infty} \subset \mathcal{G}(A)$, we have $(x, y) = \lim_{n \rightarrow \infty} (x_n, A(x_n)) \in \mathcal{G}(A)$; namely, $x \in \mathcal{D}(A)$ and $y = A(x)$.

The **closure** of an operator $(\mathcal{D}(A), A)$ is the smallest closed extension of $(\mathcal{D}(A), A)$ which we denoted by $(\mathcal{D}(\bar{A}), \bar{A})$.

Theorem 4.4.32 (Closed Graph Theorem)

If $(\mathcal{D}(A), A)$ is a closed linear operator such that $\mathcal{D}(A)$ is dense in X , then $(\mathcal{D}(A), A)$ is bounded.

Definition 4.4.33

The **resolvent set** $\rho(A)$ of a closed linear operator $(\mathcal{D}(A), A)$ from X to itself is the set of all $\lambda \in \mathbb{C}$ such that $A - \lambda \text{Id} : \mathcal{D}(A) \rightarrow X$ is one-to-one and onto, and $(A - \lambda \text{Id})^{-1} \in \mathcal{L}(X)$. The **spectrum** of A is the set $\sigma(A) = \mathbb{C} \setminus \rho(A)$. If λ is in the resolvent set of A , then $R_\lambda(A) \equiv (A - \lambda \text{Id})^{-1}$ is called the **resolvent** of A .

Proposition 4.4.34

$\lambda \in \sigma(A)$ falls into one of the following three classes.

1. The **point spectrum** or **discrete spectrum** is the set $\sigma_{ps}(A)$ of all $\lambda \in \sigma(A)$ such that $A - \lambda \text{Id}$ is not one-to-one.
2. The **residual spectrum** or **compression spectrum** is the set $\sigma_{rs}(A)$ of all λ such that $A - \lambda \text{Id}$ is one-to-one but $\mathcal{R}(A - \lambda \text{Id})$ is not a dense subspace of X .
3. The **continuous spectrum** is the set $\sigma(A) \setminus (\sigma_{ps}(A) \cup \sigma_{rs}(A))$.

Definition 4.4.35

If λ is in the point spectrum, then λ is called an **eigenvalue** of A . We have that $\mathcal{N}(A - \lambda \text{Id}) \neq \{0\}$ and the elements of $\mathcal{N}(A - \lambda \text{Id}) \setminus \{0\}$ are called the **eigenvectors** associated to λ . $\mathcal{N}(A - \lambda \text{Id})$ is called the **eigenspace** associated to λ .

If x is an eigenvector associated to λ , then $Ax = \lambda x$. As we have in finite dimension, eigenvectors associated to distinct eigenvalues are independent.

Definition 4.4.36

Let $(\mathcal{D}(A), A)$ be a linear mapping from X into Y such that $\mathcal{D}(A)$ is dense in X . Let S be the set of all bounded linear functionals $w \in Y^*$ such that there exists a bounded linear functional $v \in X^*$ satisfying $w(A(x)) = v(x)$ for all $x \in \mathcal{D}(A)$. Since $\mathcal{D}(A)$ is dense, v is uniquely defined. The **adjoint** of A is the mapping (S, A^*) from Y^* to X^* defined by $A^*(w) = v$ for $w \in S$. We have that $\mathcal{D}(A^*) = S$.

Theorem 4.4.37

Let $(\mathcal{D}(A), A)$ be a linear mapping from X into Y such that $\mathcal{D}(A)$ is dense in X , then $(\mathcal{D}(A^*), A^*)$ is a closed linear operator. Moreover, if X and Y are **reflexive** (i.e. $X^{**} = X$ and $Y^{**} = Y$), then $\mathcal{D}(A^*)$ is dense in Y^* and $A^{**} = A$.

Definition 4.4.38

Let $(\mathcal{D}(A), A)$ be a bounded linear operator from X to Y . The linear operator $(\mathcal{D}(A), A)$ is **compact** if $\overline{A(V)}$ is compact for every bounded set $V \subset \mathcal{D}(A)$.

The next theorem is not surprising if we think about linear mappings between finite dimensional vector spaces.

Theorem 4.4.39

Let $(\mathcal{D}(A), A)$ be a linear operator from X to Y , then

1. $(\mathcal{D}(A), A)$ is a compact linear operator if $(\mathcal{D}(A), A)$ is bounded and $\mathcal{R}(A)$ is finite dimensional.
2. $(\mathcal{D}(A), A)$ is a compact linear operator if $\mathcal{D}(A)$ is finite dimensional.

4.4.3 Operators on Hilbert Spaces

In the present section, $H, H_1, H_2, \dots$ will be Hilbert spaces over the complex numbers.

Obviously, the definitions and results of the previous sections about operators between Banach spaces applied are also valid for operators between Hilbert spaces. However, since there is a conjugate-linear isomorphism between H^* and H provided by the Riesz Representation Theorem, it is traditional to redefine the adjoint of a linear operator between two Hilbert spaces. Recall that a linear mapping L between two vector spaces is **Conjugate-linear** if $L(\lambda_1 v_1 + \lambda_2 v_2) = \overline{\lambda_1} L(v_1) + \overline{\lambda_2} L(v_2)$ for all $\lambda_1, \lambda_2 \in \mathbb{C}$ and all vectors v_1, v_2 .

Definition 4.4.40

Let $(\mathcal{D}(A), A)$ be a linear mapping from H_1 into H_2 such that $\mathcal{D}(A)$ is dense in H_1 . Let $S \subset H_2$ be the set of all elements w of H_2 for which there exists an element $v \in H_1$

satisfying $\langle A(x), w \rangle_{H_2} = \langle x, v \rangle_{H_1}$ for $x \in \mathcal{D}(A)$. Since $\mathcal{D}(A)$ is dense, v is uniquely defined. The **(Hilbert space) adjoint** of A is the mapping $(S, A^\bullet)$ from H_2 to H_1 defined by $A^\bullet(w) = v$ for $w \in S$. We have $\mathcal{D}(A^\bullet) = S$.

Consider the conjugate-linear application

$$\begin{aligned}\Lambda_H : H &\rightarrow H^* \\ v &\mapsto \langle \cdot, v \rangle\end{aligned}$$

As a consequence of Riesz Representation Theorem, this is a one-to-one and onto mapping.

Consider a linear operator $(\mathcal{D}(A), A)$ from H_1 to H_2 such that $\mathcal{D}(A)$ is dense in H_1 . Let $(\mathcal{D}(A^*), A^*)$ be the adjoint of $(\mathcal{D}(A), A)$ in the sense of operators between Banach spaces and $(\mathcal{D}(A^\bullet), A^\bullet)$ be the adjoint of $(\mathcal{D}(A), A)$ in the sense of operators between Hilbert spaces. We have

$$A^* = \Lambda_{H_1} \circ A^\bullet \circ \Lambda_{H_2}^{-1} \quad (4.4.2)$$

on $\mathcal{D}(A^*)$. To prove this, we note that

$$w(A(x)) = A^*(w)(x) \Leftrightarrow \langle A(x), \Lambda_{H_2}^{-1}(w) \rangle_{H_2} = \langle x, \Lambda_{H_1}^{-1}(A^*(w)) \rangle_{H_1}$$

for all $x \in \mathcal{D}(A)$ and $w \in \mathcal{D}(A^*)$. Thus $\Lambda_{H_2}^{-1}(w) \in \mathcal{D}(A^\bullet)$ for $w \in \mathcal{D}(A^*)$ and

$$A^\bullet(\Lambda_{H_2}^{-1}(w)) = \Lambda_{H_1}^{-1}(A^*(w))$$

for $w \in \mathcal{D}(A^*)$ by definition of $(\mathcal{D}(A^\bullet), A^\bullet)$. This proves (4.4.2).

It follows from Definition 4.4.40 that the mapping $A \mapsto A^\bullet$ is conjugate-linear because

$$A^\bullet(w) = v \Leftrightarrow \langle A(x), w \rangle_{H_2} = \langle x, v \rangle_{H_1} \Leftrightarrow \langle \lambda A(x), w \rangle_{H_2} = \langle \lambda x, v \rangle_{H_1} = \langle x, \bar{\lambda} v \rangle_{H_1} \Leftrightarrow (\lambda A)^\bullet = \bar{\lambda} v$$

for all $\lambda \in \mathbb{C}$. We have however that the mapping $A \mapsto A^*$ is linear.

From now on, we will follow the tradition in functional analysis and use $A(\mathcal{D}(A^*), A^*)$ to also designate $A(\mathcal{D}(A^\bullet), A^\bullet)$. The context will determine if we are talking of the dual in the sense of operators between Banach spaces or between Hilbert spaces.

Theorem 4.4.41

Let $(\mathcal{D}(A), A)$ be a linear operator from H to itself such that $\mathcal{D}(A)$ is dense in H . We have $\mathcal{R}(A)^\perp = \mathcal{N}(A^*)$ and $\mathcal{R}(A^*)^\perp = \mathcal{N}(A)$. Moreover, $\mathcal{N}(A^*)^\perp = \overline{\mathcal{R}(A)}$ and $\mathcal{N}(A)^\perp = \overline{\mathcal{R}(A^*)}$.

Definition 4.4.42

Let $(\mathcal{D}(A), A)$ be a linear operator from H into itself such that $\mathcal{D}(A)$ is dense in H . We say that $(\mathcal{D}(A), A)$ is **symmetric** if $(\mathcal{D}(A^*), A^*)$ is an extension of $(\mathcal{D}(A), A)$. Thus, $\langle A(x), y \rangle = \langle x, A(y) \rangle$ for $x, y \in \mathcal{D}(A)$.

We say that $(\mathcal{D}(A), A)$ is **self-adjoint** if $(\mathcal{D}(A^*), A^*) = (\mathcal{D}(A), A)$; namely, if $\mathcal{D}(A) = \mathcal{D}(A^*)$ and $A(x) = A^*(x)$ for all $x \in \mathcal{D}(A)$.

Theorem 4.4.43

Let $(\mathcal{D}(A), A)$ be a symmetric linear operator from H to itself such that $\mathcal{D}(A)$ is dense in H . We have

1. $\langle A(x), x \rangle \in \mathbb{R}$ for all $x \in \mathcal{D}(A)$.
2. The eigenvalues of A are real.
3. Eigenvectors associated to distinct eigenvalues are orthogonal.
4. The continuous spectrum of A is a subset of $\mathbb{R}$.

If $(\mathcal{D}(A), A)$ is also self-adjoint, we have

1. The spectrum of A is a subset of $\mathbb{R}$.
2. The residual spectrum of A is empty.

Consider a self-adjoint linear operator $A \in \mathcal{K}(H)$. According to Theorem 4.4.24, every $\lambda \in \sigma(A)$ with $\lambda \neq 0$ is an eigenvalue with $\dim \mathcal{N}(A - \lambda \text{Id}) < \infty$, and these eigenvalues may be arranged in a monotone decreasing sequence $|\lambda_1| \geq |\lambda_2| \geq |\lambda_3| \geq \dots$. Moreover, if $\sigma(A)$ is infinite, then $\lambda_n \rightarrow 0$ as $n \rightarrow \infty$.

If we choose for each eigenspace $\mathcal{N}(A - \lambda \text{Id})$, where $\lambda \in \sigma(A)$ with $\lambda \neq 0$, an orthonormal basis and do the union of these basis, we get an orthonormal set $\{\phi_j\}_{j=1}^\infty$ such that $\phi_j \neq 0$ for only a finite number of j when $\sigma(A)$ is finite. For each eigenvector ϕ_j , we denote the eigenvalue associate to it by $\tilde{\lambda}_j$. For instance, for each $\phi_j \in \mathcal{N}(A - \lambda_n \text{Id})$, we will have that $\tilde{\lambda}_j = \lambda_n$.

Theorem 4.4.44 (Hilbert-Schmidt Theorem)

Consider a self-adjoint linear operator $A \in \mathcal{K}(H)$. Let $S = \{\phi_1, \phi_2, \phi_3, \dots\}$ be an orthonormal set of eigenvectors as defined in the previous paragraph.

1. S is an orthonormal basis for H .
2. $Au = \sum_{j=1}^\infty \tilde{\lambda}_j \langle u, \phi_j \rangle \phi_j$ for $u \in H$.
3. Given $f \in H$, let $f_0 = f - \sum_{j=1}^\infty \langle f, \phi_j \rangle \phi_j$. We have that $f_0 \in \mathcal{N}(A)$.

If λ is not an eigenvalue, there is a unique solution to $Au - \lambda u = f$ given by

$$u = \sum_{j=1}^\infty \frac{1}{\tilde{\lambda}_j - \lambda} \langle f, \phi_j \rangle \phi_j - \frac{1}{\lambda} f_0 .$$

If λ is an eigenvalue, there is a solution to $Au - \lambda u = f$ if and only if $\langle f, \phi_j \rangle = 0$ for $\tilde{\lambda}_j = \lambda$. If this is satisfied, there are infinitely many solutions given by

$$u = \sum_{\substack{j=1 \\ \tilde{\lambda}_j \neq \lambda}}^{\infty} \frac{1}{\tilde{\lambda}_j - \lambda} \langle f, \phi_j \rangle \phi_j - \frac{1}{\lambda} f_0 + \sum_{\substack{j=1 \\ \tilde{\lambda}_j = \lambda}}^{\infty} c_j \phi_j ,$$

where the c_j 's are arbitrary constants.

Theorem 4.4.45

If $A \in \mathcal{K}(H)$, there exists a sequence $\{A_n\}_{n=1}^{\infty}$ of operators $A_n \in \mathcal{L}(H)$, each of finite rank, converging to A in the induced operator norm; namely, $\lim_{n \rightarrow \infty} \|A_n - A\| = 0$.

We complete this section with an important class of compact operators.

Theorem 4.4.46

Let Ω be a measurable space and $K \in L^2(\Omega \times \Omega)$. Then $T_K : L^2(\Omega) \rightarrow L^2(\Omega)$ defined by

$$(T_K f)(x) = \int_{\Omega} K(x, y) f(y) dy \quad , \quad \text{for } x \in \Omega \text{ and } f \in L^2(\Omega) ,$$

is a compact operator and $\|T_K\| \leq \|K\|_2$.

The function $K \in L^2(\Omega \times \Omega)$ in the previous theorem is called a **Hilbert-Schmidt kernel**.

Proof.

i) We first prove that T_K is effectively a bounded operator from $L^2(\Omega)$ into itself.

Since $K : \Omega \rightarrow \mathbb{R}$ is measurable, we have that $\mathbf{y} \mapsto K(\mathbf{x}, \mathbf{y})$ is also measurable for almost all $\mathbf{x}$. Given any $f \in L^2(\Omega)$, we have that $\mathbf{y} \mapsto K(\mathbf{x}, \mathbf{y})f(\mathbf{y})$ is measurable for almost all $\mathbf{x}$.

From Fubini's Theorem, we have that $\int_{\Omega} |K(\mathbf{x}, \mathbf{y})|^2 d\mathbf{y}$ exists almost everywhere in $\mathbf{x}$ because $K \in L^2(\Omega \times \Omega)$. Hence, given any $f \in L^2(\Omega)$, we have from Schwarz inequality in $L^2(\Omega)$ that

$$\int_{\Omega} |K(\mathbf{x}, \mathbf{y})| |f(\mathbf{y})| d\mathbf{y} \leq \left(\int_{\Omega} |K(\mathbf{x}, \mathbf{y})|^2 d\mathbf{y} \right)^{1/2} \left(\int_{\Omega} |f(\mathbf{y})|^2 d\mathbf{y} \right)^{1/2} < \infty$$

almost everywhere. Hence, $\mathbf{y} \mapsto K(\mathbf{x}, \mathbf{y})f(\mathbf{y})$ is in $L^1(\Omega)$ for almost all $\mathbf{x}$. So, $T_K f : \Omega \rightarrow \mathbb{R}$ is well defined almost everywhere. Moreover, from Fubini's Theory, $T_K f : \Omega \rightarrow \mathbb{R}$ is measurable.

From

$$|(T_K f)(x)| \leq \int_{\Omega} |K(x, y)| |f(y)| dy \leq \left(\int_{\Omega} |K(x, y)|^2 dy \right)^{1/2} \left(\int_{\Omega} |f(y)|^2 dy \right)^{1/2}$$

for almost all x , we get

$$\|T_K f\|_2^2 = \int_{\Omega} |(T_K f)(x)|^2 dx \leq \int_{\Omega} \int_{\Omega} |K(x, y)|^2 dy dx \int_{\Omega} |f(y)|^2 dy = \|K\|_2^2 \|f\|_2^2 < \infty$$

for all $f \in L^2(\Omega)$. Thus $T_K f \in L^2(\Omega)$ for all $f \in L^2(\Omega)$ and

$$\|T_K\| \leq \|K\|_2 = \left(\int_{\Omega} \int_{\Omega} |K(\mathbf{x}, \mathbf{y})|^2 d\mathbf{y} d\mathbf{x} \right)^{1/2}. \quad (4.4.3)$$

ii) We now prove that T_K is compact.

Let $\{\phi_j\}_{j=0}^{\infty}$ be an orthonormal basis of $L^2(\Omega)$. We consider the functions $\psi_{j,k} : \Omega \times \Omega \rightarrow \mathbb{R}$ defined by $\psi_{j,k}(\mathbf{x}, \mathbf{y}) = \phi_j(\mathbf{x})\phi_k(\mathbf{y})$ for $(\mathbf{x}, \mathbf{y}) \in \Omega \times \Omega$. The functions $\psi_{j,k}$ are measurable. Using Fubini's Theorem and Theorem 4.1.13, one can show that $\{\psi_{j,k}\}_{(j,k) \in \mathbb{N} \times \mathbb{N}}$ is an orthonormal basis of $L^2(\Omega \times \Omega)$.

We can write K as $K = \sum_{(j,k) \in \mathbb{N} \times \mathbb{N}} a_{j,k} \psi_{j,k}$ for some $a_{j,k} \in \mathbb{R}$. The convergence of the previous series is in $L^2(\Omega \times \Omega)$. In particular, $\|K\|_2^2 = \sum_{(j,k) \in \mathbb{N} \times \mathbb{N}} a_{j,k}^2$ since $\{\psi_{j,k}\}_{(j,k) \in \mathbb{N} \times \mathbb{N}}$ is an orthonormal basis. The series converges independently of the order of the summation. In particular, for every $\epsilon > 0$, there exists $M > 0$ such that $\sum_{i+j > m} a_{j,k}^2 < \epsilon$ for $m \geq M$. Thus $\sum_{i+j > N} a_{j,k}^2 \rightarrow 0$ as $N \rightarrow \infty$.

Let $K_N(\mathbf{x}, \mathbf{y}) = \sum_{i+j \leq N} a_{j,k} \psi_{j,k}(\mathbf{x}, \mathbf{y})$ for $N \geq 0$. Since the image of T_{K_N} is included in the span of $\{\phi_j\}_{j=0}^N$, we have that $\dim \mathcal{R}(T_{K_N}) \leq N + 1 < \infty$. Thus T_{K_N} is a compact operator according to Theorem 4.4.23. We can prove, as we proved (4.4.3), that

$$\|T_K - T_{K_N}\| \leq \|K - K_N\|_2 = \sum_{j+k > N} a_{j,k}^2.$$

Thus T_{K_N} converges to T_K as $N \rightarrow \infty$. Since the space of compact operators is closed with respect to the induce operator norm according to Theorem 4.4.21, we have that T_K is compact. ■

4.5 Lax-Milgram Theorem

Definition 4.5.1

Let H be an Hilbert space over $\mathbb{C}$. A **bounded (or continuous) sesquilinear form** on H is a mapping $B : H \times H \rightarrow \mathbb{C}$ such that

1. For each $u \in H$, the mapping $v \mapsto B(u, v)$ is conjugate-linear.
2. For each $v \in H$, the mapping $u \mapsto B(u, v)$ is linear.
3. There exists a constant C such that $|B(u, v)| \leq C\|u\| \|v\|$ for all $(u, v) \in H \times H$, where $\|\cdot\|$ is the norm induced by the scalar product $\langle \cdot, \cdot \rangle$ on H .

The bounded sesquilinear form B is **symmetric** if $B(u, v) = \overline{B(v, u)}$ for all $u, v \in H$.

The bounded sesquilinear form B is **coercive** if there exists a positive constant c such that $\operatorname{Re} B(u, u) \geq c\|u\|^2$ for all $u \in H$. If $B : H \times H \rightarrow \mathbb{R}$, then the condition is simply $B(u, u) \geq c\|u\|^2$.

Proposition 4.5.2

Let H be an Hilbert space and K be a non-empty, convex and closed subset of H . Given $w \in H$, there exists a unique element $u \in K$ such that

$$\|w - u\| = \min_{v \in K} \|w - v\| . \quad (4.5.1)$$

Moreover, this element u has the property of being the unique element of K satisfying

$$\operatorname{Re} \langle w - u, v - u \rangle \leq 0 \quad (4.5.2)$$

for all $v \in K$.

Proof.

i) Since $w - K$ is a non-empty, convex and closed subset of H , it has an unique element of smallest norm according to Theorem 4.1.7. If $\tilde{u} \in w - K$ is this element, then $u = w - \tilde{u}$.

ii) Suppose that u satisfies (4.5.1) and take $v \in K$. Let $v_t = (1 - t)u + tv$ for $0 \leq t \leq 1$. we have that $v_t \in K$ for all $t \in [0, 1]$ because K is convex. Thus

$$\begin{aligned} \|w - u\|^2 &\leq \|w - v_t\|^2 = \|(w - u) - t(v - u)\|^2 \\ &= \|w - u\|^2 - 2t \operatorname{Re} \langle w - u, v - u \rangle + t^2 \|u - v\|^2 . \end{aligned}$$

We get that

$$2 \operatorname{Re} \langle w - u, v - u \rangle \leq t \|u - v\|^2$$

for $0 < t \leq 1$. If $t \rightarrow 0$, we get (4.5.2).

Conversely, if $u \in K$ satisfies (4.5.2), then

$$\begin{aligned} \|w - u\|^2 - \|w - v\|^2 &= \|w - u\|^2 - \|(w - u) + (u - v)\|^2 \\ &= -2 \operatorname{Re} \langle w - u, u - v \rangle - \|u - v\|^2 = 2 \operatorname{Re} \langle w - u, v - u \rangle - \|u - v\|^2 \leq 0 \end{aligned}$$

for $v \in K$. Thus (4.5.1) is satisfied.

The uniqueness of u satisfying (4.5.2) comes from the uniqueness of u satisfying (4.5.1). If there are two distinct elements u_1 and u_2 of K satisfying (4.5.2), they must also satisfy (4.5.1) since (4.5.1) and (4.5.2) are equivalent. This is a contradiction of the uniqueness of the unique element of minimal norm in $w - K$. ■

Remark 4.5.3

The uniqueness of the element u satisfying (4.5.2) could also be proved directly. Suppose that there exist $u_1, u_2 \in K$ such that $\operatorname{Re} \langle w - u_1, v - u_1 \rangle \leq 0$ and $\operatorname{Re} \langle w - u_2, v - u_2 \rangle \leq 0$ for all $v \in K$. If we substitute $v = u_2$ in the first inequality, $v = u_1$ in the second inequality and add the resulting inequalities, we get $\langle u_1 - u_2, u_1 - u_2 \rangle \leq 0$. Thus $u_1 = u_2$. ♠

Theorem 4.5.4 (Stampacchia)

Let $B : H \times H \rightarrow \mathbb{R}$ be a coercive bounded sesquilinear form on an Hilbert space H . Moreover, let K be a non-empty, convex and closed subset of H . Then, for any $\tilde{u} \in H^*$, there exists a unique $u \in K$ such that

$$\operatorname{Re} B(v - u, u) \geq \operatorname{Re} \tilde{u}(v - u) \quad (4.5.3)$$

for all $v \in K$. Moreover, if the sesquilinear form B is also symmetric, then u is the unique element of K such that

$$\frac{1}{2}B(u, u) - \operatorname{Re} \tilde{u}(u) = \min_{v \in K} \left\{ \frac{1}{2}B(v, v) - \operatorname{Re} \tilde{u}(v) \right\}. \quad (4.5.4)$$

Proof.

i) Since $v \mapsto B(v, u)$ is a bounded linear function on H for each $u \in H$, we may use the Riesz Representation Theorem to define a unique mapping $A : H \rightarrow H$ such that $B(v, u) = \langle v, Au \rangle$ for all $v \in H$. Au is the unique element of H such that the previous relation is satisfied.

We now prove that A is a bounded, one-to-one linear mapping.

1. By definition of A , we have for $a_1, a_2 \in \mathbb{C}$ and $u_1, u_2 \in H$ that

$$\begin{aligned} \langle v, A(a_1 u_1 + a_2 u_2) \rangle &= B(v, a_1 u_1 + a_2 u_2) = \overline{a_1} B(v, u_1) + \overline{a_2} B(v, u_2) \\ &= \overline{a_1} \langle v, Au_1 \rangle + \overline{a_2} \langle v, Au_2 \rangle = \langle v, a_1 Au_1 + a_2 Au_2 \rangle \end{aligned}$$

for all $v \in H$. Thus $A(a_1 u_1 + a_2 u_2) = a_1 Au_1 + a_2 Au_2$ by uniqueness of the representation of the elements in H^* by elements in H as given by the Riesz Representation Theorem. This show that A is a linear mapping.

2. Since

$$\|Au\|^2 = \langle Au, Au \rangle = B(Au, u) \leq C\|u\|\|Au\|$$

for all $u \in H$, where C is the constant in the definition of a bounded sesquilinear form, we get

$$\|Au\| \leq C\|u\| \quad (4.5.5)$$

for all $u \in H$. Thus, the induced norm of A satisfies $\|A\| \leq C$ and, therefore, A is a bounded linear mapping.

3. We have from the coercive property of B that A is one-to-one. In fact, we have that

$$\operatorname{Re} \langle u_1 - u_2, A(u_1 - u_2) \rangle = \operatorname{Re} B(u_1 - u_2, u_1 - u_2) \geq c\|u_1 - u_2\|^2 \quad (4.5.6)$$

for any $u_1, u_2 \in H$, where c is the constant given in the definition of a coercive bounded sesquilinear form. If $Au_1 = Au_2$, we then get

$$0 = \operatorname{Re} \langle u_1 - u_2, A(u_1 - u_2) \rangle \geq c\|u_1 - u_2\|^2.$$

Thus $u_1 = u_2$.

ii) According to Riesz Representation Theorem, Theorem 4.1.9, there exists a unique $\check{u} \in H$ such that $\tilde{u}(v) = \langle v, \check{u} \rangle$ for all $v \in H$.

Proving that there exists a unique $u \in K$ satisfying (4.5.3) is equivalent to proving that there exists a unique $u \in K$ satisfying

$$\operatorname{Re} \langle v - u, Au \rangle \geq \operatorname{Re} \langle v - u, \check{u} \rangle \quad (4.5.7)$$

for all $v \in K$. Let p be a positive number. Proving that there exists a unique $u \in K$ satisfying (4.5.7) is equivalent to proving that there exists a unique $u \in K$ satisfying

$$\operatorname{Re} \langle v - u, (p\check{u} - pAu + u) - u \rangle \leq 0 \quad (4.5.8)$$

for all $v \in K$.

Let $P : K \rightarrow K$ be the function defined by $P(w) = u$ where u satisfies (4.5.2). Because of the uniqueness of u , the function P is well defined. Proving that there exists a unique $u \in K$ satisfying (4.5.8) is thus equivalent to proving that there exists a unique fixed point u of

$$\begin{aligned} F : K &\rightarrow K \\ v &\mapsto P(p\check{u} - pAv + v) \end{aligned}$$

To prove this last statement, it is enough to prove that F is a strict contraction on the closed set K . That is what we now do. We will choose the constant p to get a strict contraction.

We first prove that $\|P(w_1) - P(w_2)\| \leq \|w_1 - w_2\|$ for all $w_1, w_2 \in H$. We have that $P(w_1)$ and $P(w_2)$ are defined by $\operatorname{Re} \langle w_1 - P(w_1), v - P(w_1) \rangle \leq 0$ and $\operatorname{Re} \langle w_2 - P(w_2), v - P(w_2) \rangle \leq 0$ for all $v \in K$ respectively. If we substitute $v = P(w_2)$ in the first inequality, $v = P(w_1)$ in the second inequality and add the resulting inequalities, we get

$$\begin{aligned} &\operatorname{Re} \langle w_1 - P(w_1), P(w_2) - P(w_1) \rangle + \operatorname{Re} \langle w_2 - P(w_2), P(w_1) - P(w_2) \rangle \\ &= \operatorname{Re} \langle -w_1 + w_2 + P(w_1) - P(w_2), P(w_1) - P(w_2) \rangle \leq 0 \\ &\Leftrightarrow \langle P(w_1) - P(w_2), P(w_1) - P(w_2) \rangle = \operatorname{Re} \langle P(w_1) - P(w_2), P(w_1) - P(w_2) \rangle \\ &\leq \operatorname{Re} \langle w_1 - w_2, P(w_1) - P(w_2) \rangle \end{aligned}$$

Thus,

$$\begin{aligned} \|P(w_1) - P(w_2)\|^2 &= \langle P(w_1) - P(w_2), P(w_1) - P(w_2) \rangle \\ &\leq \operatorname{Re} \langle w_1 - w_2, P(w_1) - P(w_2) \rangle \leq \|w_1 - w_2\| \|P(w_1) - P(w_2)\|, \end{aligned}$$

where the last inequality comes from Schwarz inequality. Hence $\|P(w_1) - P(w_2)\| \leq \|w_1 - w_2\|$.

It follows that

$$\begin{aligned} \|F(v_1) - F(v_2)\|^2 &= \|P(p\check{u} - pAv_1 + v_1) - P(p\check{u} - pAv_2 + v_2)\|^2 \\ &\leq \|p(Av_2 - Av_1) - (v_2 - v_1)\|^2 \\ &= p^2 \|Av_2 - Av_1\|^2 - 2p \operatorname{Re} \langle Av_2 - Av_1, v_2 - v_1 \rangle + \|v_2 - v_1\|^2 \\ &\leq p^2 C^2 \|v_2 - v_1\|^2 - 2pc \|v_2 - v_1\|^2 + \|v_2 - v_1\|^2 \end{aligned}$$

$$= (p^2 C^2 - 2pc + 1) \|v_2 - v_2\|^2$$

for all $v_1, v_2 \in K$, where we have used (4.5.5) and (4.5.6) for the last inequality. If we take p such that $0 < p^2 C^2 - 2pc + 1 < 1$, we get that F is a strict contraction on K . Such a p exists because

$$0 > p(pC^2 - 2c) \geq -c^2/C^2 \geq -1$$

for all p is between the two roots, 0 and $2c/C^2$, of the quadratic polynomial $p(pC^2 - 2c)$. The reader should verify that we must have $0 < c \leq C$.

iii) When B is also a symmetric bounded sesquilinear form, proving that there exists a unique $u \in K$ satisfying (4.5.3) is equivalent to proving that there exists a unique $u \in K$ satisfying (4.5.4).

Let

$$\langle v, u \rangle_B = B(v, u)$$

for all $u, v \in H$. Since B is a coercive and symmetric bounded sesquilinear form on H , $\langle \cdot, \cdot \rangle_B$ is a scalar product on H equivalent to the standard scalar product on H . Namely, $\langle \cdot, \cdot \rangle_B$ and $\langle \cdot, \cdot \rangle$ define the same topology on H . We may therefore use the Riesz Representation Theorem to conclude that there exists a unique $\tilde{u} \in H$ such that $\tilde{u}(v) = \langle v, \tilde{u} \rangle_B$ for all $v \in H$. Hence, (4.5.3) becomes

$$\operatorname{Re} \langle v - u, \tilde{u} - u \rangle_B \leq 0 \quad (4.5.9)$$

for all $v \in K$.

According to Proposition 4.5.2, proving that there exists a unique $u \in K$ satisfying (4.5.9) is equivalent to proving that there exists a unique $u \in K$ satisfying

$$\|\tilde{u} - u\|_B = \min_{v \in K} \|\tilde{u} - v\|_B ,$$

where $\|v\|_B = \sqrt{B(v, v)}$ for $v \in H$. This is equivalent to proving that there exists a unique $v \in K$ where

$$v \mapsto B(\tilde{u} - v, \tilde{u} - v) = B(\tilde{u}, \tilde{u}) - 2 \operatorname{Re} B(v, \tilde{u}) + B(v, v)$$

reaches its minimum, which is equivalent to proving that there exists a unique $v \in K$ where

$$v \mapsto -2 \operatorname{Re} B(v, \tilde{u}) + B(v, v) = B(v, v) - 2 \operatorname{Re} \tilde{u}(v)$$

reaches its minimum. This yields (4.5.4). ■

Corollary 4.5.5 (Lax-Milgram Theorem)

Let $B : H \times H \rightarrow \mathbb{C}$ be a bounded, coercive sesquilinear form on an Hilbert space H . Then, for every $\tilde{u} \in H^*$, there exists a unique $u \in H$ such that

$$B(v, u) = \tilde{u}(v) \quad (4.5.10)$$

for all $v \in H$. Moreover, if the bilinear form B is also symmetric, then u is the unique

element of H such that

$$\frac{1}{2}B(u, u) - \operatorname{Re} \tilde{u}(u) = \min_{v \in H} \left\{ \frac{1}{2}B(v, v) - \operatorname{Re} \tilde{u}(v) \right\} . \quad (4.5.11)$$

Proof.

This is Stampacchia's theorem with $K = H$. To prove (4.5.10), we note that (4.5.3) now says that there exists a unique $u \in H$ such that $\operatorname{Re} B(v - u, u) \geq \operatorname{Re} \tilde{u}(v - u)$ for all $v \in H$. Since $H - u = H$, the previous sentence is equivalent to saying that there exists a unique $u \in H$ such that

$$\operatorname{Re} B(v, u) \geq \operatorname{Re} \tilde{u}(v) \quad (4.5.12)$$

for all $v \in H$.

If $v = w$ in (4.5.12), we get

$$\operatorname{Re} B(w, u) \geq \operatorname{Re} \tilde{u}(w) . \quad (4.5.13)$$

If $v = -w$ in (4.5.12), we get $\operatorname{Re} B(-w, u) \geq \operatorname{Re} \tilde{u}(-w)$. Therefore,

$$\operatorname{Re} B(w, u) = -\operatorname{Re} B(-w, u) \leq -\operatorname{Re} \tilde{u}(-w) = \operatorname{Re} \tilde{u}(w) . \quad (4.5.14)$$

(4.5.13) and (4.5.14) implies that $\operatorname{Re} B(w, u) = \operatorname{Re} \tilde{u}(w)$ for all $w \in H$. Thus, $w \mapsto B(w, u)$ and $\tilde{u}$ denote the same linear functional². Hence, (4.5.10) is true.

If $u' \neq u$ satisfies $B(v, u') = \tilde{u}(v)$ for $v \in H$, then (4.5.12) is satisfied with u' instead of u . This is a contradiction that u satisfying (4.5.12) is unique.

Obviously, (4.5.11) is (4.5.4) with $K = H$. ■

Remark 4.5.6

If we are only interested in the existence and uniqueness of $u \in H$ satisfying (4.5.10), there is a short proof of this result. It is almost the first part of the proof of Theorem 4.5.4.

According to Riesz Representation Theorem, there exists a unique $\check{u} \in H$ such that $\tilde{u}(v) = \langle v, \check{u} \rangle$ for all $v \in H$.

Similarly, since $v \mapsto B(v, u)$ is a bounded linear function on H for each $u \in H$, we may use the Riesz Representation Theorem to define a unique mapping $A : H \rightarrow H$ such that $B(v, u) = \langle v, Au \rangle$ for all $v \in H$. Au is the unique element of H such that the previous relation is satisfied.

To prove the existence and uniqueness of $u \in H$ satisfying (4.5.10), it suffices to prove that A is a linear bijection of H onto itself. The requested u in (4.5.10) is $A^{-1}(\check{u})$. The proof that A is a bounded, one-to-one linear mapping is given in the first part of the proof of Theorem 4.5.4. The coercive nature of B implies that $A(H)$ is closed as we now explain. If $\{Au_j\}_{j=1}^{\infty}$ converges to $w \in H$, we have that

$$\|A(u_j)\| \|u_j\| \geq |\langle u_j, A(u_j) \rangle| = |B(u_j, u_j)| \geq c \|u_j\|^2$$

² $\operatorname{Re} B(w, u) = \operatorname{Re} \tilde{u}(w)$ for all $w \in H$ if and only if $\operatorname{Re} B(iw, u) = \operatorname{Re} \tilde{u}(iw)$ for all $w \in H$ if and only if $\operatorname{Re}(iB(w, u)) = \operatorname{Re}(i\tilde{u}(w))$ for all $w \in H$ if and only if $\operatorname{Im} B(w, u) = \operatorname{Im} \tilde{u}(w)$ for all $w \in H$.

for $j \geq 1$. Thus $\|A(u_j)\| \geq c\|u_j\|$ for all j . Since $\{Au_j\}_{j=1}^\infty$ converges, it is a Cauchy sequence. Thus $\{u_j\}_{j=1}^\infty$ is a Cauchy sequence in the Hilbert space H . Hence, the sequence $\{u_j\}_{j=1}^\infty$ converges. Let u be the limit of this converging sequence. Since A is continuous, we have

$$w = \lim_{j \rightarrow \infty} Au_j = A\left(\lim_{j \rightarrow \infty} u_j\right) = A(u) .$$

Thus $w \in A(H)$.

Finally, the range of A is dense in H . Suppose that $w \in A(H)^\perp$, then

$$c\|w\|^2 \leq |B(w, w)| = |\langle w, Aw \rangle| = 0 .$$

Since $\mathcal{R}(A)$ is closed and dense in H , we have that $A(H) = H$. ♠

Remark 4.5.7

To prove that the mapping

$$\begin{aligned} T : H^* &\rightarrow H \\ \tilde{u} &\mapsto u \end{aligned}$$

defined by (4.5.10) in Corollary 4.5.5 is a bounded linear mapping, we show that T is the composition of two bounded linear mappings.

Let $R : H^* \rightarrow H$ be the bounded linear mapping defined by $R(\tilde{u})$ is the unique element given by the Riesz Representation Theorem such that $\tilde{u}(v) = \langle v, R(\tilde{u}) \rangle$ for all $v \in H$. We have $\|R\| = 1$ because $\|\tilde{u}\| = \|R(\tilde{u})\|$ for all $\tilde{u} \in H^*$. Let $A : H \rightarrow H$ be the bounded mapping defined in Remark 4.5.6. Since A is a linear bijection of H onto H , we have by the Open Mapping Theorem that A^{-1} is continuous. Finally, we have that $T = A^{-1} \circ R$. ♠

4.6 Exercises

The problems that we suggest cover only L^2 spaces and in particular the classical Fourier series. For exercises on the general theory of functional analysis, the reader should consult books like [3, 27, 30, 31]. Some of the books on partial differential equations listed in the bibliography have sections on functional analysis with some basic exercises. In general, they are not substitute to good books in function analysis. There is one exception, [3] has a nice introduction to the important results of functional analysis and also include a lot of exercises with detailed solutions for many of them. We strongly recommend this book.

Suggested exercises:

- In [38]: numbers 1 to 7 in Section 5.1; numbers 1, 2, 9 to 15 in Section 5.2; number 2 in Section 5.3; numbers 1 to 16 in Section 5.4.

Chapter 5

Review of some Special Functions

In this chapter, we very briefly present some special functions which are used to express the general solutions of linear second order partial differential equations like the heat equation, the wave equation and the potential equation. Good references for the material of this chapter are [33, 28].

The “Legendre polynomials” that we will introduce later are defined as the solutions of a specific differential equation found using the “Frobenius method” near a “regular point.” Likewise, the Bessel functions that we will also introduce later are defined as the solutions of a specific differential equation found using the “Frobenius method” near a “regular singular point.” We begin by presenting the Frobenius method.

5.1 Frobenius Method

Under what conditions is the solution y of an ordinary differential equation analytic about a point a ? Namely, under what conditions do we have $y = \sum_{n=0}^{\infty} a_n(x-a)^n$, a series with a positive radius of convergence, as a solution of an ordinary differential equation? The following example shows that we cannot always assume that this is possible.

Example 5.1.1

Consider the ordinary differential equation

$$x^2y' + (x-1)y + 1 = 0 . \quad (5.1.1)$$

Suppose that we have a solution of the form $y = \sum_{n=0}^{\infty} a_n x^n$ with a positive radius of convergence.

Then $y' = \sum_{n=1}^{\infty} n a_n x^{n-1}$ on the interval of convergence. If we substitute these series for y and y' in (5.1.1), we get

$$0 = \sum_{n=1}^{\infty} n a_n x^{n+1} + (x-1) \sum_{n=0}^{\infty} a_n x^n + 1 = \sum_{n=1}^{\infty} n a_n x^{n+1} + \sum_{n=0}^{\infty} a_n x^{n+1} - \sum_{n=0}^{\infty} a_n x^n + 1$$

$$\begin{aligned}
&= \sum_{n=2}^{\infty} (n-1)a_{n-1}x^n + \sum_{n=1}^{\infty} a_{n-1}x^n - \sum_{n=0}^{\infty} a_n x^n + 1 \\
&= \sum_{n=2}^{\infty} (na_{n-1} - a_n)x^n + (a_0 - a_1)x - a_0 + 1 .
\end{aligned}$$

Hence, we get $a_0 = 1$ from $1 - a_0 = 0$, $a_1 = a_0 = 1$ from $a_0 - a_1 = 0$, $a_2 = 2a_1 = 2$ from $2a_1 - a_2 = 0$, $a_3 = 3a_2 = 3 \cdot 2 = 3!$ from $3a_2 - a_3 = 0$, and so on. If we assume that $a_{n-1} = (n-1)!$, then we get $a_n = n!$ from $na_{n-1} - a_n = 0$. By induction, we have that $a_n = n!$ for all non-negative integers. We find that $y = \sum_{n=0}^{\infty} n!x^n$. The ratio test shows that this series diverges for $x \neq 0$. Therefore, we cannot assume that there is an analytic solution of (5.1.1) in the neighbourhood of the origin. ♣

Consider ordinary differential equations of the form

$$y'' + P(x)y' + Q(x)y = 0 . \quad (5.1.2)$$

If $P(x)$ and $Q(x)$ are analytic functions at $x = 0$, we say that 0 is a **regular point**. If $xP(x)$ and $x^2Q(x)$ are analytic functions at $x = 0$, we say that 0 is a **regular singular point**.

Suppose that 0 is a regular singular point; say $xP(x) = \sum_{j=0}^{\infty} \alpha_j x^j$ and $x^2Q(x) = \sum_{j=0}^{\infty} \beta_j x^j$ for $|x| < R$. Moreover, suppose λ_1 and λ_2 are the roots of the **indicial equation** $F(\lambda) = \lambda^2 + (\alpha_0 - 1)\lambda + \beta_0 = 0$ with $\lambda_1 \geq \lambda_2$.

One solution of the ordinary differential equation is given by

$$y_1(x) = x^{\lambda_1} (a_0 + a_1 x + a_2 x^2 + \dots) ,$$

where the series converges absolutely for $|x| < R$. A second linearly independent solution of the ordinary differential equation is given by

$$y_2(x) = \begin{cases} x^{\lambda_2} (b_0 + b_1 x + b_2 x^2 + \dots) & \text{if } \lambda_1 - \lambda_2 \notin \mathbb{N} \\
y_1(x) \ln(x) + x^{\lambda_2} (b_1 x + b_2 x^2 + \dots) & \text{if } \lambda_1 = \lambda_2 \\
ky_1(x) \ln(x) + x^{\lambda_2} (b_0 + b_1 x + b_2 x^2 + \dots) & \text{if } 0 < \lambda_1 - \lambda_2 \in \mathbb{N} \end{cases} \quad (5.1.3)$$

where again the series converges absolutely for $|x| < R$ and k is a constant.

There are some situations where we may avoid manipulating the formula for y_2 which has a term in $\ln(x)$. If y_1 is a known function, we can use the method of reduction of order to find a second linearly independent solution; it suffices to replace y by $y_2(x) = y_1(x)u(x)$ in the ordinary differential equation and solve for u .

If $0 < s = \lambda_1 - \lambda_2 \in \mathbb{N}$, the solution y_2 with $k = 0$ gives a second linearly independent solution when b_0 and b_s are free.

There is also a method to find y_2 when $\lambda_1 = \lambda_2$ which does not require the expression with the $\ln(x)$. To find the first series solution y_1 , we substitute the series

$$y(t, \lambda) = x^\lambda \sum_{n=0}^{\infty} c_n x^n \quad (5.1.4)$$

in (5.1.2) to get an expression of the form

$$c_0 F(\lambda) x^\lambda + x^\lambda \sum_{n=1}^{\infty} (c_n F(n + \lambda) + G(\lambda, c_{n-p}, c_{n-p+1}, \dots, c_{n-1})) x^n = 0, \quad (5.1.5)$$

where G is some function, $p > 0$ is a fixed integer and $c_j = 0$ for $j < 0$. If $\lambda = \lambda_1$, then c_0 is free and the recurrence relation

$$c_n = \frac{G(\lambda_1, c_{n-p}, c_{n-p+1}, \dots, c_{n-1})}{F(n + \lambda_1)}$$

for $n \geq 1$ gives all the other coefficients of the solution. When $\lambda_1 = \lambda_2$, the indicial equation is $F(\lambda) = (\lambda - \lambda_1)^2$. If we substitute (5.1.4) with

$$c_n = \frac{G(\lambda, c_{n-p}, c_{n-p+1}, \dots, c_{n-1})}{F(n + \lambda)}$$

for $n \geq 1$ in (5.1.2), we get from (5.1.5) that

$$\frac{d^2 y}{dx^2}(x, \lambda) + P(x) \frac{dy}{dx}(x, \lambda) + Q(x)y(x, \lambda) = c_0(\lambda - \lambda_1)^2 x^\lambda.$$

If we derive this relation with respect to λ , we get

$$\begin{aligned} \frac{d^2}{dx^2} \left(\frac{\partial y}{\partial \lambda}(x, \lambda) \right) + P(x) \frac{d}{dx} \left(\frac{\partial y}{\partial \lambda}(x, \lambda) \right) + Q(x) \left(\frac{\partial y}{\partial \lambda}(x, \lambda) \right) \\ = c_0 \left(2(\lambda - \lambda_1) + (\lambda - \lambda_1)^2 \ln(x) \right) x^\lambda. \end{aligned}$$

Hence $\frac{\partial y}{\partial \lambda}(x, \lambda)|_{\lambda=\lambda_1}$ is another solution of (5.1.2) which is linearly independent of y_1 .

Example 5.1.2

Solve the ordinary differential equation

$$4xy'' + 2y' + y = 0.$$

We have an equation of the form $y'' + P(x)y' + Q(x)y = 0$, where $P(x) = \frac{1}{2x}$ and $Q(x) = \frac{1}{4x}$. Thus $xP(x) = \frac{1}{2}$ and $x^2Q(x) = \frac{x}{4}$. We find $\alpha_0 = 1/2$ and $\beta_0 = 0$. The indicial equation is $\lambda^2 + (\alpha_0 - 1)\lambda + \beta_0 = \lambda^2 - \frac{1}{2}\lambda = \lambda\left(\lambda - \frac{1}{2}\right) = 0$. Its roots are $\lambda_1 = \frac{1}{2}$ and $\lambda_2 = 0$. Since $\lambda_1 - \lambda_2$ is not an integer, we get two linearly independent solutions $y_1 = x^{\lambda_1} \sum_{n=0}^{\infty} a_n x^n$ and $y_2 = x^{\lambda_2} \sum_{n=0}^{\infty} b_n x^n$.

If we substitute $y = x^\lambda \sum_{n=0}^{\infty} c_n x^n$ into $4xy'' + 2y' + y = 0$, we get

$$0 = 4xy'' + 2y' + y = 4 \sum_{n=0}^{\infty} (\lambda + n)(\lambda + n - 1) c_n x^{\lambda+n-1} + 2 \sum_{n=0}^{\infty} (\lambda + n) c_n x^{\lambda+n-1} + \sum_{n=0}^{\infty} c_n x^{\lambda+n}$$

$$\begin{aligned}
&= 4 \sum_{n=-1}^{\infty} (\lambda + n + 1)(\lambda + n) c_{n+1} x^{\lambda+n} + 2 \sum_{n=-1}^{\infty} (\lambda + n + 1) c_{n+1} x^{\lambda+n} + \sum_{n=0}^{\infty} c_n x^{\lambda+n} \\
&= 2\lambda(2\lambda - 1) c_0 x^{\lambda-1} + \sum_{n=0}^{\infty} (2(\lambda + n + 1)(2\lambda + 2n + 1) c_{n+1} + c_n) x^{\lambda+n} .
\end{aligned}$$

We get $2\lambda(2\lambda - 1)c_0 = 0$ and the recursion relation $c_{n+1} = \frac{-1}{2(\lambda + n + 1)(2\lambda + 2n + 1)} c_n$ for $n \geq 0$.

For $\lambda = 1/2$, we have that $2\lambda(2\lambda - 1)c_0 = 0c_0 = 0$ whatever the value of c_0 . Thus, c_0 is free. The recursion relation becomes $c_{n+1} = \frac{-1}{(2n+3)(2n+2)} c_n$ for $n \geq 0$. We find $c_1 = -\frac{c_0}{3 \times 2} = -\frac{c_0}{3!}$, $c_2 = -\frac{c_1}{5 \times 4} = \frac{c_0}{5 \times 4 \times 3 \times 2} = \frac{c_0}{5!}$, $c_3 = -\frac{c_2}{7 \times 6} = -\frac{c_0}{7 \times 6 \times 5 \times 4 \times 3 \times 2} = -\frac{c_0}{7!}$, and so on. By induction, we find that $c_n = (-1)^n \frac{c_0}{(2n+1)!}$ for $n \geq 0$. The first solution is

$$y_1 = c_0 x^{1/2} \sum_{n=0}^{\infty} (-1)^n \frac{x^n}{(2n+1)!} = c_0 \sum_{n=0}^{\infty} (-1)^n \frac{(\sqrt{x})^{2n+1}}{(2n+1)!} = c_0 \sin(\sqrt{x}) .$$

For $\lambda = 0$, we have that $2\lambda(2\lambda - 1)c_0 = 0c_0 = 0$ whatever the value of c_0 . Thus, c_0 is also free. The recursion relation becomes $c_{n+1} = \frac{-1}{(2n+2)(2n+1)} c_n$ for $n \geq 0$. We find $c_1 = -\frac{c_0}{2 \times 1} = -\frac{c_0}{2!}$, $c_2 = -\frac{c_1}{4 \times 3} = \frac{c_0}{4 \times 3 \times 2} = \frac{c_0}{4!}$, $c_3 = -\frac{c_2}{6 \times 5} = -\frac{c_0}{6 \times 5 \times 4 \times 3 \times 2} = -\frac{c_0}{6!}$, and so on. By induction, we find that $c_n = (-1)^n \frac{c_0}{(2n)!}$ for $n \geq 0$. A second solution linearly independent from the first one is

$$y_2 = c_0 \sum_{n=0}^{\infty} (-1)^n \frac{x^n}{(2n)!} = c_0 \sum_{n=0}^{\infty} (-1)^n \frac{(\sqrt{x})^{2n}}{(2n)!} = c_0 \cos(\sqrt{x}) .$$

The general solution is $y = C \sin(\sqrt{x}) + D \cos(\sqrt{x})$ for arbitrary constants C and D . ♣

Example 5.1.3

Solve the ordinary differential equation

$$2xy'' + y' + xy = 0 .$$

We have an equation of the form $y'' + P(x)y' + Q(x)y = 0$, where $P(x) = \frac{1}{2x}$ and $Q(x) = 2$. Thus $xP(x) = \frac{1}{2}$ and $x^2Q(x) = 2x^2$. We find $\alpha_0 = 1/2$ and $\beta_0 = 0$. The indicial equation is $\lambda^2 + (\alpha_0 - 1)\lambda + \beta_0 = \lambda^2 - \frac{1}{2}\lambda = \lambda\left(\lambda - \frac{1}{2}\right) = 0$. Its roots are $\lambda_1 = 1/2$ and $\lambda_2 = 0$. Since $\lambda_1 - \lambda_2$ is not an integer, we have two linearly independent solutions $y_1 = x^{\lambda_1} \sum_{n=0}^{\infty} a_n x^n$ and $y_2 = x^{\lambda_2} \sum_{n=0}^{\infty} b_n x^n$.

If we substitute $y = x^\lambda \sum_{n=0}^{\infty} c_n x^n$ into $2xy'' + y' + xy = 0$, we get

$$\begin{aligned} 0 &= 2xy'' + y' + xy = 2 \sum_{n=0}^{\infty} (\lambda + n)(\lambda + n - 1) c_n x^{\lambda+n-1} + \sum_{n=0}^{\infty} (\lambda + n) c_n x^{\lambda+n-1} + \sum_{n=0}^{\infty} c_n x^{\lambda+n+1} \\ &= 2 \sum_{n=-2}^{\infty} (\lambda + n + 2)(\lambda + n + 1) c_{n+2} x^{\lambda+n+1} + \sum_{n=-2}^{\infty} (\lambda + n + 2) c_{n+2} x^{\lambda+n+1} + \sum_{n=0}^{\infty} c_n x^{\lambda+n+1} \\ &= \lambda(2\lambda - 1) c_0 x^{\lambda-1} + (2\lambda + 1)(\lambda + 1) c_1 x^\lambda \\ &\quad + \sum_{n=0}^{\infty} ((\lambda + n + 2)(2\lambda + 2n + 3) c_{n+2} + c_n) x^{\lambda+n+1}. \end{aligned}$$

We get $\lambda(2\lambda - 1) c_0 = 0$, $(2\lambda + 1)(\lambda + 1) c_1 = 0$ and the recursion relation

$$c_{n+2} = -\frac{1}{(\lambda + n + 2)(2\lambda + 2n + 3)} c_n$$

for $n \geq 0$.

For $\lambda = 1/2$, we have that $\lambda(2\lambda - 1) c_0 = 0 c_0 = 0$ and $(2\lambda + 1)(\lambda + 1) c_1 = 3 c_1 = 0$. Thus, c_0 is free and $c_1 = 0$. The recursion relation becomes $c_{n+2} = -\frac{1}{(n+2)(2n+5)} c_n$ for $n \geq 0$. We find that $c_3 = \frac{-1}{3 \times 7} c_1 = 0$, $c_5 = \frac{-1}{5 \times 11} c_3 = 0$, and so on. Thus, we get $c_{2n-1} = 0$ for $n > 0$ by induction. We also find that $c_2 = -\frac{1}{2 \times 5} c_0$, $c_4 = -\frac{1}{4 \times 9} c_2 = \frac{c_0}{2 \times 4 \times 5 \times 9}$, $c_6 = -\frac{1}{6 \times 13} c_4 = -\frac{c_0}{2 \times 4 \times 6 \times 5 \times 9 \times 13}$, and so on. By induction, we get

$$c_{2n} = (-1)^n \frac{c_0}{2^n n! \times 5 \times 9 \times 13 \times \dots \times (4n+1)}$$

for $n > 0$. With $c_0 = 1$, we get a first solution

$$y_1 = x^{1/2} \sum_{n=0}^{\infty} (-1)^n \frac{1}{2^n n! \times 5 \times 9 \times 13 \times \dots \times (4n+1)} x^{2n}.$$

For $\lambda = 0$, we have that $\lambda(2\lambda - 1) c_0 = 0 c_0 = 0$ and $(2\lambda + 1)(\lambda + 1) c_1 = c_1 = 0$. Thus, c_0 is free and $c_1 = 0$. The recursion relation becomes $c_{n+2} = -\frac{1}{(n+2)(2n+3)} c_n$ for $n \geq 0$. We find that $c_3 = \frac{-1}{3 \times 5} c_1 = 0$, $c_5 = \frac{-1}{5 \times 9} c_3 = 0$, and so on. Thus, we get $c_{2n-1} = 0$ for $n > 0$ by induction. We also find that $c_2 = -\frac{c_0}{2 \times 3}$, $c_4 = -\frac{1}{4 \times 7} c_2 = \frac{c_0}{2 \times 4 \times 3 \times 7}$, $c_6 = -\frac{1}{6 \times 11} c_4 = -\frac{c_0}{2 \times 4 \times 6 \times 3 \times 7 \times 11}$, and so on. By induction, we get

$$c_{2n} = (-1)^n \frac{c_0}{2^n n! \times 3 \times 7 \times 11 \times \dots \times (4n-1)}$$

for $n > 0$. With $c_0 = 1$, we get a second solution linearly independent of the first one

$$y_1 = 1 + \sum_{n=1}^{\infty} (-1)^n \frac{1}{2^n n! \times 3 \times 7 \times 11 \times \dots \times (4n-1)} x^{2n}.$$

The general solution is $y = C y_1(x) + D y_2(x)$ for arbitrary constants C and D . ♣

Example 5.1.4

Solve the ordinary differential equation $y'' - y = 0$.

We have an equation of the form $y'' + P(x)y' + Q(x)y = 0$, where $P(x) = 0$ and $Q(x) = -1$. Thus $xP(x) = 0$ and $x^2Q(x) = -x^2$. We find $\alpha_0 = \beta_0 = 0$. The indicial equation is $\lambda^2 + (\alpha_0 - 1)\lambda + \beta_0 = \lambda^2 - \lambda = \lambda(\lambda - 1) = 0$. Its roots are $\lambda_1 = 1$ and $\lambda_2 = 0$.

Since $\lambda_1 - \lambda_2$ is an integer, one solution is of the form $y_1 = x \sum_{n=0}^{\infty} a_n x^n = \sum_{n=0}^{\infty} a_n x^{n+1}$ and another solution linearly independent of y_1 is of the form $y_2 = k y_1 \ln(x) + \sum_{n=0}^{\infty} b_n x^n$. However, as we will see, we are in the situation where we may take $k = 0$ in y_2 because b_0 and b_1 are free. In this case, we seek solutions of the form $y = \sum_{n=0}^{\infty} c_n x^n$. This will give us two linearly independent solutions. If we substitute this expression of y into $y'' - y = 0$, we get

$$\begin{aligned} y'' - y &= \sum_{n=2}^{\infty} n(n-1)c_n x^{n-2} - \sum_{n=0}^{\infty} c_n x^n = \sum_{n=0}^{\infty} (n+1)(n+2)c_{n+2} x^n - \sum_{n=0}^{\infty} c_n x^n \\ &= \sum_{n=0}^{\infty} ((n+1)(n+2)c_{n+2} + c_n) x^n = 0. \end{aligned}$$

We get the recursion relation $c_{n+2} = \frac{c_n}{(n+1)(n+2)}$ for $n \geq 0$, where b_0 and b_1 are free. We have $c_2 = \frac{c_0}{2} = \frac{c_0}{2!}$, $c_4 = \frac{c_2}{3 \times 4} = \frac{c_0}{2 \times 3 \times 4} = \frac{c_0}{4!}$, $c_6 = \frac{c_4}{5 \times 6} = \frac{c_0}{2 \times 3 \times 4 \times 5 \times 6} = \frac{c_0}{6!}$, and so on. By induction, we get $c_n = \frac{c_0}{n!}$ for n even. Moreover, $c_3 = \frac{c_1}{2 \times 3} = \frac{c_1}{3!}$, $c_5 = \frac{c_3}{4 \times 5} = \frac{c_1}{2 \times 3 \times 4 \times 5} = \frac{c_1}{5!}$, $c_7 = \frac{c_5}{6 \times 7} = \frac{c_1}{2 \times 3 \times 4 \times 5 \times 6 \times 7} = \frac{c_1}{7!}$, and so on. By induction, we get $c_n = \frac{c_1}{n!}$ for n odd.

By selecting two linearly independent pairs (c_0, c_1) , we get two linearly independent solutions. The choice $(c_0, c_1) = (0, 1)$ yields the solution of the form y_1 above and the choice $(c_0, c_1) = (1, 0)$ yields the solution of the form y_2 above with $k = 0$. We will however not choose those pairs. With $c_0 = c_1 = 1$, we get the solution

$$y_1 = \sum_{n=0}^{\infty} \frac{x^{2n}}{(2n)!} + \sum_{n=0}^{\infty} \frac{x^{2n+1}}{(2n+1)!} = \sum_{n=0}^{\infty} \frac{x^n}{n!} = e^x,$$

and with $c_0 = -c_1 = 1$, we get the solution

$$y_2 = \sum_{n=0}^{\infty} \frac{x^{2n}}{(2n)!} - \sum_{n=0}^{\infty} \frac{x^{2n+1}}{(2n+1)!} = \sum_{n=0}^{\infty} \frac{(-1)^n x^n}{n!} = e^{-x}.$$

Those are the two solutions normally found using the standard method to solve second order differential equation with constant coefficients. The general solution is $y(x) = A e^x + B e^{-x}$ for arbitrary constants A and B . ♣

Example 5.1.5

Solve the ordinary differential equation

$$(1 - x^2)y'' - 2xy' + 2y = 0.$$

We have an equation of the form $y'' + P(x)y' + Q(x)y = 0$, where $P(x) = \frac{-2x}{1-x^2} = -2x \sum_{n=0}^{\infty} x^{2n}$ and $Q(x) = \frac{2}{1-x^2} = 2 \sum_{n=0}^{\infty} x^{2n}$ for $|x| < 1$. Thus $xP(x) = -2 \sum_{n=0}^{\infty} x^{2n+2}$ and $x^2Q(x) = 2 \sum_{n=0}^{\infty} x^{2n+2}$ for $|x| < 1$. We find $\alpha_0 = \beta_0 = 0$. The indicial equation is $\lambda^2 + (\alpha_0 - 1)\lambda + \beta_0 = \lambda^2 - \lambda = \lambda(\lambda - 1) = 0$. Its roots are $\lambda_1 = 1$ and $\lambda_2 = 0$. As for the previous example, $\lambda_1 - \lambda_2$ is an integer and we are in the situation where we may take $k = 0$ in y_2 given by (5.1.3) because b_0 and b_1 are free as we will see. We seek solutions of the form $y = \sum_{n=0}^{\infty} c_n x^n$. This will give us two linearly independent solutions. If we substitute this expression for y into $(1-x^2)y'' - 2xy' + 2y = 0$, we get

$$\begin{aligned} 0 &= (1-x^2)y'' - 2xy' + 2y = (1-x^2) \sum_{n=2}^{\infty} n(n-1)c_n x^{n-2} - 2x \sum_{n=1}^{\infty} n c_n x^{n-1} + 2 \sum_{n=0}^{\infty} c_n x^n \\ &= \sum_{n=2}^{\infty} n(n-1)c_n x^{n-2} - \sum_{n=2}^{\infty} n(n-1)c_n x^n - 2 \sum_{n=1}^{\infty} n c_n x^n + 2 \sum_{n=0}^{\infty} c_n x^n \\ &= \sum_{n=0}^{\infty} (n+2)(n+1)c_{n+2} x^n - \sum_{n=2}^{\infty} n(n-1)c_n x^n - 2 \sum_{n=1}^{\infty} n c_n x^n + 2 \sum_{n=0}^{\infty} c_n x^n \\ &= 2(c_0 + c_2) + 6c_3 x + \sum_{n=2}^{\infty} ((n+2)(n+1)c_{n+2} - (n+2)(n-1)c_n) x^n. \end{aligned}$$

We have that c_0 and c_1 are free, $c_2 = -c_0$, $c_3 = 0$ and we find the recursion relation

$$c_{n+2} = \frac{(n+2)(n-1)}{(n+2)(n+1)} c_n = \frac{n-1}{n+1} c_n$$

for $n \geq 2$. Hence $c_5 = \frac{c_3}{2} = 0$, $c_7 = \frac{2c_5}{3} = 0$, and so on. Thus $c_n = 0$ for $n > 2$ and odd. Moreover, $c_4 = \frac{1}{3}c_2 = -\frac{1}{3}c_0$, $c_6 = \frac{3}{5}c_4 = -\frac{1}{5}c_0$, $c_8 = \frac{5}{7}c_6 = -\frac{1}{7}c_0$, and so on. By induction, we find that $c_n = -\frac{c_0}{n-1}$ for $n > 1$ and even. A first solution is given by $c_1 = 1$ and $c_0 = 0$. We get $y_1(x) = x$. A second solution linearly independent of the first one is given by $c_1 = 0$ and $c_0 = 1$. We get $y_2 = \sum_{n=0}^{\infty} \frac{-x^{2n}}{2n-1}$ for $|x| < 1$. The general solution is $y = Cx + D \sum_{n=0}^{\infty} \frac{-x^{2n}}{2n-1}$ for $|x| < 1$ and arbitrary constants C and D . ♣

Example 5.1.6

Solve the ordinary differential equation

$$xy'' + y' + xy = 0.$$

We have an equation of the form $y'' + P(x)y' + Q(x)y = 0$, where $P(x) = \frac{1}{x}$ and $Q(x) = 1$. Thus $xP(x) = 1$ and $x^2Q(x) = x^2$. We find $\alpha_0 = 1$ and $\beta_0 = 0$. The indicial equation is $\lambda^2 + (\alpha_0 - 1)\lambda + \beta_0 = \lambda^2 = 0$. There is only one root of algebraic multiplicity 2; namely, $\lambda = 0$.

One solution is of the form $y_1 = \sum_{n=0}^{\infty} a_n x^n$ and another solution linearly independent of the first one is of the form $y_2 = \ln(x)y_1(x) + \sum_{n=1}^{\infty} b_n x^n$.

If we substitute $y = \sum_{n=0}^{\infty} a_n x^n$ into $xy'' + y' + xy = 0$, we get

$$\begin{aligned} 0 &= xy'' + y' + xy = \sum_{n=2}^{\infty} n(n-1)a_n x^{n-1} + \sum_{n=1}^{\infty} n a_n x^{n-1} + \sum_{n=0}^{\infty} a_n x^{n+1} \\ &= \sum_{n=0}^{\infty} (n+2)(n+1)a_{n+2} x^{n+1} + \sum_{n=-1}^{\infty} (n+2)a_{n+2} x^{n+1} + \sum_{n=0}^{\infty} a_n x^{n+1} \\ &= a_1 + \sum_{n=0}^{\infty} ((n+2)^2 a_{n+2} + a_n) x^{n+1}. \end{aligned}$$

We have that a_0 is free and $a_1 = 0$. The recursion relation becomes $a_{n+2} = -\frac{1}{(n+2)^2} a_n$ for $n \geq 0$. We find that $a_3 = \frac{-1}{3^2} a_1 = 0$, $a_5 = \frac{-1}{5^2} a_3 = 0$, and so on. Thus $a_{2n-1} = 0$ for $n > 0$ by induction. We also have that $a_2 = \frac{-a_0}{2^2}$, $a_4 = -\frac{1}{4^2} a_2 = \frac{a_0}{2^2 \times 4^2}$, $a_6 = -\frac{1}{6^2} a_4 = \frac{-a_0}{2^2 \times 4^2 \times 6^2}$, and so on. By induction, we get that $a_{2n} = (-1)^n \frac{a_0}{(2^n n!)^2}$ for $n > 0$. With $a_0 = 1$, we get a first solution

$$y_1 = \sum_{n=0}^{\infty} (-1)^n \frac{1}{(2^n n!)^2} x^{2n}.$$

If we substitute $y = \ln(x)y_1(x) + \sum_{n=1}^{\infty} b_n x^n$ into $xy'' + y' + xy = 0$, we get

$$\begin{aligned} 0 &= xy'' + y' + xy = x \left(\ln(x)y_1''(x) + \frac{2y_1'(x)}{x} - \frac{y_1(x)}{x^2} \right) + \sum_{n=2}^{\infty} n(n-1)b_n x^{n-1} \\ &\quad + \left(\ln(x)y_1'(x) + \frac{y_1(x)}{x} \right) + \sum_{n=1}^{\infty} n b_n x^{n-1} + x \ln(x)y_1(x) + \sum_{n=1}^{\infty} b_n x^{n+1} \\ &= \ln(x) \underbrace{(xy_1''(x) + y_1'(x) + xy_1(x))}_{=0} + 2y_1'(x) + \sum_{n=0}^{\infty} (n+2)(n+1)b_{n+2} x^{n+1} \\ &\quad + \sum_{n=-1}^{\infty} (n+2)b_{n+2} x^{n+1} + \sum_{n=1}^{\infty} b_n x^{n+1} \\ &= 4 \sum_{n=1}^{\infty} (-1)^n \frac{n}{(2^n n!)^2} x^{2n-1} + b_1 + 4b_2 x + \sum_{n=1}^{\infty} ((n+2)^2 b_{n+2} + b_n) x^{n+1}. \end{aligned}$$

If we collect the terms in x^n with n even, and those in x^n with n odd, we get

$$0 = b_1 + \sum_{n=1}^{\infty} ((2n+1)^2 b_{2n+1} + b_{2n-1}) x^{2n}$$

and

$$\begin{aligned} 0 &= 4 \sum_{n=1}^{\infty} (-1)^n \frac{n}{(2^n n!)^2} x^{2n-1} + 4b_2 x + \sum_{n=2}^{\infty} ((2n)^2 b_{2n} + b_{2n-2}) x^{2n-1} \\ &= (4b_2 - 1)x + 4 \sum_{n=2}^{\infty} \left((-1)^n \frac{n}{(2^n n!)^2} + (2n)^2 b_{2n} + b_{2n-2} \right) x^{2n-1} . \end{aligned}$$

Thus, $b_1 = 0$ and $b_{2n+1} = -\frac{1}{(2n+1)^2} b_{2n-1}$ for $n > 0$. By induction, we get $b_{2n+1} = 0$ for $n \geq 0$.

Moreover, $b_2 = 1/4$ and $b_{2n} = -\frac{1}{(2n)^2} \left(b_{2n-2} + (-1)^n \frac{4n}{(2^n n!)^2} \right)$ for $n > 1$. Hence,

$$\begin{aligned} b_4 &= -\frac{1}{4^2} \left(\frac{1}{4} + \frac{8}{(2^2 2!)^2} \right) = -\frac{1}{2^4 (2!)^2} \left(1 + \frac{1}{2} \right) \\ b_6 &= -\frac{1}{6^2} \left(b_4 - \frac{12}{(2^3 3!)^2} \right) = -\frac{1}{2^6 (3!)^2} \left(\frac{2^6 (3!)^2}{6^2} b_4 - \frac{1}{3} \right) \\ &= -\frac{1}{2^6 (3!)^2} \left(-\frac{2^6 (3!)^2}{2^4 (2!)^2 6^2} \left(1 + \frac{1}{2} \right) - \frac{1}{3} \right) = \frac{1}{2^6 (3!)^2} \left(1 + \frac{1}{2} + \frac{1}{3} \right) , \end{aligned}$$

and so on. By induction, we find that $b_{2n} = \frac{(-1)^{n-1}}{2^{2n} (n!)^2} \left(1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n} \right)$ for $n > 0$.

A second solution linearly independent of y_1 is given by

$$y_2 = \ln(x) y_1(x) + \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{2^{2n} (n!)^2} \left(1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n} \right) x^{2n} .$$

The general solution is $y = C y_1(x) + D y_2(x)$ for arbitrary constants C and D .

In fact, we have that $y_1(x) = J_0(x)$, the “Bessel function of the first kind of order zero”, and $y_2(x) = Y_0(x)$, the “Bessel function of the second kind of order zero”. We will study these functions in a future section. ♣

Example 5.1.7

Solve the ordinary differential equation

$$xy'' + y' + 2y = 0 .$$

We have an equation of the form $y'' + P(x)y' + Q(x)y = 0$, where $P(x) = \frac{1}{x}$ and $Q(x) = \frac{2}{x}$. Thus $xP(x) = 1$ and $x^2Q(x) = 2x$. We find $\alpha_0 = 1$ and $\beta_0 = 0$. The indicial equation is $\lambda^2 + (\alpha_0 - 1)\lambda + \beta_0 = \lambda^2 = 0$. There is only one root of algebraic multiplicity 2; namely, $\lambda = 0$.

This time, we will not use the second formula in (5.1.3) to find two linearly independent solutions. Instead, we will use the method involving the derivative with respect to λ mentioned at the beginning of the section.

We seek a solution of the form $y = \sum_{n=0}^{\infty} c_n x^{n+\lambda}$. If we substitute this expression into $xy'' + y' + 2y = 0$, we get

$$\begin{aligned} 0 &= xy'' + y' + 2y = \sum_{n=0}^{\infty} (\lambda + n)(\lambda + n - 1)c_n x^{\lambda+n-1} + \sum_{n=0}^{\infty} (\lambda + n)c_n x^{\lambda+n-1} + 2 \sum_{n=0}^{\infty} c_n x^{\lambda+n} \\ &= \sum_{n=-1}^{\infty} (\lambda + n + 1)(\lambda + n)c_{n+1} x^{\lambda+n} + \sum_{n=-1}^{\infty} (\lambda + n + 1)c_{n+1} x^{\lambda+n} + 2 \sum_{n=0}^{\infty} c_n x^{\lambda+n} \\ &= \lambda^2 c_0 x^{\lambda-1} + \sum_{n=0}^{\infty} ((\lambda + n + 1)^2 c_{n+1} + 2c_n) x^{\lambda+n} . \end{aligned}$$

From the recurrence relation $c_{n+1} = -\frac{2c_n}{(\lambda + n + 1)^2}$ for $n \geq 0$, we find by induction that

$$c_n = \frac{(-2)^n c_0}{(\lambda + n)^2 (\lambda + n - 1)^2 \dots (\lambda + 1)^2}$$

for $n \geq 1$. With these values for c_n , the function

$$y(x) = c_0 x^\lambda + \sum_{n=1}^{\infty} \frac{(-2)^n c_0}{(\lambda + n)^2 (\lambda + n - 1)^2 \dots (\lambda + 1)^2} x^{n+\lambda}$$

satisfies $xy'' + y' + 2y = \lambda^2 c_0 x^{\lambda-1}$. If $\lambda = 0$, we get the solution

$$y_1(x) = c_0 + \sum_{n=1}^{\infty} \frac{(-2)^n c_0}{n^2 (n-1)^2 \dots 1^2} x^n = c_0 \sum_{n=0}^{\infty} \frac{(-2)^n}{(n!)^2} x^n .$$

Note that we may have $\lambda^2 c_0 = 0$ with $c_0 \neq 0$ because $\lambda = 0$; namely, c_0 is free.

To get a second solution linearly independent of the first, we consider

$$\begin{aligned} y(x) &= c_0 x^\lambda + \sum_{n=1}^{\infty} \frac{(-2)^n c_0}{(\lambda + n)^2 (\lambda + n - 1)^2 \dots (\lambda + 1)^2} x^{n+\lambda} \\ &= x^\lambda \left(c_0 + \sum_{n=1}^{\infty} \frac{(-2)^n c_0}{(\lambda + n)^2 (\lambda + n - 1)^2 \dots (\lambda + 1)^2} x^n \right) . \end{aligned}$$

This second solution is given by

$$\begin{aligned} y_2(x) &= \frac{\partial y}{\partial \lambda}(x, \lambda) \Big|_{\lambda=0} = \left(x^\lambda \ln(x) \left(c_0 + \sum_{n=1}^{\infty} \frac{(-2)^n c_0}{(\lambda + n)^2 (\lambda + n - 1)^2 \dots (\lambda + 1)^2} x^n \right) \right) \Big|_{\lambda=0} \\ &\quad + \left(x^\lambda \sum_{n=1}^{\infty} (-2)^n c_0 \frac{\partial}{\partial \lambda} \left((\lambda + n)^{-2} (\lambda + n - 1)^{-2} \dots (\lambda + 1)^{-2} \right) x^n \right) \Big|_{\lambda=0} \\ &= \ln(x) y_1(x) + \sum_{n=1}^{\infty} (-2)^n c_0 \left((-2(\lambda + n)^{-1} - 2(\lambda + n - 1)^{-1} - \dots - 2(\lambda + 1)^{-1}) \right. \\ &\quad \left. (\lambda + n)^{-2} (\lambda + n - 1)^{-2} \dots (\lambda + 1)^{-2} \right) \Big|_{\lambda=0} x^n \\ &= \ln(x) y_1(x) + c_0 \sum_{n=1}^{\infty} \frac{(-2)^{n+1}}{(n!)^2} \left(1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n} \right) x^n \end{aligned}$$

The general solution is $y = C y_1(x) + D y_2(x)$ for arbitrary constants C and D . ♣

5.2 Sturm-Liouville Problems

As we will see later, results from the theory of Sturm-Liouville problems are used to show that the Legendre polynomials and the Bessel functions form two complete orthogonal sets of functions in some spaces of functions.

All linear second order ordinary differential equation can be written as

$$(r(x)y')' + (q(x) + \lambda p(x))y = 0 .$$

If we consider this ordinary differential equation on an interval $]a, b[$ and add some conditions on y or y' at the end points a and b , we get a **Sturm-Liouville problem**.

The **regular Sturm Liouville problem** is to solve

$$(r(x)y')' + (q(x) + \lambda p(x))y = 0 \tag{5.2.1}$$

for $-\infty < a < x < b < \infty$ with the boundary conditions

$$\alpha_1 y(a) + \alpha_2 y'(a) = 0 \quad \text{and} \quad \beta_1 y'(b) + \beta_2 y(b) = 0 ,$$

where $(\alpha_1, \alpha_2) \neq (0, 0)$ and $(\beta_1, \beta_2) \neq (0, 0)$; r, q, p and r' are continuous real functions on $[a, b]$; and $p(x) > 0$ for all $x \in [a, b]$.

There are three other types of boundary conditions which are often used.

1. If $r(a) = 0$, we request that the solution and its derivative be bounded at $x = a$ and that $\beta_1 y'(b) + \beta_2 y(b) = 0$ with $(\beta_1, \beta_2) \neq (0, 0)$.
2. If $r(b) = 0$, we request that the solution and its derivative be bounded at $x = b$ and that $\alpha_1 y'(a) + \alpha_2 y(a) = 0$ with $(\alpha_1, \alpha_2) \neq (0, 0)$.
3. If $r(a) = r(b)$, we request the periodic boundary conditions $y(a) = y(b)$ and $y'(a) = y'(b)$.

The Sturm-Liouville problem is said to be **singular** if one of the following conditions is satisfied: (1) $a = \infty$ or $b = \infty$, (2) r or p are null at a or b , or (3) r' does not exist at a or b .

Sturm-Liouville problems which are **singular** only because $r(a) = 0$ (resp. $r(b) = 0$) have many of the properties that regular Sturm-Liouville problems have if we assume that the solution is bounded at a (resp. b). Such problems are called **regular singular Sturm-Liouville problem**.

A value of λ for which the Sturm-Liouville problem has non-trivial solutions is called an **eigenvalue**, the non-trivial solutions are called the **eigenfunctions** associated to the eigenvalue λ . The following result has a simple proof.

Proposition 5.2.1

The eigenfunctions associated to distinct eigenvalues are orthogonal with respect to the weight function p ; more precisely, if y_n is an eigenfunction associate to λ_n and y_m is an eigenfunction associate to λ_m with $n \neq m$, then

$$\int_a^b p(x)y_n(x)y_m(x) dx = 0 . \quad (5.2.2)$$

Proof.

Let y_n be a solution of (5.2.1) with $\lambda = \lambda_n$ and y_m be a solution of (5.2.1) with $\lambda = \lambda_m$. We have $(r(x)y_n')' + q(x)y_n = -\lambda_n p(x)y_n$ and $(r(x)y_m')' + q(x)y_m = -\lambda_m p(x)y_m$. If we multiply the first equation by y_m and the second by y_n , and subtract the resulting second equation from the resulting first equation, then we get

$$(r(x)y_n')'y_m - (r(x)y_m')'y_n = (\lambda_m - \lambda_n)p(x)y_ny_m .$$

Since

$$\frac{d}{dx}(r(x)y_n'y_m - r(x)y_m'y_n) = (r(x)y_n')'y_m - (r(x)y_m')'y_n ,$$

we get

$$\frac{d}{dx}(r(x)y_n'y_m - r(x)y_m'y_n) = (\lambda_m - \lambda_n)p(x)y_ny_m .$$

If we integrate both sides with respect to x for $a < x < b$, then we get

$$\begin{aligned} (\lambda_m - \lambda_n) \int_a^b p(x)y_n(x)y_m(x) dx &= \left(r(x)y_n'(x)y_m(x) - r(x)y_m'(x)y_n(x) \right) \Big|_a^b \\ &= \left(r(x) \det \begin{pmatrix} y_m(x) & y_m'(x) \\ y_n(x) & y_n'(x) \end{pmatrix} \right) \Big|_a^b . \end{aligned} \quad (5.2.3)$$

i) Suppose that y_n and y_m satisfy the boundary conditions $\alpha_1 y(a) + \alpha_2 y'(a) = 0$ with $(\alpha_1, \alpha_2) \neq (0, 0)$ and $\beta_1 y(b) + \beta_2 y'(b) = 0$ with $(\beta_1, \beta_2) \neq (0, 0)$.

The boundary condition $\alpha_1 y_n(a) + \alpha_2 y_n'(a) = 0$ and $\alpha_1 y_m(a) + \alpha_2 y_m'(a) = 0$ with $(\alpha_1, \alpha_2) \neq (0, 0)$ implies that

$$\begin{pmatrix} y_m(a) & y_m'(a) \\ y_n(a) & y_n'(a) \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

has a non-trivial solution; namely $(u, v) = (\alpha_1, \alpha_2)$. This is possible (if and) only if

$$\det \begin{pmatrix} y_m(a) & y_m'(a) \\ y_n(a) & y_n'(a) \end{pmatrix} = 0 . \text{ Similarly, } \det \begin{pmatrix} y_m(b) & y_m'(b) \\ y_n(b) & y_n'(b) \end{pmatrix} = 0 .$$

It follows from (5.2.3) that $(\lambda_m - \lambda_n) \int_a^b p(x)y_n(x)y_m(x) dx = 0$. Thus,

$$\int_a^b p(x)y_n(x)y_m(x) dx = 0 \text{ for } \lambda_m \neq \lambda_n .$$

ii) Suppose that $r(a) = 0$; that y_n, y_m, y_n' and y_m' are bounded at $x = a$; and that y_n and y_m satisfy the boundary condition $\beta_1 y(b) + \beta_2 y'(b) = 0$ with $(\beta_1, \beta_2) \neq (0, 0)$.

The right hand side of (5.2.3) is null at $x = a$ because $r(a) = 0$, and y_n, y_m, y'_n, y'_m are bounded at $x = a$. The boundary condition $\beta_1 y_n(b) + \beta_2 y'_n(b) = 0$ and $\beta_1 y_m(b) + \beta_2 y'_m(b) = 0$ with $(\beta_1, \beta_2) \neq (0, 0)$ implies that

$$\begin{pmatrix} y_m(b) & y'_m(b) \\ y_n(b) & y'_n(b) \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

has a non-trivial solution; namely, $(u, v) = (\beta_1, \beta_2)$. This is possible (if and) only if $\det \begin{pmatrix} y_m(b) & y'_m(b) \\ y_n(b) & y'_n(b) \end{pmatrix} = 0$. Therefore, the right hand side of (5.2.3) is also null at $x = b$.

It follows from (5.2.3) that $(\lambda_m - \lambda_n) \int_a^b p(x) y_n(x) y_m(x) dx = 0$. Thus, $\int_a^b p(x) y_n(x) y_m(x) dx = 0$ for $\lambda_m \neq \lambda_n$.

iii) Suppose that $r(b) = 0$; that y_n, y_m, y'_n and y'_m are bounded at $x = b$; and that y_n and y_m satisfy the boundary condition $\alpha_1 y'(a) + \alpha_2 y(a) = 0$ with $(\alpha_1, \alpha_2) \neq (0, 0)$.

The proof that $\int_a^b p(x) y_n(x) y_m(x) dx = 0$ for $\lambda_m \neq \lambda_n$ is similar to the proof given in (ii).

iv) Suppose that $r(a) = r(b)$ and that the periodic boundary conditions $y(a) = y(b)$ and $y'(a) = y'(b)$ are satisfied by y_n and y_m .

It follows that

$$\left(r(x) y'_n(x) y_m(x) - r(x) y'_m(x) y_n(x) \right) \Big|_a^b = 0.$$

Hence, (5.2.3) yields again $\int_a^b p(x) y_n(x) y_m(x) dx = 0$ for $\lambda_m \neq \lambda_n$. ■

For the regular Sturm-Liouville problem, there is an infinite number of distinct eigenvalues $\lambda_0 < \lambda_1 < \lambda_2 < \dots$ converging to $+\infty$. The subspace of all eigenfunctions associated to an eigenvalue λ_n is one-dimensional. If we select an eigenfunction ϕ_n associated to the eigenvalue λ_n for each n , then we get a complete orthogonal set $\{\phi_n\}_{n=0}^\infty$ in $L^2[a, b]$ with respect to the scalar product defined by

$$\langle f, g \rangle = \int_a^b f(x) g(x) p(x) dx$$

for $f, g \in L^2[a, b]$. All these results are proved by first expressing the Sturm-Liouville problem into an integral equation of the form

$$(T_K y)(x) = \int_a^b K(x, t) y(t) dt = \frac{1}{\lambda} y(x)$$

for $y \in L^2[a, b]$, where $K \in L^2([a, b] \times [a, b])$ is a Hilbert-Schmidt kernel satisfying $K(x, t) = K(t, x)$ for almost all x and t . The function K is given by a “Green function”. Details of this proof can found in [28]. Because of this symmetry, we get from Theorem 4.4.46 that T_K is a self-adjoint, compact operator on $L^2[a, b]$. The results about $\{\phi_n\}_{n=0}^\infty$ stated above are consequences of the Hilbert-Schmidt Theorem, Theorem 4.4.44. Note that the eigenvalues

of the operator T_K are of the form $1/\lambda$ where λ is what we have defined as an eigenvalue for the Sturm-Liouville problem. Mathematics is full of these historical inconsistencies.

Example 5.2.2

Find the eigenvalues and eigenfunctions of the following Sturm-Liouville problem.

$$y'' + \lambda y = 0 \quad \text{with} \quad y'(0) = y'(5) = 0.$$

There are three cases to consider.

i) For $\lambda < 0$, the solutions are of the form $y = A e^{\sqrt{-\lambda}x} + B e^{-\sqrt{-\lambda}x}$.

ii) For $\lambda = 0$, the solutions are of the form $y = Ax + B$.

iii) For $\lambda > 0$, the solutions are of the form $y = A \cos(\sqrt{\lambda}x) + B \sin(\sqrt{\lambda}x)$.

i) For $\lambda < 0$, we have $y' = A \sqrt{-\lambda} e^{\sqrt{-\lambda}x} - B \sqrt{-\lambda} e^{-\sqrt{-\lambda}x}$. The boundary condition $y'(0) = 0$ yields $A \sqrt{-\lambda} - B \sqrt{-\lambda} = 0$ and, after a division by $\sqrt{-\lambda}$, we get $A - B = 0$. The boundary condition $y'(5) = 0$ yields $A \sqrt{-\lambda} e^{5\sqrt{-\lambda}} - B \sqrt{-\lambda} e^{-5\sqrt{-\lambda}} = 0$ and, after a division by $\sqrt{-\lambda}$, we get $A e^{5\sqrt{-\lambda}} - B e^{-5\sqrt{-\lambda}} = 0$. We seek the solutions of the linear system

$$\begin{pmatrix} 1 & -1 \\ e^{5\sqrt{-\lambda}} & -e^{-5\sqrt{-\lambda}} \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Since

$$\det \begin{pmatrix} 1 & -1 \\ e^{5\sqrt{-\lambda}} & -e^{-5\sqrt{-\lambda}} \end{pmatrix} = e^{5\sqrt{-\lambda}} - e^{-5\sqrt{-\lambda}} = (e^{10\sqrt{-\lambda}} - 1) e^{-5\sqrt{-\lambda}} \neq 0$$

for $\lambda < 0$, the only solution of the linear system is $A = B = 0$; namely, we get the trivial solution. Therefore, there is no eigenvalue and eigenfunction for $\lambda < 0$.

ii) For $\lambda = 0$, we have $y' = A$. The boundary conditions $y'(0) = 0$ and $y'(5) = 0$ give $A = 0$. There is no constrain on B . We get a family of non-trivial solutions.

We have that $\lambda = \lambda_0 = 0$ is an eigenvalue and an eigenfunction associated to this eigenvalue is $y_0(x) = 1$ for all x if we take $B = 1$.

iii) For $\lambda > 0$, we have $y' = -A \sqrt{\lambda} \sin(\sqrt{\lambda}x) + B \sqrt{\lambda} \cos(\sqrt{\lambda}x)$. The boundary condition $y'(0) = 0$ gives $B \sqrt{\lambda} = 0$. Thus $B = 0$. The boundary condition $y'(5) = 0$ gives $-A \sqrt{\lambda} \sin(5\sqrt{\lambda}) = 0$, where we have used $B = 0$. After a division by $-\sqrt{\lambda}$, we get $A \sin(5\sqrt{\lambda}) = 0$. Since we are looking for non-trivial solutions, we must assume that $A \neq 0$. We obtain $\sin(5\sqrt{\lambda}) = 0$. Thus, $5\sqrt{\lambda} = n\pi$ for $n \geq 1$.

We have found the eigenvalues $\lambda = \lambda_n = (n\pi/5)^2$ for $n \geq 1$. The eigenfunctions associated to the eigenvalue λ_n are real multiples of $y_n = \cos(n\pi x/5)$.

We have that $y'' + \lambda y = 0$ is of the form (5.2.1) with $r(x) = q(x) = p(x) = 1$ for all x . Thus, from Proposition 5.2.1, we get

$$\int_0^5 y_n(x) y_m(x) dx = 0$$

for $n \neq m$ as it can easily be verified directly. Moreover, $\{\cos(n\pi x/5)\}_{n=0}^{\infty}$ is a complete orthogonal set in $L^2[0, 5]$. This is something that we have used when studying Fourier cosine series. ♣

Example 5.2.3

Find the eigenvalues and eigenfunctions of the following Sturm-Liouville problem.

$$y'' + \lambda y = 0 \quad \text{with} \quad y'(0) = 0 \quad \text{and} \quad y(L) = 0.$$

As in the previous example, there are three cases to consider.

- i) For $\lambda < 0$, the solutions are of the form $y = A e^{\sqrt{-\lambda}x} + B e^{-\sqrt{-\lambda}x}$.
 - ii) For $\lambda = 0$, the solutions are of the form $y = Ax + B$.
 - iii) For $\lambda > 0$, the solutions are of the form $y = A \cos(\sqrt{\lambda}x) + B \sin(\sqrt{\lambda}x)$.
- i) For $\lambda < 0$, we have $y' = A \sqrt{-\lambda} e^{\sqrt{-\lambda}x} - B \sqrt{-\lambda} e^{-\sqrt{-\lambda}x}$. The boundary condition $y'(0) = 0$ yields $A \sqrt{-\lambda} - B \sqrt{-\lambda} = 0$ and, after a division by $\sqrt{-\lambda}$, we get $A - B = 0$. The boundary condition $y(L) = 0$ yields $A e^{L\sqrt{-\lambda}} + B e^{-L\sqrt{-\lambda}} = 0$.

We seek the solutions of the linear system

$$\begin{pmatrix} 1 & -1 \\ e^{L\sqrt{-\lambda}} & e^{-L\sqrt{-\lambda}} \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Since

$$\det \begin{pmatrix} 1 & -1 \\ e^{L\sqrt{-\lambda}} & e^{-L\sqrt{-\lambda}} \end{pmatrix} = e^{-L\sqrt{-\lambda}} + e^{L\sqrt{-\lambda}} > 0$$

for $\lambda < 0$, the only solution of the linear system is $A = B = 0$; namely, we get the trivial solution. There is no eigenvalue and eigenfunction for $\lambda < 0$.

ii) For $\lambda = 0$, we have $y' = A$. The boundary conditions $y'(0) = A = 0$ and $y(L) = AL + B = 0$ yields $A = B = 0$. Again, there is no non-trivial solution. Hence, 0 is not an eigenvalue.

iii) For $\lambda > 0$, we have $y' = -A \sqrt{\lambda} \sin(\sqrt{\lambda}x) + B \sqrt{\lambda} \cos(\sqrt{\lambda}x)$. The boundary condition $y'(0) = 0$ gives $B \sqrt{\lambda} = 0$. Thus $B = 0$. The boundary condition $y(L) = 0$ gives $A \cos(L\sqrt{\lambda}) = 0$, where we have used $B = 0$. Since we are looking for non-trivial solutions, we must assume that $A \neq 0$. Hence $\cos(L\sqrt{\lambda}) = 0$. Thus, $L\sqrt{\lambda} = n\pi + \pi/2$ for $n \geq 0$.

We have found the eigenvalues $\lambda = \lambda_n = ((2n+1)\pi/(2L))^2$ for $n \geq 0$. The eigenfunctions associated to the eigenvalue λ_n are of the form $y = A \cos((2n+1)\pi x/(2L))$ for $n \geq 0$, where A is a non-zero constant. ♣

5.3 Legendre Polynomials $P_n(x)$ on $[-1, 1]$

Good references on this subject and the subjects of the next sections of this chapter are [26, 33]. In particular, the reader will find in these references the proofs of the results that we give without proofs.

The **Legendre differential equation** is

$$(1 - x^2)y'' - 2xy' + n(n+1)y = 0 . \quad (5.3.1)$$

Using the Frobenius method at the regular point 0, we find that the bounded solutions (on the interval $] -1, 1[$) of this differential equation are multiples of

$$P_n(x) = \frac{1}{2^n} \sum_{m=0}^{\lfloor n/2 \rfloor} (-1)^m \binom{n}{m} \binom{2n-2m}{n} x^{n-2m} ,$$

where $\lfloor n/2 \rfloor$ is the greater integer less than or equal to $n/2$. The polynomials P_n are called the **standardized Legendre polynomials** because they satisfy the standardization condition $P_n(1) = 1$. Moreover, we have that $|P_n(x)| \leq 1$ for $-1 \leq x \leq 1$.

The polynomials P_n can be generated by the recurrence relation

$$(n+1)P_{n+1}(x) = (2n+1)xP_n(x) - nP_{n-1}(x)$$

or by the **Rodrigue's formula** given in the following proposition.

Proposition 5.3.1

$$P_n(x) = \frac{(-1)^n}{2^n n!} \frac{d^n}{dx^n} (1 - x^2)^n .$$

Proof.

Since

$$(1 - x^2)^n = \sum_{k=0}^n \frac{n!}{(n-k)!k!} (-x^2)^{n-k} ,$$

we have

$$\begin{aligned} \frac{(-1)^n}{2^n n!} \frac{d^n}{dx^n} (1 - x^2)^n &= \frac{(-1)^n}{2^n n!} \frac{d^n}{dx^n} \left(\sum_{k=0}^n \frac{n!}{(n-k)!k!} (-x^2)^{n-k} \right) \\ &= \frac{1}{2^n} \sum_{k=0}^{\lfloor n/2 \rfloor} \frac{(-1)^k}{(n-k)!k!} \left(\frac{d^n}{dx^n} x^{2n-2k} \right) \end{aligned}$$

because the derivative of order n of x^{2n-2k} is zero for $2n-2k < n$; namely, $k > n/2$. Since

$$\frac{d^n}{dx^n} (x^{2n-2k}) = (2n-2k)(2n-2k-1)\cdots(n-2k+1)x^{n-2k} = \frac{(2n-2k)!}{(n-2k)!} x^{n-2k} ,$$

we get

$$\begin{aligned} \frac{(-1)^n}{2^n n!} \frac{d^n}{dx^n} (1 - x^2)^n &= \frac{1}{2^n} \sum_{k=0}^{\lfloor n/2 \rfloor} \frac{(-1)^k (2n-2k)!}{(n-k)!k!} x^{n-2k} \\ &= \frac{1}{2^n} \sum_{k=0}^{\lfloor n/2 \rfloor} (-1)^k \left(\frac{n!}{(n-k)!k!} \right) \left(\frac{(2n-2k)!}{(n-2k)!n!} \right) x^{n-2k} \\ &= \frac{1}{2^n} \sum_{k=0}^{\lfloor n/2 \rfloor} (-1)^k \binom{n}{k} \binom{2n-2k}{n} x^{n-2k} = P_n(x) . \quad \blacksquare \end{aligned}$$

There is also a generating series for the Legendre polynomials; namely,

$$\frac{1}{\sqrt{1-2xt+t^2}} = \sum_{n=0}^{\infty} P_n(x)t^n \quad , \quad -1 < x, t < 1 .$$

The first six Legendre polynomials are $P_0(x) = 1$, $P_1(x) = x$, $P_2(x) = \frac{1}{2}(3x^2 - 1)$, $P_3(x) = \frac{1}{2}(5x^3 - 3x)$, $P_4(x) = \frac{1}{8}(35x^4 - 30x^2 + 3)$ and $P_5(x) = \frac{1}{8}(63x^5 - 70x^3 + 15x)$.

The Legendre equation (5.3.1) can be written as a regular singular Sturm-Liouville problem. More precisely, (5.3.1) is

$$\frac{d}{dx} \left((1-x^2) \frac{dy}{dx} \right) + \lambda y = 0 .$$

This is (5.2.1) with $r(x) = 1 - x^2$, $q(x) = 0$, $p(x) = 1$, $a = -1$ and $b = 1$. If we require that y be bounded on the interval $] -1, 1[$, we get a regular singular Sturm-Liouville problem whose eigenvalues are $\lambda_n = n(n+1)$ for $n \geq 0$, and where the Legendre polynomial P_n is an eigenfunction associated to the eigenvalue λ_n .

It follows from the theory for the Sturm-Liouville problem that the Legendre Polynomials are orthogonal with respect to the scalar product

$$\langle f, g \rangle = \int_{-1}^1 f(x)g(x) dx$$

and the set $\{P_n : n \geq 0\}$ is complete in $L^2[-1, 1]$. More precisely, the Legendre polynomials satisfy the orthogonality relations

$$\int_{-1}^1 P_n(x)P_m(x) dx = \begin{cases} 0 & \text{if } n \neq m \\ \frac{2}{2n+1} & \text{if } n = m \end{cases}$$

We also have the following result.

Proposition 5.3.2

The L^2 -norm of P_n is

$$\| P_n \| = \left(\int_{-1}^1 (P_n(x))^2 dx \right)^{1/2} = \left(\frac{2}{2n+1} \right)^{1/2} .$$

Proof.

Let $I = \|P_n\|^2$. Using Rodrique's formula, we have

$$I = \frac{(-1)^n}{2^n n!} \int_{-1}^1 P_n(x) \frac{d^n}{dx^n} (1-x^2)^n dx .$$

Using integration by parts with $f(x) = P_n(x)$ and $g'(x) = \frac{d^n}{dx^n}(1-x^2)^n$, so $f'(x) = P'_n(x)$ and $g(x) = \frac{d^{n-1}}{dx^{n-1}}(1-x^2)^n$, we get

$$\begin{aligned} I &= \frac{(-1)^n}{2^n n!} \left(P_n(x) \frac{d^{n-1}}{dx^{n-1}}(1-x^2)^n \Big|_{-1}^1 - \int_{-1}^1 P'_n(x) \frac{d^{n-1}}{dx^{n-1}}(1-x^2)^n dx \right) \\ &= \frac{(-1)^{n+1}}{2^n n!} \int_{-1}^1 P'_n(x) \frac{d^{n-1}}{dx^{n-1}}(1-x^2)^n dx \end{aligned}$$

because $\frac{d^k}{dx^k}(1-x^2)^n = 0$ at $x = 1$ and $x = -1$ for $k < n$. Using integration by parts as above, the reader can easily show by induction that

$$I = \frac{1}{2^n n!} \int_{-1}^1 P_n^{(n)}(x)(1-x^2) dx .$$

Since

$$P_n^{(n)}(x) = \frac{d^n}{dx^n} \left(\frac{(-1)^n}{2^n n!} \frac{d^n}{dx^n}(1-x^2)^n \right) = \frac{1}{2^n n!} \frac{d^{2n}}{dx^{2n}}(x^2-1)^n = \frac{(2n)!}{2^n n!} ,$$

we get

$$I = \frac{(2n)!}{(2^n n!)^2} \int_{-1}^1 (1-x^2) dx = \frac{2(2n)!}{(2^n n!)^2} \int_0^1 (1-x^2) dx .$$

To compute the integral $\int_0^1 (1-x^2)^n dx$, we use the trigonometric substitution $x = \sin(t)$ for $0 \leq t \leq \pi/2$.

$$\int_0^1 (1-x^2)^n dx = \int_0^{\pi/2} \cos^{2n+1}(t) dt .$$

Using integration by parts with $f(t) = \cos^{2n}(t)$ and $g'(t) = \cos(t)$, so $f'(t) = -2n \cos^{2n-1}(t) \sin(t)$ and $g(t) = \sin(t)$, we get

$$\begin{aligned} \int_0^{\pi/2} \cos^{2n+1}(t) dt &= \cos^{2n}(t) \sin(t) \Big|_0^{\pi/2} + 2n \int_0^{\pi/2} \cos^{2n-1} \sin^2(t) dt \\ &= 2n \int_0^{\pi/2} \cos^{2n-1}(t) dt - 2n \int_0^{\pi/2} \cos^{2n+1}(t) dt . \end{aligned}$$

Thus,

$$\int_0^1 (1-x^2)^n dx = \int_0^{\pi/2} \cos^{2n+1}(t) dt = \frac{2n}{2n+1} \int_0^{\pi/2} \cos^{2n-1} dt .$$

Repeating this procedure $n-1$ times yields

$$\int_0^1 (1-x^2)^n dx = \frac{(2n)(2n-2)\cdots 2}{(2n+1)(2n-1)\cdots 3} \underbrace{\int_0^{\pi/2} \cos dt}_{=1} = \frac{(2^n n!)^2}{(2n+1)!} .$$

Finally, we have that

$$I = \frac{2(2n)!}{(2^n n!)^2} \left(\frac{(2^n n!)^2}{(2n+1)!} \right) = \frac{2}{2n+1} . \quad \blacksquare$$

The Fourier-Legendre expansion of f on $[-1, 1]$ is

$$f = \sum_{n=0}^{\infty} a_n P_n \quad (5.3.2)$$

where $a_n = \frac{2n+1}{2} \int_{-1}^1 f(x) P_n(x) dx$ for $n \geq 0$. The series in (5.3.2) converges in $L^2[-1, 1]$.

With the substitution $x = \cos(\phi)$ for $0 \leq \phi \leq \pi$, we get the complete orthogonal set of functions $\{P_n(\cos(\phi)) : n \geq 0\}$ in $L^2[0, \pi]$, where the scalar product is given by

$$\langle f, g \rangle = \int_0^\pi f(\phi) g(\phi) \sin(\phi) d\phi.$$

The Fourier-Legendre expansion of f on $[0, \pi]$ with respect to this complete orthogonal set of functions is

$$\tilde{f} \equiv f \circ \cos = \sum_{n=0}^{\infty} a_n P_n \circ \cos \quad (5.3.3)$$

where

$$a_n = \frac{2n+1}{2} \int_0^\pi \tilde{f}(\phi) P_n(\cos(\phi)) \sin(\phi) d\phi = \frac{2n+1}{2} \int_{-1}^1 f(x) P_n(x) dx$$

for $n \geq 0$. The series in (5.3.3) converges in $L^2[0, \pi]$.

Example 5.3.3

Express in terms of the Legendre polynomials the function $f(x) = |x|$ for $-1 \leq x \leq 1$.

The coefficient are given by $a_n = \frac{2n+1}{2} \int_{-1}^1 |x| P_n(x) dx$ for $n \geq 0$. Since $P_n(x)$ is an odd function when n is odd, we have that $|x|P_n(x)$ is odd when n is odd and so $a_n = 0$ for n odd. It is a little bit harder to find a general formulae for the coefficients a_n with n even. We have

$$\begin{aligned} a_0 &= \frac{1}{2} \int_{-1}^1 |x| dx = \int_0^1 x dx = \frac{1}{2}, \\ a_2 &= \frac{5}{2} \int_{-1}^1 |x| \frac{3x^2 - 1}{2} dx = \frac{5}{2} \int_0^1 (3x^3 - x) dx = \frac{5}{8}, \\ a_4 &= \frac{9}{2} \int_{-1}^1 |x| \frac{35x^4 - 30x^2 + 3}{8} dx = \frac{9}{8} \int_0^1 (35x^5 - 30x^3 + 3x) dx = -\frac{3}{16}, \end{aligned}$$

and so on. If we use Rodrigue's formula, we find that

$$a_{2m} = \frac{(-1)^{2m}(4m+1)}{2^{2m+1}(2m)!} \int_{-1}^1 |x| \frac{d^{2m}}{dx^{2m}} (1-x^2)^{2m} dx = \frac{(4m+1)}{2^{2m}(2m)!} \int_0^1 x \frac{d^{2m}}{dx^{2m}} (1-x^2)^{2m} dx$$

for $m \geq 0$. If we assume that $m > 0$, we may use integration by parts with $f(x) = x$ and $g'(x) = \frac{d^{2m}}{dx^{2m}} (1-x^2)^{2m}$, so $f'(x) = 1$ and $g(x) = \frac{d^{2m-1}}{dx^{2m-1}} (1-x^2)^{2m}$, to get

$$a_{2m} = \frac{(4m+1)}{2^{2m}(2m)!} \left(x \frac{d^{2m-1}}{dx^{2m-1}} (1-x^2)^{2m} \Big|_0^1 - \int_0^1 \frac{d^{2m-1}}{dx^{2m-1}} (1-x^2)^{2m} dx \right)$$

$$= -\frac{(4m+1)}{2^{2m}(2m)!} \frac{d^{2m-2}}{dx^{2m-2}} (1-x^2)^{2m} \Big|_0^1 = \frac{(4m+1)}{2^{2m}(2m)!} \frac{d^{2m-2}}{dx^{2m-2}} (1-x^2)^{2m} \Big|_{x=0},$$

where the second and third equality come from $\frac{d^k}{dx^k} (1-x^2)^{2m} \Big|_{x=1} = 0$ for $k < 2m$. The value of $\frac{d^{2m-2}}{dx^{2m-2}} (1-x^2)^{2m}$ at $x=0$ is given by $(2m-2)!$ times the coefficient of x^{2m-2} in

$$(1-x^2)^{2m} = \sum_{k=0}^{2m} \frac{(2m)!}{(2m-k)!k!} (-x^2)^k.$$

This coefficient is associated to k such that $2k = 2m-2$; namely, $k = m-1$. Therefore

$$a_{2m} = \frac{(4m+1)}{2^{2m}(2m)!} \left(\frac{(-1)^{m-1}(2m)!}{(m+1)!(m-1)!} \right) (2m-2)! = \frac{(-1)^{m-1}(4m+1)(2m-2)!}{2^{2m}(m+1)!(m-1)!}$$

for $m > 0$. Hence

$$|x| = \frac{1}{2}P_0(x) + \frac{5}{8}P_2(x) - \frac{3}{16}P_4(x) + \dots = \frac{1}{2}P_0(x) + \sum_{m=1}^{\infty} \frac{(-1)^{m-1}(4m+1)(2m-2)!}{2^{2m}(m+1)!(m-1)!} P_{2m}(x).$$

♣

5.4 Associated Legendre Polynomials $P_n^m(x)$ on $]-1, 1[$

The **associated Legendre equation** is the ordinary differential equation

$$(1-x^2) \frac{d}{dx} \left((1-x^2) \frac{dy}{dx} \right) + ((1-x^2)\lambda - m^2)y = 0 \quad (5.4.1)$$

for $-1 < x < 1$. If we differentiate m times with respect to x the Legendre equation

$$\frac{d}{dx} \left((1-x^2) \frac{dz}{dx} \right) + \lambda z = 0 \quad (5.4.2)$$

for $-1 < x < 1$, we get the differential equation

$$(1-x^2) \frac{d^{m+2}z}{dx^{m+2}} - 2(m+1)x \frac{d^{m+1}z}{dx^{m+1}} + (\lambda - m(m+1)) \frac{d^m z}{dx^m} = 0$$

for $-1 < x < 1$. If we substitute $y = (1-x^2)^{m/2} \frac{d^m z}{dx^m}$, or equivalently $\frac{d^m z}{dx^m} = (1-x^2)^{-m/2} y$, in the previous equation and multiply by $(1-x^2)^{-m/2+1}$, we get (5.4.1). Solutions of (5.4.1) are therefore of the form $y = (1-x^2)^{m/2} \frac{d^m z}{dx^m}$, where z is a solution of the Legendre equation (5.4.2).

It can be shown that (5.4.1) has non-trivial bounded solution if and only if $\lambda = n(n+1)$ for $n \geq 0$. Moreover, when $\lambda = n(n+1)$ for some n fixed, the bounded solutions of (5.4.1) are linear combinations of the **associated Legendre polynomials** defined by

$$P_n^m(x) = (1-x^2)^{m/2} \frac{d^m P_n}{dx^m}(x)$$

for $0 \leq m \leq n$. Note that m cannot be greater than n because P_n is a polynomial of degree n and thus its derivative of order greater than n is null.

The associated Legendre polynomials will be useful in the study of Laplace equation in Chapter 9.

5.5 Gamma Function

The Gamma function is needed to define the Bessel functions in the next section. The use of the Gamma function in mathematics is not limited to the definition of the Bessel functions. In fact, the Gamma function is present in many fields of mathematics. The Gamma function plays an important role in mathematical analysis, in analytic number theory, etc.

The **gamma function** Γ is defined by

$$\Gamma(\nu) = \int_0^\infty e^{-t} t^{\nu-1} dt$$

for $\nu > 0$ and recursively by $\Gamma(\nu) = \Gamma(\nu+1)/\nu$ for $\nu < 0$, ν not a negative integer. The graph of the gamma function between -3 and 3 is given in Figure 5.1.

Some of the properties of the gamma function are $\Gamma(\nu+1) = \nu\Gamma(\nu)$ and $\Gamma(n+1) = n!$ for $n \geq 0$.

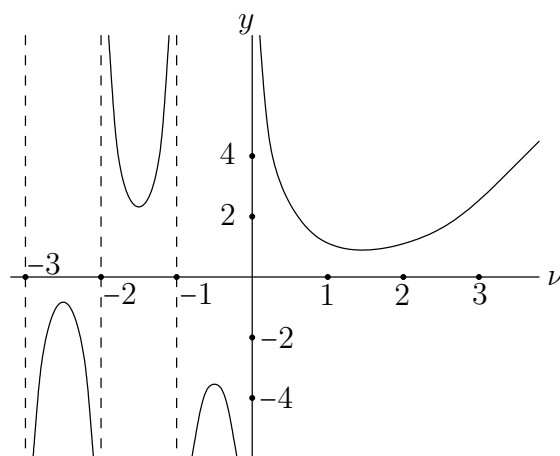


Figure 5.1: The graph of the gamma function $y = \Gamma(\nu)$

We note that $\Gamma(1/2) = \sqrt{\pi}$. This can easily be proved with a double integral. With the substitution $t = s^2$ in

$$\Gamma\left(\frac{1}{2}\right) = \int_0^\infty e^{-t} t^{-1/2} dt ,$$

we get

$$\Gamma\left(\frac{1}{2}\right) = 2 \int_0^\infty e^{-s^2} ds .$$

Hence,

$$\begin{aligned} \Gamma^2\left(\frac{1}{2}\right) &= 4 \left(\int_0^\infty e^{-u^2} du \right)^2 = 4 \left(\int_0^\infty e^{-u^2} du \right) \left(\int_0^\infty e^{-v^2} dv \right) \\ &= 4 \int_0^\infty \int_0^\infty e^{-u^2-v^2} du dv = 4 \int_0^{\pi/2} \int_0^\infty r e^{-r^2} dr d\theta = \pi , \end{aligned}$$

where we have used the polar coordinates $u = r \cos(\theta)$, $v = r \sin(\theta)$ for $0 \leq \theta \leq \pi/2$ and $r \geq 0$ to compute the double integral.

5.6 Bessel Functions

The **Bessel differential equation** is

$$x^2 y'' + x y' + (x^2 - \nu^2) y = 0 . \quad (5.6.1)$$

Using the Frobenius method at the regular singular point 0, we find the following solutions. If ν is not an integer, the functions

$$J_\nu(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k! \Gamma(\nu + k + 1)} \left(\frac{x}{2}\right)^{\nu+2k}$$

and $J_{-\nu}$ are two independent solutions of the Bessel equation. The functions J_ν are the **Bessel functions of the first kind**.

If $\nu = n$ is an integer,

$$J_n(x) = \frac{x^n}{2^n n!} \left(1 - \frac{x^2}{2^2 1! (n+1)} + \frac{x^4}{2^4 2! (n+1)(n+2)} - \frac{x^6}{2^6 3! (n+1)(n+2)(n+3)} + \dots \right)$$

and $J_{-n}(x) = (-1)^n J_n(x)$. So, J_n and J_{-n} are not linearly independent.

Let

$$Y_\nu(x) = \lim_{\rho \rightarrow \nu} \frac{J_\rho(x) \cos(\rho\pi) - J_{-\rho}(x)}{\sin(\rho\pi)} .$$

When ν is not an integer, we have

$$Y_\nu(x) = \frac{J_\nu(x) \cos(\nu\pi) - J_{-\nu}(x)}{\sin(\nu\pi)} .$$

When ν is an integer, $Y_\nu(x)$ can be computed using the Hospital rule. The functions Y_ν are the **Bessel functions of the second kind**.

If $\nu = n$ is an integer, a second solution of the Bessel differential equation, linearly independent of J_n , is given by $Y_n(x)$. Note that Y_ν is a solution of the Bessel differential equation, linearly independent of $J_\nu(x)$, even when ν is not an integer.

Example 5.6.1

Find the general solution of the ordinary differential equation

$$x^2 y'' + xy' + (\lambda^2 x^2 - \nu^2)y = 0 \quad (5.6.2)$$

by reducing it to a Bessel equation.

Let $z = \lambda x$ and $\tilde{y}(z) = y(x) = y\left(\frac{z}{\lambda}\right)$, then

$$\frac{dy}{dx} = \frac{d\tilde{y}}{dz} \frac{dz}{dx} = \lambda \frac{d\tilde{y}}{dz} \quad \text{and} \quad \frac{d^2 y}{dx^2} = \frac{d}{dx} \left(\frac{dy}{dx} \right) = \frac{d}{dz} \left(\lambda \frac{d\tilde{y}}{dz} \right) \frac{dz}{dx} = \lambda^2 \frac{d^2 \tilde{y}}{dz^2} .$$

To stick with the tradition, we will use y to denote $\tilde{y}$. The context will determine if we are working with $y(x)$ or $\tilde{y}(z)$. Hence, (5.6.2) becomes

$$0 = \left(\frac{z}{\lambda}\right)^2 \left(\lambda^2 \frac{d^2 y}{dz^2} \right) + \frac{z}{\lambda} \left(\lambda \frac{dy}{dz} \right) + \left(\lambda^2 \left(\frac{z}{\lambda}\right)^2 - \nu^2 \right) y = z^2 \frac{d^2 y}{dz^2} + z \frac{dy}{dz} + (z^2 - \nu^2)y .$$

This is a Bessel equation with the general solution $y(z) = AJ_\nu(z) + BY_\nu(z)$. Hence, the general solution of (5.6.2) is $y(x) = AJ_\nu(\lambda x) + BY_\nu(\lambda x)$ for arbitrary constants A and B . ♣

Example 5.6.2

Use the substitution $y(x) = \sqrt{x}u(x)$ and $z = 3x$ to solve (i.e. find the general solution of) the ordinary differential equation

$$x^2 y'' + \left(9x^2 + \frac{5}{36}\right)y = 0 . \quad (5.6.3)$$

Since $y' = u'x^{1/2} + ux^{-1/2}/2$ and $y'' = u''x^{1/2} + u'x^{-1/2} - ux^{-3/2}/4$, (5.6.3) becomes

$$0 = x^2 \left(u''x^{1/2} + u'x^{-1/2} - \frac{1}{4}ux^{-3/2} \right) + \left(9x^2 + \frac{5}{36} \right) ux^{1/2} = x^{5/2}u'' + x^{3/2}u' + \left(9x^2 - \frac{1}{9} \right) ux^{1/2} .$$

Hence, after multiplying both sides of the previous equality by $x^{1/2}$, we get

$$x^2 u'' + xu' + \left(9x^2 - \frac{1}{9} \right) u = 0 .$$

This is almost a Bessel equation but for the coefficient 9 in front of x^2 in the parenthesis. Let $z = 3x$, We have

$$\frac{du}{dx} = \frac{du}{dz} \frac{dz}{dx} = 3 \frac{du}{dz} \quad \text{and} \quad \frac{d^2 u}{dx^2} = \frac{d}{dx} \left(\frac{du}{dx} \right) = \frac{d}{dz} \left(3 \frac{du}{dz} \right) \frac{dz}{dx} = 9 \frac{d^2 u}{dz^2} .$$

Hence,

$$\begin{aligned} x^2 u'' + x u' + \left(9x^2 - \frac{1}{9}\right) u &= 9 \left(\frac{z}{3}\right)^2 \frac{d^2 u}{dz^2} + 3 \left(\frac{z}{3}\right) \frac{du}{dz} + \left(9 \left(\frac{z}{3}\right)^2 - \frac{1}{9}\right) u \\ &= z^2 \frac{d^2 u}{dz^2} + z \frac{du}{dz} + \left(z^2 - \frac{1}{9}\right) u = 0 . \end{aligned}$$

This is the Bessel equation with $\nu = 1/3$. The general solution of this Bessel equation is $u(z) = A J_{1/3}(z) + B J_{-1/3}(z)$. Therefore, the general solution of (5.6.3) is $y(x) = \sqrt{x} u(x) = A \sqrt{x} J_{1/3}(3x) + B \sqrt{x} J_{-1/3}(3x)$ for arbitrary constants A and B . ♣

Example 5.6.3

Find the general solution of the ordinary differential equation

$$y'' + k^2 x^2 y = 0 \tag{5.6.4}$$

for $x > 0$.

We use the substitution $y(x) = u(x)\sqrt{x}$ and $z = kx^2/2$. Since $y' = u'x^{1/2} + ux^{-1/2}/2$ and $y'' = u''x^{1/2} + u'x^{-1/2} - ux^{-3/2}/4$, we get

$$0 = \left(u''x^{1/2} + u'x^{-1/2} - \frac{1}{4}ux^{-3/2}\right) + k^2 x^2 (ux^{1/2}) = x^{1/2}u'' + x^{-1/2}u' + \left(k^2 x^{5/2} - \frac{1}{4}x^{-3/2}\right)u$$

from (5.6.4). Thus,

$$u'' + x^{-1}u' + \left(k^2 x^2 - \frac{1}{4}x^{-2}\right)u = 0 .$$

Since

$$\frac{du}{dx} = \frac{du}{dz} \frac{dz}{dx} = kx \frac{du}{dz} = \sqrt{2kz} \frac{du}{dz}$$

and

$$\begin{aligned} \frac{d^2 u}{dx^2} &= \frac{d}{dx} \left(\frac{du}{dx} \right) = \frac{d}{dz} \left(\sqrt{2kz} \frac{du}{dz} \right) \frac{dz}{dx} = \sqrt{2k} \left(\frac{1}{2} z^{-1/2} \frac{du}{dz} + z^{1/2} \frac{d^2 u}{dz^2} \right) \sqrt{2kz} \\ &= 2k \left(\frac{1}{2} \frac{du}{dz} + z \frac{d^2 u}{dz^2} \right) , \end{aligned}$$

where again we have used y to denote $\tilde{y}$ defined by $\tilde{y}(z) = y(x) = y(\sqrt{2z/k})$, we get

$$\begin{aligned} 0 &= 2k \left(\frac{1}{2} \frac{du}{dz} + z \frac{d^2 u}{dz^2} \right) + \sqrt{\frac{k}{2z}} \left(\sqrt{2kz} \frac{du}{dz} \right) + \left(k^2 \left(\frac{2z}{k} \right) - \frac{1}{4} \left(\frac{k}{2z} \right) \right) u \\ &= 2kz \frac{d^2 u}{dz^2} + 2k \frac{du}{dz} + \left(2kz - \frac{k}{8z} \right) u . \end{aligned}$$

Thus,

$$z^2 \frac{d^2 u}{dz^2} + z \frac{du}{dz} + \left(z^2 - \frac{1}{16} \right) u = 0 .$$

This is a Bessel equation with $\nu = 1/4$. Its general solution is $u = AJ_{1/4}(z) + BJ_{-1/4}(z)$. Hence, the general solution of (5.6.4) is $y(x) = \sqrt{x}u(x) = A\sqrt{x}J_{1/4}(kx^2/2) + B\sqrt{x}J_{-1/4}(kx^2/2)$ for arbitrary constants A and B . ♣

We have the following derivative rules for the Bessel functions of the first kind.

Proposition 5.6.4

$$(x^n J_n(x))' = x^n J_{n-1}(x) \quad (5.6.5)$$

and

$$(x^{-n} J_n(x))' = -x^{-n} J_{n+1}(x) \quad (5.6.6)$$

for $n \geq 0$.

Proof.

It is easy to prove (5.6.5) and (5.6.6) using the power series expansion of the Bessel function of the first kind. From

$$J_n(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k! \Gamma(n+k+1)} \left(\frac{x}{2}\right)^{n+2k},$$

we get

$$\begin{aligned} \frac{d}{dx}(x^n J_n(x)) &= \frac{d}{dx} \left(\sum_{k=0}^{\infty} \frac{(-1)^k x^{2n+2k}}{k! \Gamma(n+k+1) 2^{n+2k}} \right) = \sum_{k=0}^{\infty} \frac{(-1)^k (2n+2k) x^{2n+2k-1}}{k! \Gamma(n+k+1) 2^{n+2k}} \\ &= \sum_{k=0}^{\infty} \frac{(-1)^k x^{2n+2k-1}}{k! \Gamma(n+k) 2^{n+2k-1}} = x^n \sum_{k=0}^{\infty} \frac{(-1)^k x^{2n+2k-1}}{k! \Gamma(n+k) 2^{n+2k-1}} = x^n J_{n-1}(x) \end{aligned}$$

and

$$\frac{d}{dx}(x^{-n} J_n(x)) = \frac{d}{dx} \left(\sum_{k=0}^{\infty} \frac{(-1)^k x^{2k}}{k! \Gamma(n+k+1) 2^{n+2k}} \right) = \sum_{k=1}^{\infty} \frac{(-1)^k (2k) x^{2k-1}}{k! \Gamma(n+k+1) 2^{n+2k}}.$$

If we substitute $k = s+1$ in the last equation, we get

$$\begin{aligned} \frac{d}{dx}(x^{-n} J_n(x)) &= \sum_{s=0}^{\infty} \frac{(-1)^{s+1} x^{2s+1} (2s+2)}{(s+1)! \Gamma(n+s+2) 2^{n+2s+2}} = -x^{-n} \sum_{s=0}^{\infty} \frac{(-1)^s x^{2s+n+1}}{s! \Gamma(n+s+2) 2^{n+2s+1}} \\ &= -x^{-n} J_{n+1}(x). \quad \blacksquare \end{aligned}$$

Example 5.6.5

Compute the following integrals.

$$\text{a) } \int x^4 J_1(x) dx \quad \text{b) } \int x^3 J_0(x) dx \quad \text{c) } \int x J_2(x) dx$$

a) To compute this integral, we use integration by parts and (5.6.5). Hence,

$$\int x^4 J_1(x) dx = \int x^2 (x^2 J_1(x)) dx$$

$$\begin{aligned}
&= x^4 J_2(x) - \int 2x^3 J_2(x) \, dx \quad \left\{ \begin{array}{l} \text{If we use integration by parts with} \\ f(x) = x^2 \text{ and } g'(x) = x^2 J_1(x), \text{ so} \\ f'(x) = 2x \text{ and } g(x) = x^2 J_2(x) \text{ from} \\ (5.6.5) \text{ with } n = 2. \end{array} \right. \\
&= x^4 J_2(x) - 2 \int x^3 J_2(x) \, dx \\
&= x^4 J_2(x) - 2x^3 J_3(x) + C \quad \left\{ \begin{array}{l} \text{From (5.6.5) with } n = 3. \end{array} \right.
\end{aligned}$$

b) To compute this integral, we again use integration by parts and (5.6.5). Hence,

$$\begin{aligned}
\int x^3 J_0(x) \, dx &= \int x^2 (x J_0(x)) \, dx \\
&= -x J_1(x) - \int 2x^2 J_1(x) \, dx \quad \left\{ \begin{array}{l} \text{If we use integration by parts with} \\ f(x) = x^2 \text{ and } g'(x) = x J_0(x), \text{ so} \\ f'(x) = 2x \text{ and } g(x) = x J_1(x) \text{ from} \\ (5.6.5) \text{ with } n = 1. \end{array} \right. \\
&= x^3 J_1(x) - 2 \int x^2 J_1(x) \, dx \\
&= x^3 J_1(x) - 2x^2 J_2(x) + C \quad \left\{ \begin{array}{l} \text{From (5.6.5) with } n = 2. \end{array} \right.
\end{aligned}$$

c) To compute this integral, we use integration by parts and (5.6.6). Hence,

$$\begin{aligned}
\int x J_2(x) \, dx &= \int x^2 (x^{-1} J_2(x)) \, dx \\
&= -x J_1(x) + \int 2J_1(x) \, dx \quad \left\{ \begin{array}{l} \text{If we use integration by parts with} \\ f(x) = x^2 \text{ and } g'(x) = x^{-1} J_2(x), \text{ so} \\ f'(x) = 2x \text{ and } g(x) = -x^{-1} J_1(x) \text{ from} \\ (5.6.6) \text{ with } n = 1. \end{array} \right. \\
&= -x J_1(x) + 2 \int J_1(x) \, dx \\
&= -x J_1(x) - 2J_0(x) + C \quad \left\{ \begin{array}{l} \text{From (5.6.6) with } n = 0. \end{array} \right.
\end{aligned}$$

♣

The Bessel functions of the first kind satisfy the following recurrence relations.

$$nJ_n(x) + xJ'_n(x) = xJ_{n-1}(x) , \quad (5.6.7)$$

$$nJ_n(x) - xJ'_n(x) = xJ_{n+1}(x) , \quad (5.6.8)$$

$$J_{n-1}(x) + J_{n+1}(x) = \frac{2n}{x} J_n(x) \quad (5.6.9)$$

and

$$J_{n-1}(x) - J_{n+1}(x) = 2J'_n(x) . \quad (5.6.10)$$

(5.6.7) and (5.6.8) comes from expanding (5.6.5) and (5.6.6) respectively. The sum of (5.6.7) and (5.6.8) yields (5.6.9), and the difference of (5.6.7) and (5.6.8) yields (5.6.10).

Two special Bessel functions of the first kind are $J_{1/2}(x) = \sqrt{2/(\pi x)} \sin(x)$ and $J_{-1/2}(x) = \sqrt{2/(\pi x)} \cos(x)$.

Some of the zeros $\alpha_{n,0}$ of J_0 with the value of J_1 at these points, and some of the zeros $\alpha_{n,1}$ of J_1 with the value of J_0 at these points, are listed in the following table. The values are rounded to four decimals.

n	$\alpha_{n,0}$	$J_1(\alpha_{n,0})$	$\alpha_{n,1}$	$J_0(\alpha_{n,1})$
0			0.0000	1.0000
1	2.4048	0.5191	3.8317	-0.4028
2	5.5201	-0.3403	7.0156	0.3001
3	8.6537	0.2715	10.1735	-0.2497
4	11.7915	-0.2325	13.3237	0.2184
5	14.9309	0.2065	16.4706	-0.1965
6	18.0711	-0.1877	19.6159	0.1801
7	21.2116	0.1733		

The Bessel equation (5.6.1) can be associated to a regular singular Sturm-Liouville problem. More precisely, the substitution $x = \sqrt{\lambda} t$ and $\nu = k \in \mathbb{N}$, transforms (5.6.1) into

$$\frac{d}{dt} \left(t \frac{dy}{dt} \right) + \left(-\frac{k^2}{t} + \lambda t \right) y = 0 \quad (5.6.11)$$

for $0 < t < R$, where R is determined by the problem under consideration. This is (5.2.1) with $r(t) = t$, $q(t) = -k^2/t$, $p(t) = t$. $a = 0$ and $b = R$. If we require that y be bounded near 0 and $y(R) = 0$, we get a regular singular Sturm-Liouville problem whose eigenvalues are $\lambda = \lambda_{n,k} = (\alpha_{n,k}/R)^2$, where $n > 0$ and $\alpha_{n,k}$ is the n^{th} zero of $J_k(x)$. The eigenfunctions associated to $\lambda_{n,k}$ are multiple of $J_k(\alpha_{n,k}t/R)$. This comes from the fact that the bounded solution of (5.6.11) on $[0, R]$ are scalar multiple of $y(t) = J_k(\sqrt{\lambda}t)$ and $y(R) = 0$ implies that $\sqrt{\lambda}R = \alpha_{n,k}$. We ignore $n = 0$ for $k > 0$ because $J_k(\alpha_{0,k}t/R) = J_k(0) = 0$ for all t cannot be an eigenfunction. The set $\{J_k(\alpha_{n,k}t/R) : n \geq 1\}$ is complete and orthogonal in $L^2[0, R]$, where the scalar product is defined by

$$\langle f, g \rangle = \int_0^R f(t)g(t) t \, dt .$$

The orthogonality is a consequence of the theory for the Sturm-Liouville problems. More precisely, the Bessel functions of the first kind satisfy the following orthogonality relations.

Proposition 5.6.6

$$\int_0^R J_k \left(\frac{\alpha_{i,k}t}{R} \right) J_k \left(\frac{\alpha_{j,k}t}{R} \right) t \, dt = \begin{cases} 0 & \text{if } i \neq j \\ \frac{R^2}{2} J_{k+1}^2(\alpha_{i,k}) & \text{if } i = j \end{cases}$$

Proof.

The orthogonality comes from the theory of the Sturm-Liouville problems. We only need to prove that

$$\int_0^R J_k^2\left(\frac{\alpha_{n,k}t}{R}\right) t \, dt = \frac{R^2}{2} J_{k+1}^2(\alpha_{n,k}) . \quad (5.6.12)$$

We use the fact that $y(t) = J_k(\gamma_{n,k}t)$ with $\gamma_{n,k} = \alpha_{n,k}/R$ is a solution of (5.6.11), the Bessel equation in the Sturm-Liouville form with the conditions y bounded at the origin and $y(R) = 0$. More specifically, we have

$$\frac{d}{dt}\left(t \frac{d}{dt}(J_k(\gamma_{n,k}t))\right) + \left(-\frac{k^2}{t} + \gamma_{n,k}^2 t\right) J_k(\gamma_{n,k}t) = 0 .$$

Multiplying both sides by $2t \frac{d}{dt}(J_k(\gamma_{n,k}t))$, we get

$$2t \frac{d}{dt}(J_k(\gamma_{n,k}t)) \frac{d}{dt}\left(t \frac{d}{dt}(J_k(\gamma_{n,k}t))\right) + 2(-k^2 + \gamma_{n,k}^2 t^2) J_k(\gamma_{n,k}t) \frac{d}{dt}(J_k(\gamma_{n,k}t)) = 0 .$$

Namely,

$$\frac{d}{dt}\left(t \frac{d}{dt}(J_k(\gamma_{n,k}t))\right)^2 + (-k^2 + \gamma_{n,k}^2 t^2) \frac{d}{dt} J_k^2(\gamma_{n,k}t) = 0 .$$

Hence,

$$\begin{aligned} \left(t \frac{d}{dt}(J_k(\gamma_{n,k}t))\right)^2 \Big|_0^R &= \int_0^R \frac{d}{dt}\left(t \frac{d}{dt}(J_k(\gamma_{n,k}t))\right)^2 dt \\ &= \int_0^R (k^2 - \gamma_{n,k}^2 t^2) \frac{d}{dt} J_k^2(\gamma_{n,k}t) dt . \end{aligned} \quad (5.6.13)$$

We first compute the left hand side of (5.6.13). From (5.6.8) with $x = \gamma_{n,k}t$, we get

$$t \frac{d}{dt}(J_k(\gamma_{n,k}t)) = \gamma_{n,k}t J'_k(\gamma_{n,k}t) = k J_k(\gamma_{n,k}t) - \gamma_{n,k}t J_{k+1}(\gamma_{n,k}t) .$$

Therefore,

$$\left(t \frac{d}{dt}(J_k(\gamma_{n,k}t))\right)^2 \Big|_0^R = (kJ_k(\gamma_{n,k}t) - \gamma_{n,k}t J_{k+1}(\gamma_{n,k}t))^2 \Big|_0^R = \alpha_{n,k}^2 J_{k+1}^2(\alpha_{n,k})$$

because $\gamma_{n,k}R = \alpha_{n,k}$, $J_k(\alpha_{n,k}) = 0$ for all n , and $J_k(0) = 0$ for $k > 0$.

We now turn our attention on the right hand side of (5.6.13). Using integration by parts with $f(t) = (k^2 - \gamma_{n,k}^2 t^2)$ and $g'(t) = \frac{d}{dt}(J_k^2(\gamma_{n,k}t))$, so $f'(t) = -2\gamma_{n,k}^2 t$ and $g(t) = J_k^2(\gamma_{n,k}t)$, we get

$$\int_0^R (k^2 - \gamma_{n,k}^2 t^2) \frac{d}{dt}(J_k^2(\gamma_{n,k}t)) dt = (k^2 - \gamma_{n,k}^2 t^2) J_k^2(\gamma_{n,k}t) \Big|_0^R + \int_0^R 2\gamma_{n,k}^2 t J_k^2(\gamma_{n,k}t) dt$$

$$= 2\gamma_{n,k}^2 \int_0^R t J_k^2(\gamma_{n,k}t) dt = 2 \frac{\alpha_{n,k}^2}{R^2} \int_0^R J_k^2(\gamma_{n,k}t) t dt ,$$

where again we have used $J_k(\alpha_{n,k}) = 0$ for all n and $J_k(0) = 0$ for $k > 0$.

It follows from (5.6.13) that

$$\alpha_{n,k}^2 J_{k+1}^2(\alpha_{n,k}) = 2 \frac{\alpha_{n,k}^2}{R^2} \int_0^R J_k^2(\gamma_{n,k}t) t dt ,$$

and this gives (5.6.12) after simplification. ■

Hence, the norm of $J_k(\alpha_{n,k}t/R)$ on $[0, R]$ is given by

$$\left\| J_k \left(\frac{\alpha_{n,k}t}{R} \right) \right\|^2 = \int_0^R J_k^2 \left(\frac{\alpha_{n,k}t}{R} \right) t dt = \frac{R^2}{2} J_{k+1}^2(\alpha_{n,k}) .$$

For k fix, the Fourier-Bessel expansion of f on $]0, R[$ with respect to the functions $J_k(\alpha_{n,k}t/R)$ for $n \geq 1$ is

$$f = \sum_{n=1}^{\infty} a_n J_k \left(\frac{\alpha_{n,k}t}{R} \right)$$

for $0 < t < R$, where

$$a_n = \frac{2}{R^2 J_{k+1}^2(\alpha_{n,k})} \int_0^R f(t) J_k \left(\frac{\alpha_{n,k}t}{R} \right) t dt$$

for $n \geq 1$. The convergence of the series is in $L^2[0, R]$.

Example 5.6.7

Express $f(x) = x^2$ for $0 < x < R$ in terms of the Bessel functions $\{J_0(\gamma_{n,0}x)\}_{n=0}^{\infty}$, where $\gamma_{n,0} = \alpha_{n,0}/R$.

The coefficient of the Fourier-Bessel series are

$$\begin{aligned} a_n &= \frac{2}{R^2 J_1^2(\alpha_{n,0})} \int_0^R t^3 J_0(\gamma_{n,0}t) dt = \frac{2}{R^2 J_1^2(\alpha_{n,0})} \left(\frac{1}{\gamma_{n,0}^4} \int_0^{\alpha_{n,0}} s^3 J_0(s) ds \right) \\ &= \frac{2}{R^2 J_1^2(\alpha_{n,0}) \gamma_{n,0}^4} \left(s^3 J_1(s) - 2s^2 J_2(s) \right) \Big|_0^{\alpha_{n,0}} \\ &= \frac{2R^2}{J_1^2(\alpha_{n,0}) \alpha_{n,0}^4} \left(\alpha_{n,0}^3 J_1(\alpha_{n,0}) - 2\alpha_{n,0}^2 J_2(\alpha_{n,0}) \right) \end{aligned}$$

The second equality comes from the substitution $s = \gamma_{n,0}t$ in the integral. The integral of $s^3 J_0(s)$ computed in Example 5.6.5 (b) gives the third equality. The expression for a_n can be further simplified using (5.6.9) with $n = 1$. We get

$$J_2(\alpha_{n,0}) = J_0(\alpha_{n,0}) + J_2(\alpha_{n,0}) = \frac{2}{\alpha_{n,0}} J_1(\alpha_{n,0}) .$$

Thus

$$a_n = \frac{2R^2}{J_1^2(\alpha_{n,0}) \alpha_{n,0}^4} \left(\alpha_{n,0}^3 J_1(\alpha_{n,0}) - 4\alpha_{n,0} J_1(\alpha_{n,0}) \right) = \frac{2R^2}{\alpha_{n,0}^3 J_1(\alpha_{n,0})} \left(\alpha_{n,0}^2 - 4 \right) .$$

♣

5.7 Exercises

Suggested exercises:

- In [24]: numbers 6.1 to 6.9 in Section 6.7.
- in [38]: numbers 6, 7 and 13 in Section 10.5.
- In [33]: numbers 1, 2 and 8 in Section 26; all the numbers in Sections 27 to 29, and 32 to 35.

Chapter 6

Distributions and Strong Solutions

6.1 Introduction

Consider the linear partial differential equation

$$L(\mathbf{x}, D)u = \sum_{|\alpha| \leq m} a_\alpha(\mathbf{x}) D^\alpha u = f \quad (6.1.1)$$

where $u \in C^m(\mathbb{R}^n)$, $f \in C(\mathbb{R}^n)$ and $a_\alpha \in C^{|\alpha|}(\mathbb{R}^n)$ for all multi-indices involved in the sum.

If we multiply both sides of the equation in (6.1.1) by a function $\phi \in C_c^\infty(\mathbb{R}^n)$ (functions in $C^\infty(\mathbb{R}^n)$ with compact supports) and use integration by parts, we get

$$\begin{aligned} \int_{\mathbb{R}^n} f(\mathbf{x}) \phi(\mathbf{x}) d\mathbf{x} &= \sum_{|\alpha| \leq m} \int_{\mathbb{R}^n} a_\alpha(\mathbf{x}) D^\alpha u(\mathbf{x}) \phi(\mathbf{x}) d\mathbf{x} \\ &= \sum_{|\alpha| \leq m} (-1)^{|\alpha|} \int_{\mathbb{R}^n} u(\mathbf{x}) D^\alpha (a_\alpha(\mathbf{x}) \phi(\mathbf{x})) d\mathbf{x} . \end{aligned}$$

Conversely, if there exists a function $u \in C^m(\mathbb{R}^n)$ such that

$$\int_{\mathbb{R}^n} f(\mathbf{x}) \phi(\mathbf{x}) d\mathbf{x} = \sum_{|\alpha| \leq m} (-1)^{|\alpha|} \int_{\mathbb{R}^n} u(\mathbf{x}) D^\alpha (a_\alpha(\mathbf{x}) \phi(\mathbf{x})) d\mathbf{x} \quad (6.1.2)$$

for all functions $\phi \in C_c^\infty(\mathbb{R}^n)$, then integrations by parts allows us to rewrite this equation as

$$\int_{\mathbb{R}^n} f(\mathbf{x}) \phi(\mathbf{x}) d\mathbf{x} = \sum_{|\alpha| \leq m} \int_{\mathbb{R}^n} a_\alpha(\mathbf{x}) D^\alpha u(\mathbf{x}) \phi(\mathbf{x}) d\mathbf{x}$$

for all functions $\phi \in C_c^\infty(\mathbb{R}^n)$. Since these functions are dense in the space of continuous complex valued functions over $\mathbb{R}^n$, (6.1.2) implies that (6.1.1) is satisfied by u in the classical sense.

The previous discussion suggests the following approach to solve (6.1.1).

1. Find a function u which satisfies (6.1.2) for all function $\phi \in C_c^\infty(\mathbb{R}^n)$.

2. Show, if possible, that u found in (1) is in $C^m(\mathbb{R}^n)$.

Let $L^*(\mathbf{x}, D)$ be the linear partial differential operator defined by

$$L^*(\mathbf{x}, D)\phi = \sum_{|\alpha| \leq m} (-1)^{|\alpha|} D^\alpha (a_\alpha(\mathbf{x})\phi)$$

for all $\phi \in C_c^\infty(\mathbb{R}^n)$. The operator $L^*(\mathbf{x}, D)$ is called the **adjoint** of $L(\mathbf{x}, D)$. Hence, if $u \in C^m(\mathbb{R}^n)$ is a solution of (6.1.1), we have

$$\int_{\mathbb{R}^n} (L(\mathbf{x}, D)u(\mathbf{x}))\phi(\mathbf{x}) d\mathbf{x} = \int_{\mathbb{R}^n} u(\mathbf{x})(L^*(\mathbf{x}, D)\phi(\mathbf{x})) d\mathbf{x}$$

for all $\phi \in C_c^\infty(\mathbb{R}^n)$. This is reminiscing of the relation that we have in linear algebra between an $n \times n$ matrix A and its adjoint A^* ; namely, $\langle A\mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, A^*\mathbf{y} \rangle$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

A **classical solution** of (6.1.1) is a function $u \in C^m(\mathbb{R}^n)$ that satisfies (6.1.1). A **strong solution** or a **solution in the sense of distributions** of (6.1.1) is a measurable function $u : \mathbb{R}^n \rightarrow \mathbb{C}$ that satisfies (6.1.2). Sometime, only strong solutions of (6.1.1) can be found.

We should warn the reader that our use of strong solution is not universal. What we called strong solution is often called “weak solution” in the literature. However, we have chosen to reserve the expression “weak solution” for the solution of the “variational problem” associated to a partial differential equation as we will see in Chapter 12. The concept of Strong solution is more general than the concept of “weak solution” though they are often the same.

The simple idea of solving an integral equation like (6.1.2) instead of its associated differential equation like (6.1.1) has led to a lot of research in partial differential equations during the twentieth century. The **theory of distributions** that we briefly present in this chapter and the **Sobolev spaces** that we will study in a future chapter have come out of the research done in that century. The modern theory of numerical analysis for partial differential equations is based on the theory of distributions and the Sobolev spaces.

Good references for the material presented in this chapter are [27, 31]. You will find in these references the proofs of the results that we state without proofs in the presentation below.

6.2 Test Functions

From now on, unless otherwise stated, $\Omega \subset \mathbb{R}^n$ is an open set.

The space $C_c^\infty(\Omega)$ of infinitely differentiable functions with compact support in Ω is denoted $\mathcal{D}(\Omega)$. Recall that the support of a function $f : \Omega \rightarrow \mathbb{C}$, denoted $\text{supp } f$, is the set

$$\text{supp } f = \overline{\{\mathbf{x} \in \Omega : f(\mathbf{x}) \neq 0\}},$$

where $\overline{A}$ denotes the closure of the set $A \subset \Omega$. The functions in $\mathcal{D}(\Omega)$ are called the **test functions** on Ω .

The convergence (and so the topology) on $\mathcal{D}(\Omega)$ is defined as follows. We say that sequence $\{\phi_j\}_{j=0}^\infty$ **converges** toward ϕ in $\mathcal{D}(\Omega)$ and write $\phi_j \rightarrow \phi$ if

1. there exists a compact set K such that $\text{supp } \phi_j \subset K$ for all j and $\text{supp } \phi \subset K$.
2. $\lim_{j \rightarrow \infty} \|D^\alpha \phi_j - D^\alpha \phi\|_{\infty, K} = 0$ for all multi-indices $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n) \in \mathbb{N}^n$.

Recall that, for all continuous functions $\psi : \Omega \rightarrow \mathbb{C}$ and compact sets $K \subset \Omega$, $\|\psi\|_{\infty, K} = \max_{\mathbf{x} \in K} |\psi(\mathbf{x})|$.

A complete definition of the locally convex topology on $\mathcal{D}(\Omega)$ is presented in [31, 27]. We have only stated the definition of convergence in $\mathcal{D}(\Omega)$ that is convenient for the presentation of distribution theory.

6.3 Distributions

A **distribution** on Ω is a continuous linear functional on $\mathcal{D}(\Omega)$, where the continuity is defined as follows. An application $u : \mathcal{D}(\Omega) \rightarrow \mathbb{C}$ is continuous at $\phi \in \mathcal{D}(\Omega)$ if $\lim_{j \rightarrow \infty} |u(\phi_j) - u(\phi)| = 0$ for all sequences $\{\phi_j\}_{j=0}^\infty$ in $\mathcal{D}(\Omega)$ converging toward ϕ in $\mathcal{D}(\Omega)$. The space of all distributions on Ω is denoted $\mathcal{D}'(\Omega)$.

Let K be a compact subset of $\mathbb{R}^n$. The space $\mathcal{D}_K$ is defined by $\mathcal{D}_K = \{u \in \mathcal{D}(\mathbb{R}^n) : \text{supp } u \subset K\}$.

Proposition 6.3.1

A linear functional u on $\mathcal{D}(\Omega)$ is continuous if and only if for every compact set $K \subset \Omega$ there exist a constant C_K and an integer N_K such that $|u(\phi)| \leq C_K \max_{|\alpha| \leq N_K} \|D^\alpha \phi\|_{\infty, K}$ for all $\phi \in \mathcal{D}_K$.

The proof of this result is included in the proof of Theorem 6.8 in [31].

Remark 6.3.2

The following notation is often used. For $u \in \mathcal{D}'(\Omega)$ and $\phi \in \mathcal{D}(\Omega)$, we define $\langle u, \phi \rangle$ as $\langle u, \phi \rangle = u(\phi)$. We will use this notation occasionally. ♠

Example 6.3.3

Let u be a function in $L^1_{loc}(\mathbb{R}^n)$. The function u defines a distribution on $\mathbb{R}^n$ by

$$\langle u, \phi \rangle = u(\phi) = \int_{\mathbb{R}^n} u(\mathbf{x}) \phi(\mathbf{x}) d\mathbf{x} \quad , \quad \phi \in \mathcal{D}(\mathbb{R}^n) .$$

It is traditional to use the same letter to designate the function and the distribution. The context determines to which one we refer to.

To verify this statement, we have to show that $u : \mathcal{D}(\mathbb{R}^n) \rightarrow \mathbb{C}$ is well defined, u is linear, and u is continuous.

If $\phi \in \mathcal{D}(\mathbb{R}^n)$ and $K = \text{supp } \phi$, then

$$|u(\phi)| = \left| \int_{\mathbb{R}^n} u(\mathbf{x}) \phi(\mathbf{x}) \, d\mathbf{x} \right| \leq \int_K |u(\mathbf{x})| |\phi(\mathbf{x})| \, d\mathbf{x} \leq \|\phi(\mathbf{x})\|_{\infty, K} \int_K |u(\mathbf{x})| \, d\mathbf{x} < \infty$$

because $\|\phi(\mathbf{x})\|_{\infty, K} < \infty$ since ϕ is continuous on the compact set K , and $\int_K |u(\mathbf{x})| \, d\mathbf{x}$ is bounded since $u \in L^1_{loc}(\mathbb{R}^n)$. Thus, $u : \mathcal{D}(\mathbb{R}^n) \rightarrow \mathbb{C}$ is well defined.

$u : \mathcal{D}(\mathbb{R}^n) \rightarrow \mathbb{C}$ is linear because the integral is a linear mapping.

To show that $u : \mathcal{D}(\mathbb{R}^n) \rightarrow \mathbb{C}$ is continuous, let ϕ be any element of $\mathcal{D}(\mathbb{R}^n)$ and $\{\phi_j\}_{j=0}^\infty$ be any sequence $\mathcal{D}(\mathbb{R}^n)$ converging toward ϕ in $\mathcal{D}(\mathbb{R}^n)$. By definition of convergence in $\mathcal{D}(\mathbb{R}^n)$, there exists a compact set $K \subset \mathbb{R}^n$ such that $\text{supp } \phi_j \subset K$ for all j , $\text{supp } \phi \subset K$, and

$$\lim_{j \rightarrow \infty} \|D^\alpha \phi_j - D^\alpha \phi\|_{\infty, K} = 0$$

for all multi-indices $\alpha \in \mathbb{N}^n$. In particular, for $\alpha = \mathbf{0}$, we get

$$\lim_{j \rightarrow \infty} \|\phi_j - \phi\|_{\infty, K} = 0 .$$

Hence,

$$\begin{aligned} |u(\phi_j) - u(\phi)| &= \left| \int_{\mathbb{R}^n} u(\mathbf{x}) (\phi_j(\mathbf{x}) - \phi(\mathbf{x})) \, d\mathbf{x} \right| \leq \int_K |u(\mathbf{x})| |\phi_j(\mathbf{x}) - \phi(\mathbf{x})| \, d\mathbf{x} \\ &\leq \|\phi_j - \phi\|_{\infty, K} \int_K |u(\mathbf{x})| \, d\mathbf{x} \rightarrow 0 \end{aligned}$$

as $j \rightarrow \infty$. Note that $\int_K |u(\mathbf{x})| \, d\mathbf{x}$ is finite because $u \in L^1_{loc}(\mathbb{R}^n)$. ♣

Example 6.3.4

The **Dirac delta function** is the distribution on $\mathbb{R}^n$ defined by

$$\langle \delta, \phi \rangle = \delta(\phi) = \phi(\mathbf{0}) \quad , \quad \phi \in \mathcal{D}(\mathbb{R}^n) .$$

♣

Remark 6.3.5

The Dirac delta function is often introduced in engineering books as a function which is zero everywhere but at the origin and such that $\int_{-\infty}^\infty \delta(x) f(x) \, dx = f(0)$ for all functions f . Obviously, this does not make any sense. However, the formal manipulations done using this “definition” of the Dirac delta function can be rigorously proved with the correct definition of this function given in the previous example.

Since most of the distributions $u \in \mathcal{D}'(\mathbb{R})$ used in engineering comes from integrable functions, the integral notation $\int_{-\infty}^\infty u(x) \phi(x) \, dx$ is often used instead of $u(\phi)$ or $\langle u, \phi \rangle$ even when the distribution u is not an integrable function as is the case for the Dirac delta function. We will not use these notation in the present text. ♠

The next proposition states that we may defined the **product** of $u \in \mathcal{D}'(\Omega)$ and $f \in C^\infty(\Omega)$ as the distribution fu defined by

$$(fu)(\phi) = u(f\phi) \quad , \quad \phi \in \mathcal{D}(\Omega) .$$

Proposition 6.3.6

If $f \in C^\infty(\Omega)$ and $u \in \mathcal{D}'(\Omega)$, then

$$v(\phi) = u(f\phi) \quad , \quad \phi \in \mathcal{D}(\Omega) \quad ,$$

defines a distribution on Ω . We denote this distribution by fu , the product of the function f and the distribution u .

Proof.

i) $(fu)(\phi)$ is well defined for all $\phi \in \mathcal{D}(\Omega)$ because $f\phi \in \mathcal{D}(\Omega)$ for all $\phi \in \mathcal{D}(\Omega)$.

ii) It is easy to show that fu is a linear functional.

iii) To prove that fu is continuous, we show that fu satisfies Proposition 6.3.1.

Let K be a compact subset of Ω . Since $u \in \mathcal{D}'(\Omega)$, it follows from Proposition 6.3.1 that there exist a constant C_K and an integer N_K such that $|u(\phi)| \leq C_K \max_{|\alpha| \leq N_K} \|D^\alpha \phi\|_{\infty, K}$ for all $\phi \in \mathcal{D}_K$. According to Leibniz formula

$$D^\alpha(f\phi) = \sum_{\beta \leq \alpha} c_{\alpha, \beta} D^{\alpha-\beta} f D^\beta \phi \quad ,$$

where the coefficients $c_{\alpha, \beta}$ are constants. This is a formula similar to the binomial formula. Hence, there exists a constant C_f such that

$$\|D^\alpha(f\phi)\|_\infty \leq C_f \max_{\beta \leq \alpha} \|D^\beta \phi\|_{\infty, K}$$

for all $\phi \in \mathcal{D}_K$ and $|\alpha| \leq N_K$. The constant C_f depends only of $\|D^{\alpha-\beta} f\|_{\infty, K}$ and $c_{\alpha, \beta}$ for $\beta \leq \alpha$. We finally get that

$$|(fu)(\phi)| = |u(f\phi)| \leq C_K \max_{|\alpha| \leq N_K} \|D^\alpha(f\phi)\|_{\infty, K} \leq C_K C_f \max_{|\alpha| \leq N_K} \|D^\alpha \phi\|_{\infty, K}$$

for all $\phi \in \mathcal{D}_K$. ■

The **support** of a distribution $u \in \mathcal{D}'(\Omega)$ is the complement of the largest open set $V \subset \Omega$ such that $u(\phi) = 0$ for all $\phi \in \mathcal{D}(\Omega)$ with $\text{supp } \phi \subset V$. The space of distributions on Ω with compact support is denoted $\mathcal{E}'(\Omega)$; in other words, $\mathcal{E}'(\Omega) = \{u \in \mathcal{D}'(\Omega) : \text{supp } u \text{ is compact}\}$.

The space of distributions is equipped with the following (weak) topology. We say that the sequence $\{u_j\}_{j=0}^\infty$ of distribution on Ω **converge (weakly)** to $u \in \mathcal{D}'(\Omega)$ if $\lim_{j \rightarrow \infty} |u_j(\phi) - u(\phi)| = 0$ for all $\phi \in \mathcal{D}(\Omega)$. The space of distribution on Ω is “kind of complete” according to this topology.

Proposition 6.3.7

Let $\{u_j\}_{j=0}^\infty$ be a sequence in $\mathcal{D}'(\Omega)$ such that the sequence $\{u_j(\phi)\}_{j=0}^\infty$ of complex numbers converges for all $\phi \in \mathcal{D}(\Omega)$. Let $u : \mathcal{D}(\Omega) \rightarrow \mathbb{C}$ be the function defined by $u(\phi) = \lim_{j \rightarrow \infty} u_j(\phi)$ for all $\phi \in \mathcal{D}(\Omega)$. Then $u \in \mathcal{D}'(\Omega)$ and the sequence $\{u_j\}_{j=0}^\infty$ converges to u in $\mathcal{D}'(\Omega)$.

The proof of this result is part of the proof of Theorem 6.17 in [31].

Example 6.3.8

Consider the function

$$\chi_a(x) = \begin{cases} 1/(2a) & \text{if } |x| < a \\ 0 & \text{if } |x| > a \end{cases}$$

It is often stated in engineering books that the Dirac delta function δ is the limit of χ_a as $a \rightarrow 0$.

We can rigorously prove this statement in the context of our theory. We have that χ_a defined a distribution on $\mathbb{R}$ by

$$\chi_a(\phi) = \int_{\mathbb{R}} \chi_a(x) \phi(x) dx$$

for all $\phi \in \mathcal{D}(\mathbb{R})$. We now prove that this distribution converges to δ as $a \rightarrow 0$.

Given $\phi \in \mathcal{D}(\mathbb{R})$, let $\Phi(x) = \int_0^x \phi(x) dx$. We have that

$$\lim_{a \rightarrow 0} \chi_a(\phi) = \lim_{a \rightarrow 0} \frac{1}{2a} \int_{-a}^a \phi(x) dx = \lim_{a \rightarrow 0} \frac{1}{2a} (\Phi(a) - \Phi(-a)) = \Phi'(0) = \phi(0) = \delta(\phi).$$

Since this is true for all $\phi \in \mathcal{D}(\mathbb{R})$, we get that $\chi_a \rightarrow \delta$ in $\mathcal{D}'(\mathbb{R})$. ♣

6.4 Derivatives of Distributions

Proposition 6.4.1

If $u \in \mathcal{D}'(\Omega)$ and α is a multi-index, then

$$v(\phi) = (-1)^{|\alpha|} u(D^\alpha \phi) \quad , \quad \phi \in \mathcal{D}(\Omega) \quad ,$$

defines a distribution on Ω . We denote this distribution by $D^\alpha u$, the **partial derivative** of $u \in \mathcal{D}'(\Omega)$.

Proof.

- i) $(D^\alpha u)(\phi)$ is well defined for all $\phi \in \mathcal{D}(\Omega)$ because $D^\alpha \phi \in \mathcal{D}(\Omega)$ for all $\phi \in \mathcal{D}(\Omega)$.
- ii) It is easy to show that $D^\alpha u$ is a linear functional.

iii) To prove that $D^\alpha u$ is continuous, we show that $D^\alpha u$ satisfies Proposition 6.3.1.

Let K be a compact subset of Ω . Since $u \in \mathcal{D}'(\Omega)$, it follows from Proposition 6.3.1 that there exist a constant C_K and an integer N_K such that $|u(\phi)| \leq C_K \max_{|\beta| \leq N_K} \|D^\beta \phi\|_{\infty, K}$ for all $\phi \in \mathcal{D}_K$. Hence

$$|(D^\alpha u)(\phi)| = |(-1)^{|\alpha|} u(D^\alpha \phi)| \leq C_K \max_{|\beta| \leq N_K} \|D^{\beta+\alpha} \phi\|_{\infty, K} \leq C_K \max_{|\beta| \leq N_K + |\alpha|} \|D^\beta \phi\|_{\infty, K}$$

for all $\phi \in \mathcal{D}_K$. ■

The distributions are infinitely differentiable. The following example justifies this definition of derivative.

Example 6.4.2

Suppose that $u \in C^k(\mathbb{R}^n)$. Show that the partial derivatives $D^\alpha u$ of the distribution u on $\mathbb{R}^n$ with $|\alpha| \leq k$ corresponds to the standard partial derivatives of the function u . More precisely, show that the derivative of the distribution u on $\mathbb{R}^n$ is the distribution given by the standard derivative of the function u .

Note that $u \in C^k(\mathbb{R}^n)$ implies that standard derivative $D^\alpha u$ of u is in $C(\mathbb{R}^n)$ for all $|\alpha| \leq k$ and so it is in $L^1_{loc}(\mathbb{R}^n)$ for all $|\alpha| \leq k$. Thus, the standard derivative $D^\alpha u$ can define a distribution for all $|\alpha| \leq k$.

Let $\alpha = (1, 0, \dots, 0)$, the reasoning for other first order partial derivatives is similar. There are two ways to look at $D^\alpha u$.

i) The function u can be used to define a distribution on $\mathbb{R}^n$ because $u \in L^1_{loc}(\mathbb{R}^n)$. Hence, the derivative of u in the sense of distributions is

$$\langle D^\alpha u, \phi \rangle = -\langle u, D^\alpha \phi \rangle = - \int_{\mathbb{R}^n} u(\mathbf{x}) \frac{\partial \phi}{\partial x_1}(\mathbf{x}) d\mathbf{x} \quad , \quad \phi \in \mathcal{D}(\Omega) \quad .$$

ii) The function $D^\alpha u$ can be used to define a distribution on $\mathbb{R}^n$ because $D^\alpha u \in L^1_{loc}(\mathbb{R}^n)$. Hence, by definition,

$$\langle D^\alpha u, \phi \rangle = \int_{\mathbb{R}^n} (D_{x_1} u)(\mathbf{x}) \phi(\mathbf{x}) d\mathbf{x} \quad , \quad \phi \in \mathcal{D}(\Omega) \quad .$$

However, for all $\phi \in \mathcal{D}(\mathbb{R}^n)$, we have

$$\begin{aligned} - \int_{\mathbb{R}^n} u(\mathbf{x}) \frac{\partial \phi}{\partial x_1}(\mathbf{x}) d\mathbf{x} &= - \int_{\mathbb{R}^{n-1}} \left(\int_{-\infty}^{\infty} u(x_1, \mathbf{y}) \frac{\partial \phi}{\partial x_1}(x_1, \mathbf{y}) dx_1 \right) d\mathbf{y} \\ &= - \int_{\mathbb{R}^{n-1}} \left(\lim_{x_1 \rightarrow +\infty} u(x_1, \mathbf{y}) \phi(x_1, \mathbf{y}) - \lim_{x_1 \rightarrow -\infty} u(x_1, \mathbf{y}) \phi(x_1, \mathbf{y}) \right. \\ &\quad \left. - \int_{-\infty}^{\infty} D^\alpha u(x_1, \mathbf{y}) \phi(x_1, \mathbf{y}) dx_1 \right) d\mathbf{y} \\ &= \int_{\mathbb{R}^{n-1}} \left(\int_{-\infty}^{\infty} D^\alpha u(x_1, \mathbf{y}) \phi(x_1, \mathbf{y}) dx_1 \right) d\mathbf{y} = \int_{\mathbb{R}^n} D^\alpha u(\mathbf{x}) \phi(\mathbf{x}) d\mathbf{x} \quad , \end{aligned}$$

where we have used integration by parts for the differentiable functions u and ϕ , and

$$\lim_{x_1 \rightarrow +\infty} u(x_1, \mathbf{y}) \phi(x_1, \mathbf{y}) = \lim_{x_1 \rightarrow -\infty} u(x_1, \mathbf{y}) \phi(x_1, \mathbf{y}) = 0$$

because ϕ has a compact support, so the support of ϕ is bounded.

Thus, the action of the derivative of the distribution u given in (i) is equal to the action of the distribution $D^\alpha u$ given in (ii) on $\mathcal{D}(\mathbb{R}^n)$.

By Induction, one can show that this is also true for higher order partial derivatives. ♣

Example 6.4.3

Let $k(x_1, x_2) = \frac{1}{2\pi(x_1 + ix_2)}$ and $L = D_{x_1} + i D_{x_2}$.

Given $\phi \in \mathcal{D}(\mathbb{R}^2)$, we have that

$$\begin{aligned} \langle Lk, \phi \rangle &= -\langle k, L\phi \rangle = \frac{-1}{2\pi} \int_{\mathbb{R}^2} \left(\frac{\partial \phi}{\partial x_1} + i \frac{\partial \phi}{\partial x_2} \right) \frac{1}{x_1 + ix_2} d\mathbf{x} \\ &= \lim_{r \rightarrow 0} \frac{-1}{2\pi} \int_{\|\mathbf{x}\|_2 \geq r} \left(\frac{\partial \phi}{\partial x_1} + i \frac{\partial \phi}{\partial x_2} \right) \frac{1}{x_1 + ix_2} d\mathbf{x}, \end{aligned}$$

where the limit is justified by the Lebesgue Dominate Convergence Theorem. Using the Green's Theorem on the disk defined by $\|\mathbf{x}\|_2 \leq r$, we get

$$\begin{aligned} \int_{\|\mathbf{x}\|_2 \geq r} \left(\frac{\partial \phi}{\partial x_1} + i \frac{\partial \phi}{\partial x_2} \right) \frac{1}{x_1 + ix_2} d\mathbf{x} &= -i \int_{\|\mathbf{x}\|_2 \geq r} \left(\frac{\partial}{\partial x_1} \left(\frac{i\phi}{x_1 + ix_2} \right) - \frac{\partial}{\partial x_2} \left(\frac{\phi}{x_1 + ix_2} \right) \right) d\mathbf{x} \\ &= -i \int_{\|\mathbf{x}\|_2 = r} \frac{\phi}{x_1 + ix_2} dx_1 + \frac{i\phi}{x_1 + ix_2} dx_2 \end{aligned}$$

where the orientation of the circle $\|\mathbf{x}\|_2 = r$ is clockwise. If we use the parametric representation $(x_1, x_2) = (r \cos(\theta), r \sin(\theta))$ for $0 \leq \theta \leq 2\pi$, a parametric representation associated to the counterclockwise orientation of the circle, we get

$$\begin{aligned} \int_{\|\mathbf{x}\|_2 \geq r} \left(\frac{\partial \phi}{\partial x_1} + i \frac{\partial \phi}{\partial x_2} \right) \frac{1}{x_1 + ix_2} d\mathbf{x} &= i \int_0^{2\pi} \frac{\phi(r \cos(\theta), r \sin(\theta))}{re^{i\theta}} (-r \sin(\theta) + ir \cos(\theta)) d\theta \\ &= - \int_0^{2\pi} \phi(r \cos(\theta), r \sin(\theta)) d\theta. \end{aligned}$$

Thus,

$$\langle Lk, \phi \rangle = \lim_{r \rightarrow 0} \frac{1}{2\pi} \int_0^{2\pi} \phi(r \cos(\theta), r \sin(\theta)) d\theta = \frac{1}{2\pi} \int_0^{2\pi} \phi(\mathbf{0}) d\theta = \phi(\mathbf{0}) = \delta(\phi)$$

for all $\phi \in \mathcal{D}(\mathbb{R}^2)$. This means that $Lk = \delta$. ♣

Lemma 6.4.4

Let $]a, b[$ be an open interval of $\mathbb{R}$. If $f \in L^1_{loc}(]a, b[)$ satisfies

$$\langle f, \phi' \rangle = \int_a^b f(x) \phi'(x) dx = 0 \quad (6.4.1)$$

for all $\phi \in \mathcal{D}(]a, b[)$, then there exists a constant C such that $f = C$ almost everywhere on $]a, b[$.

Proof.

Choose $\psi \in \mathcal{D}(]a, b[)$ such that $\int_a^b \psi(x) dx = 1$. For each $\phi \in \mathcal{D}(]a, b[)$, let

$$\theta(x) = \int_a^x \left(\phi(y) - \psi(y) \int_a^b \phi(z) dz \right) dy.$$

We prove that $\theta \in \mathcal{D}(]a, b[)$. Let $K = \text{supp } \psi \cup \text{supp } \phi$ and choose $m, M \in]a, b[$ such $m < x < M$ for all $x \in K$. This is possible because ϕ and ψ have compact supports in the open set $]a, b[$. We then have that $\theta(x) = 0$ for $x < m$ because $\phi(x) = \psi(x) = 0$ for $x < m$, and $\theta(x) = 0$ for $x > M$ because

$$\begin{aligned} \int_a^x \left(\phi(y) - \psi(y) \int_a^b \phi(z) dz \right) dy &= \int_a^b \left(\phi(y) - \psi(y) \int_a^b \phi(z) dz \right) dy \\ &= \int_a^b \phi(y) dy - \left(\int_a^b \psi(y) dy \right) \left(\int_a^b \phi(z) dz \right) = 0 \end{aligned}$$

for all $x > M$, where we have used $\phi(x) = \psi(x) = 0$ for $x > M$.

It follows from (6.4.1) that

$$\begin{aligned} 0 &= \langle f, \theta' \rangle = \int_a^b f(x) \left(\phi(x) - \psi(x) \int_a^b \phi(z) dz \right) dx \\ &= \int_a^b f(x) \phi(x) dx - \left(\int_a^b f(x) \psi(x) dx \right) \left(\int_a^b \phi(z) dz \right) \\ &= \int_a^b f(x) \phi(x) dx - \left(\int_a^b f(z) \psi(z) dz \right) \left(\int_a^b \phi(x) dx \right) \\ &= \int_a^b \left(f(x) - \left(\int_a^b f(z) \psi(z) dz \right) \right) \phi(x) dx \end{aligned}$$

for all $\phi \in \mathcal{D}(]a, b[)$. Thus

$$f(x) - \left(\int_a^b f(z) \psi(z) dz \right) = 0$$

almost everywhere in $]a, b[$. The constant C in the conclusion of the lemma is given by $C = \int_a^b f(z) \psi(z) dz$. The readers should convince themselves that this is independent of ψ . ■

The previous proposition can be restated as follows. If $u \in \mathcal{D}'(]a, b[)$ is defined by $u(\phi) = \int_a^b f(x)\phi(x) dx$ for $\phi \in \mathcal{D}(]a, b[)$, where $f \in L^1_{loc}(]a, b[)$, and $Du = 0$ in the sense of distributions, then there exists a constant C such that $f(x) = C$ almost everywhere on $]a, b[$. Thus, we may say that u is a distribution given by the constant function $f(x) = C$ for all $x \in]a, b[$.

Proposition 6.4.5

Suppose that $u \in C(\Omega)$. Moreover, suppose that the derivative in the sense of distributions $D^\alpha u$ can be represented by a continuous function $g_\alpha : \Omega \rightarrow \mathbb{C}$ for $|\alpha| \leq k$; namely,

$$\langle D^\alpha u, \phi \rangle = (-1)^{|\alpha|} \langle u, D^\alpha \phi \rangle = \langle g_\alpha, \phi \rangle = \int_\Omega g_\alpha(\mathbf{x}) \phi(\mathbf{x}) d\mathbf{x}$$

for all $\phi \in \mathcal{D}(\Omega)$. Then $u \in C^k(\Omega)$.

Proof.

k = 1) Let $g_\alpha \in C(\Omega)$ be the function representing $D^\alpha u$ for $\alpha = (1, 0, \dots, 0)$, the partial derivative of the distribution u on Ω with respect to x_1 . The reasoning is similar for the other multi-indices α such that $|\alpha| = 1$.

Since the standard derivative is a local concept, we may assume that $\Omega = \prod_{j=1}^n]a_j, b_j[$.

Choose $\mathbf{y} \in \Omega$. Let

$$v(\mathbf{x}) = \int_{y_1}^{x_1} g_\alpha(s, \tilde{\mathbf{x}}) ds + u(y_1, \tilde{\mathbf{x}})$$

for $\mathbf{x} = (x_1, \tilde{\mathbf{x}}) \in \Omega$. Since $g_\alpha \in C(\Omega)$, we have that

$$\begin{aligned}]a_1, b_1[&\rightarrow \mathbb{C} \\ x &\mapsto v(x, \tilde{\mathbf{x}}) \end{aligned}$$

is differentiable in the classical sense for all $\tilde{\mathbf{x}} \in \prod_{j=2}^n]a_j, b_j[$. More precisely, $\frac{\partial v}{\partial x_1}(x, \tilde{\mathbf{x}}) = g_\alpha(x, \tilde{\mathbf{x}})$ for all $(x, \tilde{\mathbf{x}}) \in \Omega$. The goal is to show that $u = v$ in Ω .

Consider

$$\int_\Omega (u(\mathbf{x}) - v(\mathbf{x})) D^\alpha \phi(\mathbf{x}) d\mathbf{x} = \int_\Omega u(\mathbf{x}) D^\alpha \phi(\mathbf{x}) d\mathbf{x} - \int_\Omega v(\mathbf{x}) D^\alpha \phi(\mathbf{x}) d\mathbf{x} \quad (6.4.2)$$

for $\phi \in \mathcal{D}(\Omega)$. By definition of the derivative in the sense of distributions, we have

$$- \int_\Omega u(\mathbf{x}) D^\alpha \phi(\mathbf{x}) d\mathbf{x} = \int_\Omega g_\alpha(\mathbf{x}) \phi(\mathbf{x}) d\mathbf{x} \quad (6.4.3)$$

for all $\phi \in \mathcal{D}(\Omega)$ by hypothesis. Moreover, using integration by parts, we have

$$\int_\Omega v(\mathbf{x}) D^\alpha \phi(\mathbf{x}) d\mathbf{x} = \int_{\prod_{j=2}^n]a_j, b_j[} \int_{a_1}^{b_1} v(x_1, \tilde{\mathbf{x}}) \frac{\partial \phi}{\partial x_1}(x_1, \tilde{\mathbf{x}}) dx_1 d\tilde{\mathbf{x}}$$

$$\begin{aligned}
&= \int_{\prod_{j=2}^n]a_j, b_j[} \left(v(x_1, \tilde{\mathbf{x}}) \phi(x_1, \tilde{\mathbf{x}}) \Big|_{x_1=a_1}^{x_1=b_1} - \int_{a_1}^{b_1} \frac{\partial v}{\partial x_1}(x_1, \tilde{\mathbf{x}}) \phi(x_1, \tilde{\mathbf{x}}) dx_1 \right) d\tilde{\mathbf{x}} \\
&= - \int_{\Omega} \frac{\partial v}{\partial x_1}(\mathbf{x}) \phi(\mathbf{x}) d\mathbf{x} = - \int_{\Omega} g_{\alpha}(\mathbf{x}) \phi(\mathbf{x}) d\mathbf{x}
\end{aligned} \tag{6.4.4}$$

for all $\phi \in \mathcal{D}(\Omega)$, where we have use the fact that ϕ has a compact support in Ω to obtain

$$v(x_1, \tilde{\mathbf{x}}) \phi(x_1, \tilde{\mathbf{x}}) \Big|_{x_1=a_1}^{x_1=b_1} = \lim_{x \rightarrow b_1} v(x_1, \tilde{\mathbf{x}}) \phi(x_1, \tilde{\mathbf{x}}) - \lim_{x \rightarrow a_1} v(x_1, \tilde{\mathbf{x}}) \phi(x_1, \tilde{\mathbf{x}}) = 0 .$$

It follows from (6.4.2), (6.4.3) and (6.4.4) that

$$\int_{\Omega} (u(\mathbf{x}) - v(\mathbf{x})) D^{\alpha} \phi(\mathbf{x}) d\mathbf{x} = 0 \tag{6.4.5}$$

for all $\phi \in \mathcal{D}(\Omega)$.

If we consider $\phi \in \mathcal{D}(\Omega)$ of the form $\phi(\mathbf{x}) = \phi_1(x_1) \phi_2(\tilde{\mathbf{x}})$ for all $\mathbf{x} = (x_1, \tilde{\mathbf{x}}) \in \Omega$, where $\phi_1 \in \mathcal{D}(]a_1, b_1[)$ and $\phi_2 \in \mathcal{D}\left(\prod_{j=2}^n]a_j, b_j[\right)$, we can rewrite (6.4.5) as

$$\int_{\prod_{j=2}^n]a_j, b_j[} \left(\int_{a_1}^{b_1} (u(x_1, \tilde{\mathbf{x}}) - v(x_1, \tilde{\mathbf{x}})) D^{\alpha} \phi_1(x_1) dx_1 \right) \phi_2(\tilde{\mathbf{x}}) d\tilde{\mathbf{x}} = 0$$

for all $\phi_1 \in \mathcal{D}(]a_1, b_1[)$ and $\phi_2 \in \mathcal{D}\left(\prod_{j=2}^n]a_j, b_j[\right)$. It follows that, for almost all $\tilde{\mathbf{x}} \in \prod_{j=2}^n]a_j, b_j[$, we have

$$\int_{a_1}^{b_1} (u(x_1, \tilde{\mathbf{x}}) - v(x_1, \tilde{\mathbf{x}})) D^{\alpha} \phi_1(x_1) dx_1 = 0$$

for all $\phi_1 \in \mathcal{D}(]a_1, b_1[)$. Hence, from Lemma 6.4.4, there exists a constant C such that $u-v = C$ almost everywhere in Ω . Since v and u are continuous on Ω , it follows that $u(\mathbf{x}) - v(\mathbf{x}) = C$ for all $\mathbf{x} \in \Omega$. However, $u(y_1, x_2, \dots, x_n) = v(y_1, x_2, \dots, x_n)$. Thus $C = 0$.

$\mathbf{k} > 1$) A proof by induction on k takes care of higher order partial derivatives. ■

6.5 Convolution

Theorem 6.5.1 (Generalized Young's Inequality)

Let (X, μ) be a measure space, $1 \leq p \leq \infty$, and $K : X \times X \rightarrow \mathbb{C}$ be a measurable function. Suppose that there exists $C > 0$ such that

$$\sup_{x \in X} \int_X |K(x, y)| d\mu(y) \leq C \quad \text{and} \quad \sup_{y \in X} \int_X |K(x, y)| d\mu(x) \leq C .$$

If $f \in L^p(X)$, then $Tf : X \rightarrow \mathbb{C}$ defined by

$$(Tf)(x) = \int_X K(x, y) f(y) \, d\mu(y)$$

is defined almost everywhere, $Tf \in L^p(X)$ and $\|Tf\|_p \leq C\|f\|_p$.

Proof.

$0 < p < \infty$) Let q be the **conjugate exponent** of p ; namely. $\frac{1}{p} + \frac{1}{q} = 1$. Since

$$K(x, y)f(y) = K(x, y)^{1/q} K(x, y)^{1/p} f(y)$$

for all $x, y \in X$, we get from Hölder's inequality that

$$\begin{aligned} \int_X |K(x, y)f(y)| \, d\mu(y) &\leq \left(\int_X |K(x, y)| \, d\mu(y) \right)^{1/q} \left(\int_X |K(x, y)| |f(y)|^p \, d\mu(y) \right)^{1/p} \\ &\leq C^{1/q} \left(\int_X |K(x, y)| |f(y)|^p \, d\mu(y) \right)^{1/p} \end{aligned}$$

for all $x \in X$. Hence, from Fubini's theorem, we have

$$\begin{aligned} \int_X \left(\int_X |K(x, y)f(y)| \, d\mu(y) \right)^p \, d\mu(x) &\leq C^{p/q} \int_X \left(\int_X |K(x, y)| |f(y)|^p \, d\mu(y) \right) \, d\mu(x) \\ &= C^{p/q} \int_X \left(\int_X |K(x, y)| \, d\mu(x) \right) |f(y)|^p \, d\mu(y) \leq C^{1+p/q} \int_X |f(y)|^p \, d\mu(y) < \infty . \end{aligned} \quad (6.5.1)$$

Thus

$$\int_X |K(x, y)f(y)| \, d\mu(y) < \infty$$

for almost all $x \in X$. It follows that $y \mapsto K(x, y)f(y) \in L^1(\mu)$ for almost all $x \in X$. So Tf is defined almost everywhere. Moreover, since

$$|(Tf)(x)| \leq \int_X |K(x, y)f(y)| \, d\mu(y) ,$$

it follows from (6.5.1) that

$$\int_X |(Tf)(x)|^p \, d\mu(x) \leq C^{1+p/q} \int_X |f(y)|^p \, d\mu(y) .$$

After taking the p^{th} root on both sides of the previous inequality, we get

$$\|Tf\|_p \leq C^{1/p+1/q} \|f\|_p = C \|f\|_p .$$

$p = 1$) We have from Fubini's theorem that

$$\int_X \int_X |K(x, y)f(y)| \, d\mu(y) \, d\mu(x) = \int_X \left(\int_X |K(x, y)| \, d\mu(x) \right) |f(y)| \, d\mu(y)$$

$$\leq C \int_X |f(y)| \, d\mu(y) < \infty . \quad (6.5.2)$$

Thus

$$\int_X |K(x, y)f(y)| \, d\mu(y) < \infty$$

for almost all $x \in X$. It follows that $y \mapsto K(x, y)f(y) \in L^1(\mu)$ for almost all $x \in X$. So Tf is defined almost everywhere. Moreover, since

$$|(Tf)(x)| \leq \int_X |K(x, y)f(y)| \, d\mu(y) ,$$

we have from (6.5.2) that $\|Tf\| \leq C\|f\|$.

$p = \infty$) We have

$$\int_X |K(x, y)f(y)| \, d\mu(y) \leq \|f\|_\infty \int_X |K(x, y)| \, d\mu(y) \leq C\|f\|_\infty$$

for all $x \in X$. It follows that $y \mapsto K(x, y)f(y) \in L^1(\mu)$ for all $x \in X$. So Tf is defined (almost) everywhere. Moreover, $\|(Tf)(x)\|_\infty \leq C\|f\|_\infty$.

The norm $\|\cdot\|_\infty$ above is the essential supremum of the measurable function. ■

The **convolution** of f and g in $L^1(\mathbb{R}^n)$ is defined by

$$(f * g)(\mathbf{x}) = \int_{\mathbb{R}^n} f(\mathbf{x} - \mathbf{y}) g(\mathbf{y}) \, d\mathbf{y}$$

for all $\mathbf{x} \in \mathbb{R}^n$ where the integral converges. It can be proved that the integral is defined almost everywhere.

Theorem 6.5.2 (Young's Inequality)

if $f \in L^1(\mathbb{R}^n)$ and $g \in L^p(\mathbb{R}^n)$ for $1 \leq p \leq \infty$, then $f * g \in L^p(\mathbb{R}^n)$ and

$$\|f * g\|_p \leq \|f\|_1 \|g\|_p .$$

Proof.

It is a special case of the Generalized Young's Inequality where $X = \mathbb{R}^n$ and $K(x, y) = f(\mathbf{x} - \mathbf{y})$. ■

The following proposition is often used to expand to a larger class of functions results proved for a special class of functions.

Proposition 6.5.3

Suppose that $\phi \in L^1(\mathbb{R}^n)$ and let $a = \int_{\mathbb{R}^n} \phi(\mathbf{x}) \, d\mathbf{x}$. For $\epsilon > 0$, let

$$\phi_\epsilon(\mathbf{x}) = \epsilon^{-n} \phi(\epsilon^{-1} \mathbf{x}) .$$

If $f \in L^p(\mathbb{R}^n)$ with $1 \leq p < \infty$, then $f * \phi_\epsilon \rightarrow af$ in $L^p(\mathbb{R}^n)$ as $\epsilon \rightarrow 0$.
 If $f \in L^\infty(\mathbb{R}^n)$ and f is uniformly continuous on an open set $V \subset \mathbb{R}^n$, then $f * \phi_\epsilon \rightarrow af$ uniformly on K as $\epsilon \rightarrow 0$ for all compact sets $K \subset V$.

Proof.

The substitution $\mathbf{y} = \epsilon^{-1}\mathbf{x}$ shows that

$$\int_{\mathbb{R}^n} \phi_\epsilon(\mathbf{x}) d\mathbf{x} = \int_{\mathbb{R}^n} \epsilon^{-n} \phi(\epsilon^{-1}\mathbf{x}) d\mathbf{x} = \int_{\mathbb{R}^n} \phi(\mathbf{y}) d\mathbf{y} = a$$

for every $\epsilon > 0$.

i) If $f \in L^p(\mathbb{R}^n)$ with $1 \leq p < \infty$, the substitution $\mathbf{y} = \epsilon \mathbf{z}$ yields

$$(f * \phi_\epsilon)(\mathbf{x}) - af(\mathbf{x}) = \int_{\mathbb{R}^n} (f(\mathbf{x} - \mathbf{y}) - f(\mathbf{x})) \phi_\epsilon(\mathbf{y}) d\mathbf{y} = \int_{\mathbb{R}^n} (f(\mathbf{x} - \epsilon \mathbf{z}) - f(\mathbf{x})) \phi(\mathbf{z}) d\mathbf{z} .$$

Using Jensen's Inequality (Theorem 3.3 in [30]), we get

$$\begin{aligned} |(f * \phi_\epsilon)(\mathbf{x}) - af(\mathbf{x})|^p &\leq \left(\int_{\mathbb{R}^n} |f(\mathbf{x} - \epsilon \mathbf{z}) - f(\mathbf{x})| |\phi(\mathbf{z})| d\mathbf{z} \right)^p \\ &\leq \int_{\mathbb{R}^n} |f(\mathbf{x} - \epsilon \mathbf{z}) - f(\mathbf{x})|^p |\phi(\mathbf{z})| d\mathbf{z} . \end{aligned}$$

If we define $f_z \in L^p(\mathbb{R}^n)$ by $f_z(\mathbf{x}) = f(\mathbf{x} - \mathbf{z})$ for $\mathbf{x}, \mathbf{z} \in \mathbb{R}^n$, we then have that

$$\|f * \phi_\epsilon - af\|_p^p \leq \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |f(\mathbf{x} - \epsilon \mathbf{z}) - f(\mathbf{x})|^p |\phi(\mathbf{z})| d\mathbf{z} d\mathbf{x} \leq \int_{\mathbb{R}^n} \|f_{\epsilon \mathbf{z}} - f\|_p^p |\phi(\mathbf{z})| d\mathbf{z} .$$

However,

$$\|f_{\epsilon \mathbf{z}} - f\|_p^p |\phi(\mathbf{z})| \leq (2\|f\|_p)^p |\phi(\mathbf{z})| \in L^1(\mathbb{R}^n)$$

and

$$\|f_{\epsilon \mathbf{z}} - f\|_p^p |\phi(\mathbf{z})| \rightarrow 0 \quad \text{as } \epsilon \rightarrow 0$$

for all $\mathbf{z}$ because $\|f_{\epsilon \mathbf{z}} - f\|_p \rightarrow 0$ as $\epsilon \rightarrow 0$ (Theorem 9.5 in [30]). It follows from the Lebesgue Dominate Convergence Theorem that

$$\|f * \phi_\epsilon - af\|_p \rightarrow 0 \quad \text{as } \epsilon \rightarrow 0 .$$

ii) Suppose that $f \in L^\infty(\mathbb{R}^n)$ and f is uniformly continuous on $V \subset \mathbb{R}^n$. Let $K \subset V$ be a compact set.

Given $\delta > 0$, choose a compact set $W \subset \mathbb{R}^n$ such that

$$\int_{\mathbb{R}^n \setminus W} |\phi(\mathbf{x})| d\mathbf{x} < \frac{\delta}{4\|f\|_\infty} .$$

This is possible because $\phi \in L^1(\mathbb{R}^n)$.

Since f is uniformly continuous on the open set V , we may choose ϵ small enough such that $K + \epsilon W \subset V$ and

$$\sup_{\mathbf{x} \in K, \mathbf{y} \in W} |f(\mathbf{x} + \epsilon \mathbf{y}) - f(\mathbf{x})| < \frac{\delta}{2\|\phi\|_1}.$$

Then,

$$\begin{aligned} \sup_{\mathbf{x} \in K} |(f * \phi_\epsilon)(\mathbf{x}) - af(\mathbf{x})| &\leq \sup_{\mathbf{x} \in K} \int_{\mathbb{R}^n} |f(\mathbf{x} - \epsilon \mathbf{y}) - f(\mathbf{x})| |\phi(\mathbf{y})| d\mathbf{y} \\ &\leq \sup_{\mathbf{x} \in K, \mathbf{y} \in W} |f(\mathbf{x} - \epsilon \mathbf{y}) - f(\mathbf{x})| \int_W |\phi(\mathbf{y})| d\mathbf{y} + 2\|f\|_\infty \int_{\mathbb{R}^n \setminus W} |\phi(\mathbf{y})| d\mathbf{y} \\ &\leq \frac{\delta}{2\|\phi\|_1} \int_W |\phi(\mathbf{y})| d\mathbf{y} + 2\|f\|_\infty \frac{\delta}{4\|f\|_\infty} < \delta. \end{aligned}$$

The role of ϵ and δ are interchanged from the usual role that they play in a proof of convergence. Instead of “for all ϵ there exists $\delta \dots$ ”, we have “for all δ there exists $\epsilon \dots$ ” ■

Definition 6.5.4

The functions ϕ_ϵ for $\epsilon > 0$ defined in Proposition 6.5.3 are called **approximate identity** when $a = 1$.

Given $\phi : \mathbb{R}^n \rightarrow \mathbb{C}$, we define $\phi^\vee$ and ϕ_z in $\mathcal{D}(\mathbb{R}^n)$ by $\phi^\vee(\mathbf{x}) = \phi(-\mathbf{x})$ for $\mathbf{x} \in \mathbb{R}^n$ and $\phi_z(\mathbf{x}) = \phi(\mathbf{x} - \mathbf{z})$ for $\mathbf{x}, \mathbf{z} \in \mathbb{R}^n$.

This motivates the following definition. Given $u \in \mathcal{D}'(\mathbb{R}^n)$, we define $u^\vee$ and u_z in $\mathcal{D}'(\mathbb{R}^n)$ by $u^\vee(\phi) = \langle u^\vee, \phi \rangle = \langle u, \phi^\vee \rangle$ and $u_z(\phi) = \langle u_z, \phi \rangle = \langle u, \phi_z \rangle$ for $\phi \in \mathcal{D}(\mathbb{R}^n)$.

For two functions f and g in $L^1(\mathbb{R}^n)$, we have

$$(f * g)(\mathbf{x}) = \int_{\mathbb{R}^n} f(\mathbf{y}) (g^\vee)_x(\mathbf{y}) d\mathbf{y}.$$

It therefore makes sense to define the **convolution** of $u \in \mathcal{D}'(\mathbb{R}^n)$ with $\phi \in \mathcal{D}(\mathbb{R}^n)$ by

$$(u * \phi)(\mathbf{x}) = \langle u, (\phi^\vee)_x \rangle$$

for all $\mathbf{x} \in \mathbb{R}^n$. We will see later that this definition of convolution is the usual definition of convolution when $u \in L^1(\mathbb{R}^n)$.

Suppose that $u \in \mathcal{E}'(\mathbb{R}^n)$ and that K is the compact support of u . Choose $\psi \in \mathcal{D}(\mathbb{R}^n)$ such that $\psi(z) = 1$ for all $z \in K$. We have that $\psi u = u$ because $\langle \psi u, \phi \rangle = \langle u, \psi \phi \rangle = \langle u, \phi \rangle$ for all $\phi \in \mathcal{D}(\mathbb{R}^n)$. This is obviously independent of the function ψ chosen. We can use this property to extend u in a unique way to a continuous linear functional on $C^\infty(\mathbb{R}^n)$, where the convergence in $C^\infty(\mathbb{R}^n)$ is defined by uniform convergence on compact sets. We define $u : C^\infty(\mathbb{R}^n) \rightarrow \mathbb{C}$ by $u(\phi) = u(\psi \phi)$ for $\phi \in C^\infty(\mathbb{R}^n)$.

Proposition 6.5.5

1. If $u \in \mathcal{D}'(\mathbb{R}^n)$ and $\psi \in \mathcal{D}(\mathbb{R}^n)$, then $u * \psi$ is in $C^\infty(\mathbb{R}^n)$.
2. If $u \in \mathcal{E}'(\mathbb{R}^n)$ and $\psi \in \mathcal{D}(\mathbb{R}^n)$, then $u * \psi \in \mathcal{D}(\mathbb{R}^n)$.
3. If $u \in \mathcal{E}'(\mathbb{R}^n)$ and $\psi \in C^\infty(\mathbb{R}^n)$, then $u * \psi \in C^\infty(\mathbb{R}^n)$.

Proposition 6.5.6

If $u \in \mathcal{D}'(\mathbb{R}^n)$ and $\phi \in \mathcal{D}(\mathbb{R}^n)$, then $D^\alpha(u * \phi) = u * D^\alpha \phi = D^\alpha u * \phi$ for all multi-indices α .

We may also define the convolution of two distributions if one of them has compact support. Consider $u, v \in \mathcal{D}'(\mathbb{R}^n)$. If one of them has compact support, then the function $L : \mathcal{D}(\mathbb{R}^n) \rightarrow C^\infty(\mathbb{R}^n)$ defined by $L(\phi) = u * (v * \phi)$ for $\phi \in \mathcal{D}(\mathbb{R}^n)$ is well defined. To justify this statement, we note that if v has compact support, then $v * \phi \in \mathcal{D}(\mathbb{R}^n)$ by (2) of Proposition 6.5.5 and $u * (v * \phi) \in C^\infty(\mathbb{R}^n)$ by (1) of Proposition 6.5.5. Moreover, if u has compact support and v does not have compact support, then $v * \phi \in C^\infty(\mathbb{R}^n)$ by (1) of Proposition 6.5.5 and again $u * (v * \phi) \in C^\infty(\mathbb{R}^n)$ by (3) of Proposition 6.5.5.

It follows from Theorem 6.33 in [31] that there exists a unique distribution on $\mathbb{R}^n$, denoted $u * v$, such that

$$L(\phi) = u * (v * \phi) = (u * v) * \phi$$

for all $\phi \in \mathcal{D}(\mathbb{R}^n)$. In particular, we have

$$\langle (u * v), \phi \rangle = ((u * v) * \phi^\vee)(\mathbf{0}) = (u * (v * \phi^\vee))(\mathbf{0}) = \langle u, (v * \phi^\vee)^\vee \rangle$$

for all $\phi \in \mathcal{D}(\mathbb{R}^n)$.

Proposition 6.5.7

If $u \in \mathcal{E}'(\mathbb{R}^n)$ and $\psi \in C^\infty(\mathbb{R}^n)$, then $u * \psi = \psi * u$.

Proof.

First, we have to expand our definition of integration. Suppose that Ω is a measurable space with measure μ , that E is a **topological vector space**¹ and that $f : \Omega \rightarrow E$ is a measurable function. The integral of f over Ω is the element $y \in E$, if such a y exists, such that

$$u(y) = \int_{\Omega} u \circ f \, d\mu \tag{6.5.3}$$

for all $u \in E^*$. We assume that E^* separates points in E ² to ensure the uniqueness of y if it exists. There is a more traditional approach to define the integral in (6.5.3) using the

¹We will not give a definition of topological vector spaces. This will lead us too deep into analysis. The reader can find a definition of topological vector space in [27, 31, 39] and many other graduate textbooks in Analysis. A Banach space is a topological vector space.

²Namely, for any y_1 and y_2 in E with $y_1 \neq y_2$, there exists $u \in E^*$ such that $u(y_1) \neq u(y_2)$.

equivalent of “Riemann sums.” This approach can be found in more traditional textbooks in functional analysis.

Let

$$g(x) = \int_{\mathbb{R}^n} \psi(\mathbf{w}) \phi(\mathbf{x} - \mathbf{w}) d\mathbf{w}$$

for $\mathbf{x} \in \mathbb{R}^n$. According to the definition of $u * \psi$, we have

$$((u * \psi) * \phi)(\mathbf{z}) = (u * (\psi * \phi))(\mathbf{z}) = (u * g)(\mathbf{z})$$

for $\mathbf{z} \in \mathbb{R}^n$ because ψ and ϕ are of class $C^\infty(\mathbb{R}^n)$ with one of them having a compact support. So $\psi * \phi \in \mathcal{D}(\mathbb{R}^n)$. We then have that

$$((u * \psi) * \phi)(\mathbf{z}) = (u * g)(\mathbf{z}) = \langle u, (g^\vee)_\mathbf{z} \rangle = \langle u, h \rangle ,$$

where

$$h(\mathbf{x}) = g(\mathbf{z} - \mathbf{x}) = \int_{\mathbb{R}^n} \psi(\mathbf{w}) \phi(\mathbf{z} - \mathbf{x} - \mathbf{w}) d\mathbf{w}$$

for all $\mathbf{x} \in \mathbb{R}^n$. If $f(\mathbf{w})$ for $\mathbf{w} \in \mathbb{R}^n$ is the function in $\mathcal{D}(\mathbb{R}^n)$ defined by

$$f(\mathbf{w})(\mathbf{x}) = \psi(\mathbf{w}) \phi(\mathbf{z} - \mathbf{x} - \mathbf{w})$$

for all $\mathbf{x} \in \mathbb{R}^n$, then we get

$$((u * \psi) * \phi)(\mathbf{z}) = \langle u, h \rangle = \int_{\mathbb{R}^n} u \circ f d\mathbf{w} ,$$

where we have used (6.5.3) with $\Omega = \mathbb{R}^n$ and $E = \mathcal{D}(\mathbb{R}^n)$. If $q(\mathbf{w})$ for $\mathbf{w} \in \mathbb{R}^n$ is the function in $\mathcal{D}(\mathbb{R}^n)$ defined by

$$q(\mathbf{w})(\mathbf{x}) = \phi(\mathbf{z} - \mathbf{x} - \mathbf{w})$$

for all $\mathbf{x} \in \mathbb{R}^n$, then

$$\begin{aligned} ((u * \psi) * \phi)(\mathbf{z}) &= \int_{\mathbb{R}^n} \psi(\mathbf{w}) \langle u, q(\mathbf{w}) \rangle d\mathbf{w} = \int_{\mathbb{R}^n} \psi(\mathbf{w}) \langle u, (\phi^\vee)_{\mathbf{z}-\mathbf{w}} \rangle d\mathbf{w} \\ &= \int_{\mathbb{R}^n} \psi(\mathbf{w}) (u * \phi)(\mathbf{z} - \mathbf{w}) d\mathbf{w} = (\psi * (u * \phi))(\mathbf{z}) \end{aligned}$$

for $\mathbf{z} \in \mathbb{R}^n$ and all $\phi \in \mathcal{D}(\mathbb{R}^n)$. Thus, $\psi * (u * \phi) = (u * \psi) * \phi$ for all $\phi \in \mathcal{D}(\mathbb{R}^n)$. But, by definition of the convolution of two distributions with one having a compact support, this states that $\psi * u = u * \psi$. ■

6.6 Fourier Transforms

If $f \in L^1(\mathbb{R}^n)$, the **Fourier transform** of f is the function defined by

$$\mathcal{F}(f)(\mathbf{y}) \equiv \hat{f}(\mathbf{y}) = \int_{\mathbb{R}^n} e^{-i\mathbf{x}\cdot\mathbf{y}} f(\mathbf{x}) d\mu(\mathbf{x}) \quad , \quad \mathbf{y} \in \mathbb{R}^n , \quad (6.6.1)$$

where μ is the Lebesgue measure m divided by $(2\pi)^{n/2}$. To be able to state aesthetically pleasing relations between Fourier transform and convolution, we redefined the **convolution** of two functions $f, g \in L^1(\mathbb{R}^n)$ as

$$(f * g)(\mathbf{x}) = \int_{\mathbb{R}^n} f(\mathbf{x} - \mathbf{y}) g(\mathbf{y}) d\mu(\mathbf{y})$$

for all $\mathbf{x} \in \mathbb{R}^n$ where the integral converges. We have added the factor $(2\pi)^{-n/2}$ to the previous definition of convolution. We will use this definition of convolution only inside this section and revert to the previous definition in the rest of the text. The reader shall therefore not be surprised by the appearance of the factor $(2\pi)^{n/2}$ in the formulae involving convolution and Fourier transform later in the text.

Example 6.6.1

Suppose that $f \in C(\mathbb{R})$ has a compact support.

We first show that the function $g : \mathbb{R} \rightarrow \mathbb{R}$ defined by $g(x) = \sum_{k \in \mathbb{Z}} f(x + 2\pi k)$ is a continuous periodic function of period 2π . Since f has a compact support, the summation is finite on every bounded open set. It follows that $g \in C(\mathbb{R})$ since continuity is a local property. Since

$$g(x + 2\pi) = \sum_{k \in \mathbb{Z}} f(x + 2\pi(k + 1)) = \sum_{k \in \mathbb{Z}} f(x + 2\pi k) = g(x)$$

for all x , we have that g is periodic of period 2π .

We now show that the coefficient a_n of the Fourier series $\sum_{n \in \mathbb{Z}} a_n e^{nxi}$ of g is given by $a_n = \frac{1}{2\pi} \hat{f}(n)$ for all n . It is a simple computation from the definition of the coefficient of the series.

$$\begin{aligned} a_n &= \frac{1}{2\pi} \int_{-\pi}^{\pi} g(x) e^{-nxi} dx = \frac{1}{2\pi} \sum_{k \in \mathbb{Z}} \int_{-\pi}^{\pi} f(x + 2\pi k) e^{-nxi} dx \\ &= \frac{1}{2\pi} \sum_{k \in \mathbb{Z}} \int_{(2k-1)\pi}^{(2k+1)\pi} f(y) e^{-n(y-2\pi k)i} dy = \frac{1}{2\pi} \sum_{k \in \mathbb{Z}} \int_{(2k-1)\pi}^{(2k+1)\pi} f(y) e^{-ny i} dy \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} f(y) e^{-ny i} dy = \frac{1}{2\pi} \hat{f}(n) \end{aligned}$$

for all $n \in \mathbb{Z}$. Note that the previous summations are finite.

It follows that

$$\frac{1}{2\pi} \sum_{n \in \mathbb{Z}} \hat{f}(n) = \sum_{n \in \mathbb{Z}} f(n) . \quad (6.6.2)$$

Since g is continuous we have that

$$\frac{1}{2\pi} \sum_{n \in \mathbb{Z}} \hat{f}(n) e^{nxi} = \sum_{n \in \mathbb{Z}} a_n e^{nxi} = g(x) = \sum_{n \in \mathbb{Z}} f(x + n)$$

for all x . If $x = 0$, we get (6.6.2). The formula (6.6.2) is known as the **Poisson summation formula**. ♣

Example 6.6.2

As an example, we compute the Fourier transform of $f(x) = e^{-\|x\|^2}$ for $x \in \mathbb{R}^n$. This Fourier transform will play an important role in the study of the heat equation. Since

$$\hat{f}(y) = \int_{\mathbb{R}^n} e^{-i x \cdot y} e^{-\|x\|^2} d\mu(x) = \prod_{j=1}^n \left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-i x_j y_j - x_j^2} dx_j \right),$$

it suffices to compute

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-i x y - x^2} dx.$$

If we use the substitution $x = r/2$, we get

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-i x y - x^2} dx = \frac{1}{2\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(2iry + r^2)/4} dr = \frac{1}{2\sqrt{2\pi}} e^{-y^2/4} \int_{-\infty}^{\infty} e^{-(r+yi)^2/4} dr.$$

To compute the integral, we note that $e^{-z^2/4}$ with $z \in \mathbb{C}$ is an entire function and use Cauchy's theorem for integrals along a close path to conclude that $\int_{\Gamma} e^{-z^2/4} dz = 0$ where Γ is the close path defined by $\Gamma = \Gamma_1 + \Gamma_2 - \Gamma_3 - \Gamma_4$ with $\Gamma_1 = \{r + 0i : -R \leq r \leq R\}$, $\Gamma_2 = \{R + si : 0 \leq s \leq y\}$, $\Gamma_3 = \{r + yi : -R \leq r \leq R\}$ and $\Gamma_4 = \{-R + si : 0 \leq s \leq y\}$ (Figure 6.1). We have that

$$\int_{\Gamma_1} e^{-z^2/4} dz = \int_{-R}^R e^{-r^2/4} dr \rightarrow \int_{-\infty}^{\infty} e^{-r^2/4} dr = 2 \int_{-\infty}^{\infty} e^{-t^2} dt = 2\sqrt{\pi} \quad \text{as } R \rightarrow \infty,$$

$$\int_{\Gamma_3} e^{-z^2/4} dz = \int_{-R}^R e^{-(r+yi)^2/4} dr \rightarrow \int_{-\infty}^{\infty} e^{-(r+yi)^2/4} dr \quad \text{as } R \rightarrow \infty,$$

$$\left| \int_{\Gamma_2} e^{-z^2/4} dz \right| \leq \int_0^y |e^{-(R+si)^2/4}| ds = e^{-R^2/4} \int_0^y e^{s^2/4} ds \rightarrow 0 \quad \text{as } R \rightarrow \infty,$$

and similarly

$$\left| \int_{\Gamma_4} e^{-z^2/4} dz \right| \leq \int_0^y |e^{-(-R+si)^2/4}| ds = e^{-R^2/4} \int_0^y e^{s^2/4} ds \rightarrow 0 \quad \text{as } R \rightarrow \infty.$$

Thus,

$$\int_{-\infty}^{\infty} e^{-(r+yi)^2/4} dr = 2\sqrt{\pi}.$$

We finally get that

$$\hat{f}(y) = \prod_{j=1}^n \left(\frac{1}{\sqrt{2}} e^{-y_j^2/4} \right) = \frac{1}{2^{n/2}} e^{-\|y\|^2/4}.$$

More generally, we can compute the Fourier transform of $f(x) = e^{-a\|x\|^2}$ for $x \in \mathbb{R}^n$ and a a positive constant. We have that

$$\hat{g}(y) = \int_{\mathbb{R}^n} e^{-i x \cdot y} e^{-a\|x\|^2} d\mu(x).$$

If we use the change of variables $w = \sqrt{a} x$, we get

$$\hat{g}(y) = \frac{1}{a^{n/2}} \int_{\mathbb{R}^n} e^{-i w \cdot (a^{-1/2} y)} e^{-\|w\|^2} d\mu(w) = \frac{1}{a^{n/2}} \hat{f}(a^{-1/2} y) = \frac{1}{(2a)^{n/2}} e^{-\|y\|^2/(4a)}. \quad (6.6.3)$$

♣

The Fourier transform has the following properties.

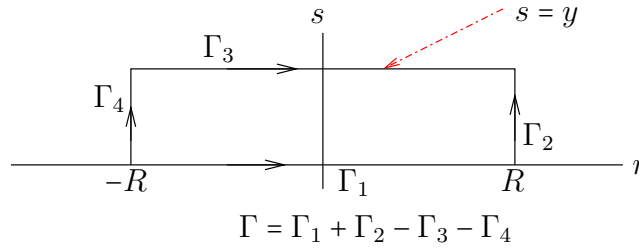


Figure 6.1: According to Cauchy's theorem for integrals along a close path, we have that $\int_{\Gamma} = \int_{\Gamma_1} + \int_{\Gamma_2} - \int_{\Gamma_3} - \int_{\Gamma_4} = 0$.

Proposition 6.6.3

1. If f and g are in $L^1(\mathbb{R}^n)$, then $(f * g)^\wedge = \hat{f} \hat{g}$.
2. If f is in $L^1(\mathbb{R}^n)$, then $(f_z)^\wedge(\mathbf{y}) = e^{-i\mathbf{z} \cdot \mathbf{y}} \hat{f}(\mathbf{y})$
3. If $f \in C_c^m(\mathbb{R}^n)$, α is a multi-index such that $|\alpha| \leq m$, and p is the polynomial $p(\mathbf{x}) = (-i\mathbf{x})^\alpha$, then $(D^\alpha f)^\wedge = p \hat{f}$ and $(pf)^\wedge = D^\alpha \hat{f}$.

Proof.

Left to the reader. ■

In a course for engineer students, the Fourier transform and the “inverse Fourier transform” are introduced without considerations for the domain and range of these transformations. The following theorem states under which conditions the inversion formula is valid. Before stating this theorem, we need to define the space $C_0(\mathbb{R}^n)$ of continuous functions that **vanish at infinity**.

The space $C_0(\mathbb{R}^n)$ is the space of continuous real valued functions on $\mathbb{R}^n$ satisfying the following property. For every $\epsilon > 0$, there exists a compact set K such that $|f(\mathbf{x})| < \epsilon$ for $\mathbf{x} \in \mathbb{R}^n \setminus K$.

Proposition 6.6.4

If $f \in L^1(\mathbb{R}^n)$, then $\hat{f} \in C_0(\mathbb{R}^n)$ and $\|\hat{f}\|_\infty \leq \|f\|_1$.

Proposition 6.6.5 (Inverse Fourier Transform)

If $f \in L^1(\mathbb{R}^n)$, $\hat{f} \in L^1(\mathbb{R}^n)$ and

$$g(\mathbf{y}) = \int_{\mathbb{R}^n} e^{i\mathbf{x} \cdot \mathbf{y}} \hat{f}(\mathbf{x}) d\mu(\mathbf{x})$$

for all $\mathbf{y} \in \mathbb{R}^n$, then $g \in C_0(\mathbb{R}^n)$. Moreover, $f = g$ almost everywhere in $\mathbb{R}^n$ and

$$\|f\|_\infty \leq \|\hat{f}\|_1.$$

Given $f \in L^1(\mathbb{R}^n)$, let

$$\mathcal{F}^{-1}(f)(\mathbf{y}) \equiv \check{f}(\mathbf{y}) = \int_{\mathbb{R}^n} e^{i\mathbf{x} \cdot \mathbf{y}} f(\mathbf{x}) d\mu(\mathbf{x}) \quad , \quad \mathbf{y} \in \mathbb{R}^n .$$

The previous theorem states that $\mathcal{F}^{-1}(\mathcal{F}(f)) = f$ almost everywhere when $f \in L^1(\mathbb{R}^n)$ and $\hat{f} \in L^1(\mathbb{R}^n)$.

It is possible to extend the Fourier transform to $L^2(\mathbb{R}^n)$. This will be useful when talking about Sobolev spaces later.

Theorem 6.6.6 (Plancherel)

To each $f \in L^2(\mathbb{R}^n)$ can be associated a function $\hat{f} \in L^2(\mathbb{R}^n)$ such that

1. For $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$. $\hat{f}$ is the Fourier transform given in (6.6.1).
2. $\|f\|_2 = \|\hat{f}\|_2$ for all $f \in L^2(\mathbb{R}^n)$. In fact, $\int_{\mathbb{R}^n} fg d\mathbf{x} = \int_{\mathbb{R}^n} \hat{f}\hat{g} d\mathbf{x}$ for all $f, g \in L^2(\mathbb{R}^n)$.
3. The linear mapping $\mathcal{F} : L^2(\mathbb{R}^n) \rightarrow L^2(\mathbb{R}^n)$ defined by $\mathcal{F}(f) = \hat{f}$ for all $f \in L^2(\mathbb{R}^n)$ is an Hilbert space isomorphism.

In fact, (2) and (3) in Theorem 6.6.6 implies that $\mathcal{F} : L^2(\mathbb{R}^n) \rightarrow L^2(\mathbb{R}^n)$ is an isometry.

The space $\mathcal{S}(\mathbb{R}^n)$ of **rapidly decreasing functions** on $\mathbb{R}^n$ (also called **Schwartz class**) is the space of all functions $f \in C^\infty(\mathbb{R}^n)$ such that

$$\sup_{\mathbf{x} \in \mathbb{R}^n} |\mathbf{x}^\alpha D^\beta f(\mathbf{x})| < \infty$$

for all multi-indices α and β . A locally convex topology is defined on $\mathcal{S}(\mathbb{R}^n)$ with the help of the norms defined by

$$\sup_{|\alpha| \leq k} \sup_{\mathbf{x} \in \mathbb{R}^n} (1 + \|\mathbf{x}\|_2^2)^k |D^\alpha f(\mathbf{x})|$$

for $k \geq 0$. We will not elaborate on this subject but the interested reader may consult [31, 27] for more information.

If $f \in \mathcal{S}(\mathbb{R}^n)$, we have that $\mathbf{x}^\alpha D^\beta f$ itself vanishes at infinity for all multi-indices α and β . Moreover, $\mathbf{x}^\alpha D^\beta f \in L^1(\mathbb{R}^n)$ for all multi-indices α and β . To prove this, choose $k > n/2$ and $C > 0$. Since f vanishes at infinity, there exists $r > 0$ such that $\|\mathbf{x}\|^{2k} |f(\mathbf{x})| < C$ for all $\|\mathbf{x}\| > r$. Hence,

$$\begin{aligned} \int_{\mathbb{R}^n} |f(\mathbf{x})| d\mathbf{x} &= \int_{\|\mathbf{x}\| \leq r} |f(\mathbf{x})| d\mathbf{x} + \int_{\|\mathbf{x}\| > r} |f(\mathbf{x})| d\mathbf{x} \\ &= \int_{\|\mathbf{x}\| \leq r} |f(\mathbf{x})| d\mathbf{x} + \int_{\|\mathbf{x}\| > r} \|\mathbf{x}\|^{2k} |f(\mathbf{x})| \|\mathbf{x}\|^{-2k} d\mathbf{x} \end{aligned}$$

$$\leq \int_{\|\mathbf{x}\| \leq r} |f(\mathbf{x})| d\mathbf{x} + C \int_{\|\mathbf{x}\| > r} \|\mathbf{x}\|^{-2k} d\mathbf{x} < \infty$$

because f is locally integrable, the integral of $|f|$ on $\{\mathbf{x} : \|\mathbf{x}\| \leq r\}$ is finite, and $\int_{\|\mathbf{x}\| > r} \|\mathbf{x}\|^{-2k} d\mathbf{x} < \infty$ since $2k > n$. For all multi-indices α and β , this reasoning can be applied to $\mathbf{x}^\alpha D^\beta f$ instead of f to show that $\mathbf{x}^\alpha D^\beta f \in L^1(\mathbb{R}^n)$.

Hence, the Fourier transform of rapidly decreasing functions and their derivatives exists.

Proposition 6.6.7

If $f \in L^1(\mathbb{R}^n)$ has compact support, then $\hat{f}$ is the restriction to $\mathbb{R}^n$ of an entire function on $\mathbb{C}^n$. If $f \in \mathcal{D}(\mathbb{R}^n)$, then $\hat{f}$ is a rapidly decreasing function on any subset of the form $\{\mathbf{z} \in \mathbb{C}^n : \|\operatorname{Im} \mathbf{z}\|_2 \leq c\}$ for $c > 0$. Recall that $\operatorname{Im} \mathbf{z} = (\operatorname{Im} z_1 \ \operatorname{Im} z_2 \ \dots \ \operatorname{Im} z_n)^\top$ for $\mathbf{z} \in \mathbb{C}^n$.

Proof.

The integral

$$\int_{\mathbb{R}^n} e^{-i\mathbf{z} \cdot \mathbf{x}} f(\mathbf{x}) d\mu(\mathbf{x}) ,$$

converges for all $\mathbf{z} \in \mathbb{C}^n$ because f has a compact support so the integration is over a bounded set and the exponential factor is bounded on this set. Hence, we may define $\hat{f}(\mathbf{z})$ for $\mathbf{z} \in \mathbb{C}^n$ by

$$\hat{f}(\mathbf{z}) = \int_{\mathbb{R}^n} e^{-i\mathbf{z} \cdot \mathbf{x}} f(\mathbf{x}) d\mu(\mathbf{x}) .$$

$\hat{f}$ is an entire function because $\mathbf{z} \mapsto e^{i\mathbf{z} \cdot \mathbf{x}}$ is an entire function and, for any multi-index α , complex derivatives D_z^α and integration can be interchanged.

If $f \in \mathcal{D}(\mathbb{R}^n)$, choose r large enough such that $\operatorname{supp} f \subset \{\mathbf{x} : \|\mathbf{x}\| < r\}$. Then, for any multi-index α ,

$$\begin{aligned} |\mathbf{z}^\alpha \hat{f}(\mathbf{z})| &= |(i)^{|\alpha|} (D^\alpha f)^\wedge(\mathbf{z})| = \left| \int_{\mathbb{R}^n} e^{-i\mathbf{z} \cdot \mathbf{x}} D^\alpha f(\mathbf{x}) d\mu(\mathbf{x}) \right| \\ &\leq \int_{\|\mathbf{x}\| < r} |e^{-i\mathbf{z} \cdot \mathbf{x}}| |D^\alpha f(\mathbf{x})| d\mu(\mathbf{x}) \leq e^{r\|\operatorname{Im} \mathbf{z}\|_2} \int_{\|\mathbf{x}\| < r} |D^\alpha f(\mathbf{x})| d\mu(\mathbf{x}) \\ &= e^{r\|\operatorname{Im} \mathbf{z}\|_2} \int_{\mathbb{R}^n} |D^\alpha f(\mathbf{x})| d\mu(\mathbf{x}) \end{aligned} \tag{6.6.4}$$

for all $\mathbf{z} \in \mathbb{C}^n$. Hence, for $\|\operatorname{Im} \mathbf{z}\|_2 \leq c$, we have

$$|\mathbf{z}^\alpha \hat{f}(\mathbf{z})| \leq e^{rc} \int_{\mathbb{R}^n} |D^\alpha f(\mathbf{x})| d\mu(\mathbf{x}) < \infty .$$

Given any two multi-indices α and β , to prove that $\mathbf{z}^\alpha D^\beta \hat{f}(\mathbf{z})$ is uniformly bounded for $\|\operatorname{Im} \mathbf{z}\|_2 \leq c$, we use the relation

$$\mathbf{z}^\alpha D^\beta \hat{f}(\mathbf{z}) = \mathbf{z}^\alpha ((-i\mathbf{x})^\beta f)^\wedge(\mathbf{z}) = (-i)^{|\beta|} \mathbf{z}^\alpha (\mathbf{x}^{VEC\beta} f)^\wedge(\mathbf{z})$$

to reduce the problem to the previous case with $\mathbf{x}^\beta f$ instead of f where $\mathbf{x}^\beta f \in \mathcal{D}(\mathbb{R}^n)$. ■

The Fourier transform can only be defined on a subset of the space $\mathcal{D}'(\mathbb{R}^n)$ of distributions on $\mathbb{R}^n$.

Proposition 6.6.8

1. $\mathcal{D}(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n)$ and the inclusion $i : \mathcal{D}(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n)$ is continuous.
2. $\mathcal{D}(\mathbb{R}^n)$ is dense in $\mathcal{S}(\mathbb{R}^n)$.

In the statement of the previous proposition, the topology on $\mathcal{D}(\mathbb{R}^n)$ is the topology defined in Section 6.2 and the topology on $\mathcal{S}(\mathbb{R}^n)$ is the locally convex topology that we have mentioned above.

If u is a continuous linear functional on $\mathcal{S}(\mathbb{R}^n)$, then $u \circ i$ is a continuous linear functional on $\mathcal{D}(\mathbb{R}^n)$; namely, $u \circ i \in \mathcal{D}'(\mathbb{R}^n)$. Since $\mathcal{D}(\mathbb{R}^n)$ is dense in $\mathcal{S}(\mathbb{R}^n)$, each continuous linear functional on $\mathcal{S}(\mathbb{R}^n)$ is associated in this manner to a unique element of $\mathcal{D}'(\mathbb{R}^n)$. The **tempered distributions** on $\mathbb{R}^n$ are the distributions on $\mathbb{R}^n$ associated to continuous linear functionals on $\mathcal{S}(\mathbb{R}^n)$. For this reason, the space of all the tempered distributions is denoted $\mathcal{S}'(\mathbb{R}^n)$. Another way to look at the tempered distributions is to say that $u \in \mathcal{D}'(\mathbb{R}^n)$ is a tempered distribution if it can be extended to a continuous linear functional on $\mathcal{S}(\mathbb{R}^n)$.

Example 6.6.9

Here are a couple of examples of tempered distributions.

If ν is a positive measure on $\mathbb{R}^n$ such that $\int_{\mathbb{R}^n} (1 + \|\mathbf{x}\|^2)^{-k} d\nu < \infty$ for some positive integer k , then ν defines a tempered distribution on $\mathbb{R}^n$. More precisely, $u : \mathcal{S}(\mathbb{R}^n) \rightarrow \mathbb{C}$ defined by $u(\phi) = \int_{\mathbb{R}^n} \phi(\mathbf{x}) d\nu(\mathbf{x})$ for $\phi \in \mathcal{S}(\mathbb{R}^n)$ is a temperate distribution on $\mathbb{R}^n$.

If $g : \mathbb{R}^n \rightarrow \mathbb{R}$ is a measurable functions such that $\int_{\mathbb{R}^n} |(1 + \|\mathbf{x}\|^2)^{-k} g(\mathbf{x})|^p d\mathbf{x} < \infty$ for a positive integer k and a real number $p \geq 1$, then g defines a tempered distribution on $\mathbb{R}^n$. More precisely, $u : \mathcal{S}(\mathbb{R}^n) \rightarrow \mathbb{C}$ defined by $u(\phi) = \int_{\mathbb{R}^n} \phi(\mathbf{x}) g(\mathbf{x}) d\mathbf{x}$ for $\phi \in \mathcal{S}(\mathbb{R}^n)$ is a temperate distribution on $\mathbb{R}^n$.

Every $g \in L^p(\mathbb{R}^n)$ for $1 \leq p \leq \infty$ defines a tempered distribution on $\mathbb{R}^n$. ♣

The following proposition states that all distributions with compact support are also tempered distribution but there are more tempered distributions.

Proposition 6.6.10

$$\mathcal{E}'(\mathbb{R}^n) \subset \mathcal{S}'(\mathbb{R}^n) \subset \mathcal{D}'(\mathbb{R}^n)$$

The Fourier transform of a rapidly decreasing function has the following properties.

Proposition 6.6.11

1. The Fourier transform (6.6.1) is a continuous linear bijection from $\mathcal{S}(\mathbb{R}^n)$ to itself and its inverse is continuous.

2. If ϕ and ψ are in $\mathcal{S}(\mathbb{R}^n)$, then $\phi * \psi \in \mathcal{S}(\mathbb{R}^n)$ and $(\phi * \psi)^\wedge = \hat{\phi} \hat{\psi}$.

The **convolution** of a temperate distribution $u \in \mathcal{S}'(\mathbb{R}^n)$ with a function $\phi \in \mathcal{S}(\mathbb{R}^n)$ is defined by

$$(u * \phi)(\mathbf{x}) = \langle u, (\phi^\vee)_\mathbf{x} \rangle$$

for all $\mathbf{x} \in \mathbb{R}^n$. The **Fourier transform** of a tempered distribution $u \in \mathcal{S}'(\mathbb{R}^n)$ is the distribution $\hat{u}$ defined by

$$\langle \hat{u}, \phi \rangle = \langle u, \hat{\phi} \rangle \quad , \quad \phi \in \mathcal{S}(\mathbb{R}^n) . \quad (6.6.5)$$

As for the standard distributions, we use the notation $\langle u, \phi \rangle = u(\phi)$ for $u \in \mathcal{S}'(\mathbb{R}^n)$ and $\phi \in \mathcal{S}(\mathbb{R}^n)$.

Proposition 6.6.12

1. The Fourier transform (6.6.5) is a continuous linear bijection from $\mathcal{S}'(\mathbb{R}^n)$ into itself and its inverse is continuous.
2. If $u \in \mathcal{S}'(\mathbb{R}^n)$, α is a multi-index, and p is the polynomial $p(\mathbf{x}) = (i\mathbf{x})^\alpha$, then $(D^\alpha u)^\wedge = p \hat{u}$ and $(p^\vee u)^\wedge = D^\alpha \hat{u}$.
3. If $u \in \mathcal{S}'(\mathbb{R}^n)$ and $\phi \in \mathcal{S}(\mathbb{R}^n)$, then $u * \phi \in C^\infty(\mathbb{R}^n)$ and $u * \phi \in \mathcal{S}'(\mathbb{R}^n)$.
4. If $u \in \mathcal{S}'(\mathbb{R}^n)$ and $\phi \in \mathcal{S}(\mathbb{R}^n)$, then $(u * \phi)^\wedge = \hat{\phi} \hat{u}$ and $\hat{u} * \hat{\phi} = (\phi u)^\wedge$.

Proof.

We will only prove the first part of (2). The reader should consult the references for the proofs of the other results.

By definition, we have that $(D^\alpha u)(\phi) = (-1)^\alpha u(D^\alpha \phi)$ for all $\phi \in \mathcal{S}(\mathbb{R}^n)$. Hence,

$$(D^\alpha u)^\wedge(\phi) = (D^\alpha u)(\hat{\phi}) = (-1)^{|\alpha|} u(D^\alpha \hat{\phi}) = (-1)^{|\alpha|} u((-i\mathbf{x})^\alpha \hat{\phi}) = (i\mathbf{x})^\alpha u(\hat{\phi}) = (i\mathbf{x})^\alpha \hat{u}(\phi)$$

for all $\phi \in \mathcal{S}(\mathbb{R}^n)$. Thus, $(D^\alpha u)^\wedge = p \hat{u}$ in the sense of distributions. ■

Example 6.6.13

The Dirac delta function δ is a distribution with compact support, so we can compute its Fourier transform. Since

$$\langle \hat{\delta}, \phi \rangle = \langle \delta, \hat{\phi} \rangle = \hat{\phi}(\mathbf{0}) = \int_{\mathbb{R}^n} \phi(\mathbf{x}) d\mu(\mathbf{x}) = \langle f, \phi \rangle$$

for all $\phi \in \mathcal{S}(\mathbb{R}^n)$, where $f \in \mathcal{S}'(\mathbb{R}^n)$ is the tempered distribution defined by the measure μ ; namely, $f(\phi) = \int_{\mathbb{R}^n} \phi(\mathbf{x}) d\mu(\mathbf{x}) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \phi(x) dx$ for all $\phi \in \mathcal{S}(\mathbb{R}^n)$, then $\hat{\delta} = f$.

Note that in the literature, it is often stated that $\hat{\delta} = 1$. The reason behind the difference with our statement is that they do not use the factor $(2\pi)^{-n/2}$ in their definition of the Fourier transform as we have done with the measure μ . ♣

Example 6.6.14

Using Example 6.3.8, it is also possible to prove that $\hat{\delta} = \mu$ in $\mathcal{D}'(\mathbb{R})$. This is often the proof given for this result.

Since the Fourier transform is continuous on $\mathcal{S}'(\mathbb{R})$ and since χ_a defined in Example 6.3.8 converges to δ in $\mathcal{D}(\mathbb{R})$ as $a \rightarrow 0$, and so in $\mathcal{S}(\mathbb{R})$ because of (1) in Proposition 6.6.8, we have that $(\chi_a)^\wedge$ converges to $\hat{\delta}$ in $\mathcal{S}'(\mathbb{R})$.

Since $\chi_a \in L^1(\mathbb{R})$, its Fourier transform is given by

$$(\chi_a)^\wedge(x) = \frac{1}{2a\sqrt{2\pi}} \int_{-a}^a e^{-ixy} dy = \frac{-1}{2a\sqrt{2\pi} ix} e^{-ixy} \Big|_{-a}^a = \frac{\sin(ax)}{ax\sqrt{2\pi}}$$

for $x \neq 0$ and

$$(\chi_a)^\wedge(x) = \frac{1}{2a\sqrt{2\pi}} \int_{-a}^a dy = \frac{1}{\sqrt{2\pi}}$$

for $x = 0$. Thus $\lim_{a \rightarrow 0} (\chi_a)^\wedge(x) = 1/\sqrt{2\pi}$ for all $x \in \mathbb{R}$.

Thus $\lim_{a \rightarrow 0} (\chi_a)^\wedge(x) = 1/\sqrt{2\pi}$ for all $x \in \mathbb{R}$. We have shown that $(\chi_a)^\wedge$ converges pointwise to the function $f : \mathbb{R} \rightarrow \mathbb{R}$ defined by $f(x) = 1/\sqrt{2\pi}$ for all x . We cannot immediately conclude that $\hat{\delta} = f$ because pointwise convergence is not the convergence in $\mathcal{S}'(\mathbb{R})$. However, if the pointwise limit exists, then the limit in $\mathcal{S}'(\mathbb{R})$ must be equal almost everywhere to the pointwise limit. Thus, $\hat{\delta} = f$ almost everywhere as expected. $\clubsuit$

6.7 Fundamental Solutions

Our presentation is based on [10] and make use of the Fourier transform. This approach is analytic. A more elementary and algebraic presentation is given in [35].

Let $\Omega \subset \mathbb{R}^n$ be an open set and consider the linear partial differential equation

$$L(\mathbf{x}, D)u = \sum_{|\alpha| \leq m} a_\alpha D^\alpha u = f \quad (6.7.1)$$

where $u : \Omega \rightarrow \mathbb{C}$, $f : \Omega \rightarrow \mathbb{C}$ is continuous and $a_\alpha : \Omega \rightarrow \mathbb{C}$ is of class $C^{|\alpha|}$ for all multi-indices involved in the sum.

As in the introduction of the present chapter, let $L^*(\mathbf{x}, D)$ be the adjoint of $L(\mathbf{x}, D)$ defined by

$$L^*(\mathbf{x}, D)\phi = \sum_{|\alpha| \leq m} (-1)^{|\alpha|} D^\alpha (a_\alpha \phi) \quad , \quad \phi \in \mathcal{D}(\Omega) .$$

A **strong solution** or a **solution in the sense of distributions** of (6.7.1) is a distribution u on Ω such that

$$\langle L(\mathbf{x}, D)u, \phi \rangle = \langle u, L^*(\mathbf{x}, D)\phi \rangle = \langle f, \phi \rangle$$

for all $\phi \in \mathcal{D}(\Omega)$.

Consider the linear partial differential equation with constant coefficients

$$L(\mathbf{x}, D)u = \sum_{|\alpha| \leq m} a_\alpha D^\alpha u = f, \quad (6.7.2)$$

where $a_\alpha \in \mathbb{R}$ for all multi-indices α involved in the sum, $u : \mathbb{R}^n \rightarrow \mathbb{C}$ and $f \in \mathcal{D}(\mathbb{R}^n)$.

We assume that the coordinates are such that $\{\mathbf{x} \in \mathbb{R}^n : x_n = 0\}$ is not a characteristic surface. This implies that $a_{(0,0,\dots,0,m)} \neq 0$. Without loss of generality, we may even assume that $a_{(0,0,\dots,0,m)} = 1$. We show that there exists a solution $u \in C^\infty(\mathbb{R}^n)$ for (6.7.2).

We have that $L(\mathbf{x}, D)\hat{u} = (pu)^\wedge$, where

$$p(\mathbf{y}) = \sum_{|\alpha| \leq m} a_\alpha (-i\mathbf{y})^\alpha. \quad (6.7.3)$$

Let $\lambda_1(\check{\mathbf{y}})$, $\lambda_2(\check{\mathbf{y}})$, $\dots$, $\lambda_m(\check{\mathbf{y}})$ be the roots of $g(z) = p(\check{\mathbf{y}}, z)$ for $\check{\mathbf{y}} \in \mathbb{R}^{n-1}$. By Rouché's theorem, the roots $\lambda_j : \mathbb{R}^{n-1} \rightarrow \mathbb{C}$ are continuous functions.

Lemma 6.7.1

There exists a measurable function $\kappa : \mathbb{R}^{n-1} \rightarrow [-m, m]$ such that

$$\min_{\substack{\check{\mathbf{y}} \in \mathbb{R}^{n-1} \\ 1 \leq j \leq m}} \{|\kappa(\check{\mathbf{y}}) - \operatorname{Im} \lambda_j(\check{\mathbf{y}})|\} \geq 1.$$

Proof.

Given $\check{\mathbf{y}} \in \mathbb{R}^{n-1}$, there exists a least one interval $I_k = [2k - m - 1, 2k - m + 1]$ with $0 \leq k \leq m$ such that $\operatorname{Im} \lambda_j(\check{\mathbf{y}}) \notin I_k$ for all j because there are $m + 1$ intervals and at most m distinct zeros. Let

$$V_k = \{\check{\mathbf{y}} \in \mathbb{R}^{n-1} : \operatorname{Im} \lambda_j(\check{\mathbf{y}}) \notin I_k \text{ for } 1 \leq j \leq m\}$$

for $0 \leq k \leq m$. We have that $\mathbb{R}^{n-1} = \bigcup_{1 \leq k \leq m} V_k$ and the V_k are measurable because

$$V_k = \bigcap_{1 \leq j \leq m} \lambda_j^{-1}(\mathbb{C} \setminus (\mathbb{R} + i[2k - m - 1, 2k - m + 1]))$$

and the λ_j are continuous.

The function κ is defined by $\kappa(\check{\mathbf{y}}) = 2k - m$ if $\check{\mathbf{y}} \in V_k \setminus \bigcup_{0 \leq j < k} V_j$. It is a measurable function because $\kappa^{-1}(\{2k - m\}) = V_k \setminus \bigcup_{0 \leq j < k} V_j$ is a measurable set. ■

Lemma 6.7.2

Let $q(z) = z^m + g(z)$, where g is a polynomial of degree less than m and $g(0) \neq 0$. Suppose that $\lambda_1, \lambda_2, \dots, \lambda_m$ are the roots of q . Then, $|q(0)| \geq \left(\frac{\lambda}{2}\right)^m$, where $\lambda =$

$$\min_{1 \leq j \leq m} |\lambda_j|.$$

Proof.

The factorization of q is $q(z) = (z - \lambda_1)(z - \lambda_2) \dots (z - \lambda_m)$. Hence,

$$|q(0)| = \left| \prod_{j=1}^m \lambda_j \right| = \prod_{j=1}^m |\lambda_j| \geq \lambda^m \geq \left(\frac{\lambda}{2} \right)^m. \quad \blacksquare$$

Remark 6.7.3

A more convoluted but cute proof of the previous lemma was given in [10]. The factorization of q is $q(z) = (z - \lambda_1)(z - \lambda_2) \dots (z - \lambda_m)$. Hence,

$$\left| \frac{q(z)}{q(0)} \right| = \prod_{j=1}^m \left| 1 - \frac{z}{\lambda_j} \right| \leq 2^m$$

for $|z| = \lambda$ since $|z/\lambda_j| = \lambda/|\lambda_j| \leq 1$, so $z/\lambda_j \in \overline{B_1(0)} \subset \mathbb{C}$ and $|z/\lambda_j - 1| \leq 2$. Since $m! = q^{(m)}(z)$, Cauchy Integral Formula gives

$$\begin{aligned} m! &= |q^{(m)}(0)| = \left| \frac{m!}{2\pi i} \int_{|z|=\lambda} \frac{q(z)}{z^{m+1}} dz \right| = \frac{m!}{2\pi} \left| \int_0^{2\pi} \frac{q(\lambda e^{i\theta})}{(\lambda e^{i\theta})^{m+1}} \lambda i e^{i\theta} d\theta \right| \\ &\leq \frac{m!}{2\pi \lambda^m} \int_0^{2\pi} |q(\lambda e^{i\theta})| d\theta \leq \frac{m! 2^m |q(0)|}{2\pi \lambda^m} \int_0^{2\pi} d\theta = \frac{m! 2^m |q(0)|}{\lambda^m}. \end{aligned}$$

Thus, $\left(\frac{\lambda}{2} \right)^m \leq |q(0)|.$ ♠

Theorem 6.7.4

There exists a solution of class C^∞ for (6.7.2).

Proof.

Let

$$u(\mathbf{x}) = (2\pi)^{-n/2} \int_{\mathbb{R}^{n-1}} \int_{\text{Im } y_n = \kappa(\check{\mathbf{y}})} e^{-i\mathbf{x} \cdot \mathbf{y}} \frac{\hat{f}(\mathbf{y})}{p(\mathbf{y})} dy_n d\check{\mathbf{y}}$$

for $\mathbf{x} \in \mathbb{R}^n$, where κ is defined in Lemma 6.7.1 and p is defined in (6.7.3). In the previous integral, we assume that $y_n \in \mathbb{C}$.

i) Given $\mathbf{y} \in \mathbb{R}^n$, we first prove that $|p(\mathbf{y})| \geq 2^{-m}$ for $\text{Im } y_n = \kappa(\check{\mathbf{y}})$. Let $q(z) = p(\check{\mathbf{y}}, y_n + z)$. From Lemma 6.7.2, we have that

$$|p(\mathbf{y})| = |q(0)| \geq \left(\frac{\lambda}{2} \right)^m,$$

where $\lambda = \min_{1 \leq j \leq m} \{|\lambda_j| : \lambda_j \text{ is a root of } q\}$. However, the roots of the present polynomial q are $\lambda_j(\check{\mathbf{y}}) - y_n$, where $\lambda_j(\check{\mathbf{y}})$ for $1 \leq j \leq m$ are the roots of the polynomial $g(z) = p(\check{\mathbf{y}}, z)$. Hence, for $\text{Im } y_n = \kappa(\check{\mathbf{y}})$, we get from Lemma 6.7.1 that

$$\lambda = \min_{1 \leq j \leq m} |\lambda_j(\check{\mathbf{y}}) - y_n| \geq \min_{1 \leq j \leq m} |\text{Im } \lambda_j(\check{\mathbf{y}}) - \kappa(\check{\mathbf{y}})| \geq 1 .$$

ii) From Proposition 6.6.7, $\hat{f}$ is rapidly decreasing on the set $\{\mathbf{y} \in \mathbb{C}^n : \|\text{Im } \mathbf{y}\|_2 \leq m\}$ that contains the domain of integration $\{\mathbf{y} = (\check{\mathbf{y}}, y_n) \in \mathbb{R}^{n-1} \times \mathbb{C} : \text{Im } y_n = \kappa(\check{\mathbf{y}})\}$ because $|\text{Im } y_n| \leq m$ since the range of κ is $[-m, m]$.

It follows from (i) and (ii) that $\mathbf{y} \mapsto e^{-i\mathbf{x}\cdot\mathbf{y}} \frac{\hat{f}(\mathbf{y})}{p(\mathbf{y})}$ is bounded and rapidly decreasing on the domain of integration

$$S = \{\mathbf{y} = (\check{\mathbf{y}}, y_n) \in \mathbb{R}^{n-1} \times \mathbb{C} : |y_n| \leq m\} .$$

Since $|\text{Im } y_n| \leq m$ in the domain of integration, the exponential in the definition of u is bounded by e^{mx_n} . It follows that $\mathbf{y} \mapsto D_{\mathbf{x}}^{\alpha} e^{-i\mathbf{x}\cdot\mathbf{y}} \frac{\hat{f}(\mathbf{y})}{p(\mathbf{y})}$ is rapidly decreasing on S for all multi-indices α . Hence, we may interchange derivative and integral in the definition of u . Thus u is of class C^{∞} .

We have

$$\begin{aligned} (L(\mathbf{x}, D)u)(\mathbf{x}) &= (2\pi)^{-n/2} \int_{\mathbb{R}^{n-1}} \int_{\text{Im } y_n = \kappa(\check{\mathbf{y}})} (L(\mathbf{x}, D)e^{-i\mathbf{x}\cdot\mathbf{y}}) \frac{\hat{f}(\mathbf{y})}{p(\mathbf{y})} dy_n d\check{\mathbf{y}} \\ &= (2\pi)^{-n/2} \int_{\mathbb{R}^{n-1}} \int_{\text{Im } y_n = \kappa(\check{\mathbf{y}})} e^{-i\mathbf{x}\cdot\mathbf{y}} \hat{f}(\mathbf{y}) dy_n d\check{\mathbf{y}} \end{aligned} \quad (6.7.4)$$

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

Since $y_n \mapsto e^{i\mathbf{x}\cdot\mathbf{y}} \hat{f}(\mathbf{y})$ is an entire function, we may use Cauchy's theorem for integrals along a close path to conclude that

$$\int_{y_n \in \Gamma} e^{-i\mathbf{x}\cdot\mathbf{y}} \hat{f}(\mathbf{y}) dy_n = 0 ,$$

where $\Gamma = \Gamma_1 + \Gamma_2 - \Gamma_3 - \Gamma_4$ with $\Gamma_1 = \{r + 0i : -R \leq r \leq R\}$, $\Gamma_2 = \{R + si : 0 \leq s \leq \kappa(\check{\mathbf{y}})\}$, $\Gamma_3 = \{r + \kappa(\check{\mathbf{y}})i : -R \leq r \leq R\}$ and $\Gamma_4 = \{-R + si : 0 \leq s \leq \kappa(\check{\mathbf{y}})\}$ (Figure 6.2). We have that

$$\int_{y_n \in \Gamma_1} e^{-i\mathbf{x}\cdot\mathbf{y}} \hat{f}(\mathbf{y}) dy_n \rightarrow \int_{-\infty}^{\infty} e^{-i\mathbf{x}\cdot\mathbf{y}} \hat{f}(\mathbf{y}) dy_n \quad \text{as } R \rightarrow \infty$$

and

$$\int_{y_n \in \Gamma_3} e^{-i\mathbf{x}\cdot\mathbf{y}} \hat{f}(\mathbf{y}) dy_n \rightarrow \int_{\text{Im } y_n = \kappa(\check{\mathbf{y}})} e^{-i\mathbf{x}\cdot\mathbf{y}} \hat{f}(\mathbf{y}) dy_n \quad \text{as } R \rightarrow \infty .$$

Moreover, since $\mathbf{y} \mapsto 2^{i\mathbf{x}\cdot\mathbf{y}} \hat{f}(\mathbf{y})$ is rapidly decreasing on the set S , we have that there exists a constant $C > 0$ such that $|2^{i\mathbf{x}\cdot\mathbf{y}} \hat{f}(\mathbf{y})| \leq C/\|\mathbf{y}\|_2^2$ for all $\mathbf{z} \in S$. Thus, for $R^2 > 2m^2$, we have that

$$\left| \int_{y_n \in \Gamma_2} e^{-i\mathbf{x}\cdot\mathbf{y}} \hat{f}(\mathbf{y}) dy_n \right| \leq C \int_0^{\kappa(\check{\mathbf{y}})} \frac{1}{\|(\check{\mathbf{y}}, R + is)\|_2^2} ds \leq C \int_0^m \frac{1}{R^2 - s^2} ds$$

$$\leq C \int_0^m \frac{1}{R^2 - m^2} ds = \frac{2Cm}{R^2} \rightarrow 0 \quad \text{as } R \rightarrow \infty$$

and similarly

$$\begin{aligned} \left| \int_{y_n \in \Gamma_4} e^{-i\mathbf{x} \cdot \mathbf{y}} \hat{f}(\mathbf{y}) dy_n \right| &\leq C \int_0^{\kappa(\check{\mathbf{y}})} \frac{1}{\|(\check{\mathbf{y}}, -R + is)\|_2^2} ds \leq C \int_0^m \frac{1}{R^2 - s^2} ds \\ &\leq C \int_0^m \frac{1}{R^2 - m^2} ds = \frac{2Cm}{R^2} \rightarrow 0 \quad \text{as } R \rightarrow \infty. \end{aligned}$$

Hence,

$$\int_{\text{Im } y_n = \kappa(\check{\mathbf{y}})} e^{-i\mathbf{x} \cdot \mathbf{y}} \hat{f}(\mathbf{y}) dy_n = \int_{-\infty}^{\infty} e^{-i\mathbf{x} \cdot \mathbf{y}} \hat{f}(\mathbf{y}) dy_n.$$

It follows from (6.7.4) that

$$(L(\mathbf{x}, D)u)(\mathbf{x}) = \int_{\mathbb{R}^n} e^{-i\mathbf{x} \cdot \mathbf{y}} \hat{f}(\mathbf{y}) d\mu(\mathbf{y}) = f(\mathbf{x})$$

for $\mathbf{x} \in \mathbb{R}^n$, where we have used the Formula for the Inverse Fourier transform. ■

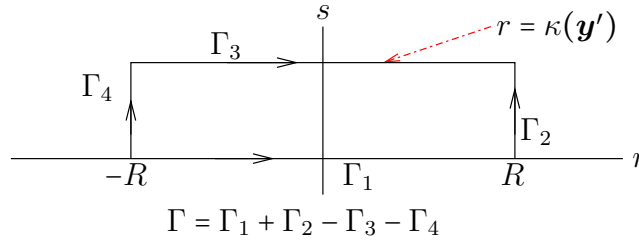


Figure 6.2: According to Cauchy's theorem for integrals along a close path, we have that $\int_{\Gamma} = \int_{\Gamma_1} + \int_{\Gamma_2} - \int_{\Gamma_3} - \int_{\Gamma_4} = 0$.

Definition 6.7.5

A **fundamental solution** for the differential operator $L(\mathbf{x}, D)$ defined in (6.7.2) is a distribution F on $\mathbb{R}^n$ such that $L(\mathbf{x}, D)F = \delta$.

Example 6.7.6

Let $R = \{\mathbf{x} \in \mathbb{R}^2 : x_i > 0 \text{ for } i = 1, 2\}$ and $L(\mathbf{x}, D) = D^{(1,1)} = \frac{\partial^2}{\partial x_1 \partial x_2}$. We have that $\mathcal{X}_R$ is a fundamental solution of the differential operator $L(\mathbf{x}, D)$ because

$$\begin{aligned} \langle L(\mathbf{x}, D)\mathcal{X}_R, \phi \rangle &= \langle \mathcal{X}_R, D^{(1,1)} \phi \rangle = \int_0^\infty \int_0^\infty \frac{\partial^2 \phi}{\partial x_1 \partial x_2}(x_1, x_2) dx_2 dx_1 \\ &= \int_0^\infty \left(\frac{\partial \phi}{\partial x_1}(x_1, x_2) \Big|_0^\infty \right) dx_1 = - \int_0^\infty \frac{\partial \phi}{\partial x_1}(x_1, 0) dx_1 = - \phi(x_1, 0) \Big|_0^\infty = \phi(0, 0) = \delta(\phi) \end{aligned}$$

for all $\phi \in \mathcal{D}(\mathbb{R}^2)$. ♣

Remark 6.7.7

If $F \in \mathcal{D}'(\mathbb{R}^n)$ is a fundamental solution for $L(\mathbf{x}, D)$, then a solution of $L(\mathbf{x}, D)u = f$, where $f \in \mathcal{D}(\mathbb{R}^n)$, is the distribution $u = F * f$ because

$$L(\mathbf{x}, D)u = L(\mathbf{x}, D)(F * f) = (L(\mathbf{x}, D)F) * f = \delta * f = f.$$

Theorem 6.7.8 (Malgrange-Ehrempreis)

Every differential linear operator with constant coefficients

$$L(\mathbf{x}, D) = \sum_{|\alpha| \leq m} a_\alpha D^\alpha \quad (6.7.5)$$

has a fundamental solution on $\mathbb{R}^n$. ♠

Proof.

We define a linear functional F on $\mathcal{D}(\mathbb{R}^n)$ by

$$F(\phi) = (2\pi)^{-n/2} \int_{\mathbb{R}^{n-1}} \int_{\text{Im } y_n = \kappa(\check{\mathbf{y}})} \frac{\hat{\phi}^\vee(\mathbf{y})}{p(\mathbf{y})} dy_n d\check{\mathbf{y}}, \quad \phi \in \mathcal{D}(\mathbb{R}^n),$$

where κ is defined in Lemma 6.7.1 and p is defined in (6.7.3). In the previous integral, we assume that $y_n \in \mathbb{C}$. To prove that $F \in \mathcal{D}'(\mathbb{R}^n)$, we have to show that F is continuous on $\mathcal{D}(\mathbb{R}^n)$. We will also show that $L(\mathbf{x}, D)F = \delta$.

i) Let $K \subset \mathbb{R}^n$ be a compact set. We will find a constant C_K and an integer N_K , such that

$$|F(\phi)| \leq C_K \sum_{|\alpha| \leq N_K} \|D^\alpha \phi\|_{\infty, K}$$

for all $\phi \in \mathcal{D}_K$. It will follow from Proposition 6.3.1 that F is continuous on $\mathcal{D}(\mathbb{R}^n)$ and so $F \in \mathcal{D}'(\mathbb{R}^n)$.

As in the proof of Theorem 6.7.4, we can use Proposition 6.6.7 to conclude that $\hat{\phi}$ is rapidly decreasing on the set $\{\mathbf{y} \in \mathbb{C}^n : \|\text{Im } \mathbf{y}\|_2 \leq m\}$. Moreover, we have shown in the proof of Theorem 6.7.4 that $|p(\mathbf{y})| \geq 2^{-m}$ for $\text{Im } y_n = \kappa(\check{\mathbf{y}})$. Hence,

$$\begin{aligned} |F(\phi)| &\leq \frac{1}{C_1(2\pi)^{n/2}} \int_{\mathbb{R}^{n-1}} \int_{-\infty}^{\infty} |\hat{\phi}^\vee(\check{\mathbf{y}}, r + \kappa(\check{\mathbf{y}})i)| dr d\check{\mathbf{y}} \\ &\leq \underbrace{\left(\frac{1}{C_1(2\pi)^{n/2}} \int_{\mathbb{R}^{n-1}} \int_{-\infty}^{\infty} (1 + \|\check{\mathbf{y}}\|^2 + r^2)^{-n-1} dr d\check{\mathbf{y}} \right)}_{=C_2} \sup_{\|\text{Im } \mathbf{y}\| \leq m} \left\{ (1 + \|\mathbf{y}\|^2)^{n+1} |\hat{\phi}^\vee(\mathbf{y})| \right\} \\ &= C_2 \sup_{\|\text{Im } \mathbf{y}\| \leq m} \left\{ (1 + \|\mathbf{y}\|^2)^{n+1} |\hat{\phi}^\vee(\mathbf{y})| \right\} \leq C_3 \sum_{|\alpha| \leq 2(n+1)} \|D^\alpha \phi\|_\infty \end{aligned}$$

for all $\phi \in \mathcal{D}_K$ and some constant C_3 . The last inequality above is proved using the relation

$$|\mathbf{y}^\alpha \hat{\phi}(\mathbf{y})| \leq e^{r\|\operatorname{Im} \mathbf{y}\|_2} \int_{\mathbb{R}^n} |D^\alpha \phi(\mathbf{x})| d\mu(\mathbf{x})$$

for $\mathbf{y} \in \mathbb{C}^n$ and $\alpha \in \mathbb{N}^n$ that we have obtained in (6.6.4) of the proof of Proposition 6.6.7, where r is such that $K \subset B_r(\mathbf{0})$. Thus, the constant C_3 depends on K .

ii) We have $\langle L(\mathbf{x}, D)F, \phi \rangle = \langle F, L^*(\mathbf{x}, D)\phi \rangle$ for $\phi \in \mathcal{D}(\mathbb{R}^n)$ where $L^*(\mathbf{x}, D) = \sum_{|\alpha| \leq m} (-1)^{|\alpha|} a_\alpha D^\alpha$. Let $p^*(\mathbf{y}) = \sum_{|\alpha| \leq m} a_\alpha (i\mathbf{y})^\alpha$. Then

$$((L^*(\mathbf{x}, D)\phi))^\wedge(-\mathbf{y}) = p^*(-\mathbf{y})\hat{\phi}(-\mathbf{y}) = p(\mathbf{y})\hat{\phi}(-\mathbf{y})$$

for $\mathbf{y} \in \mathbb{R}^n$. Hence

$$\begin{aligned} \langle L(\mathbf{x}, D)F, \phi \rangle &= \langle F, L^*(\mathbf{x}, D)\phi \rangle = (2\pi)^{-n/2} \int_{\mathbb{R}^{n-1}} \int_{\operatorname{Im} \mathbf{y}_n = \kappa(\check{\mathbf{y}})} \frac{(L^*(\mathbf{x}, D)\phi)^\wedge(-\mathbf{y})}{p(\mathbf{y})} dy_n d\check{\mathbf{y}} \\ &= (2\pi)^{-n/2} \int_{\mathbb{R}^{n-1}} \int_{\operatorname{Im} \mathbf{y}_n = \kappa(\check{\mathbf{y}})} \hat{\phi}(-\mathbf{y}) dy_n d\check{\mathbf{y}} = (2\pi)^{-n/2} \int_{\mathbb{R}^n} \hat{\phi}(-\mathbf{y}) d\mathbf{y} \\ &= \int_{\mathbb{R}^n} \hat{\phi}(\mathbf{y}) d\mu(\mathbf{y}) = \phi(\mathbf{0}) = \delta(\phi) \end{aligned}$$

for all $\phi \in \mathcal{D}(\mathbb{R}^n)$. The fourth equality comes from the Cauchy's theorem for integral along a close path as we have done in Theorem 6.7.4 for instance. Note that $\hat{\phi}$ is the restriction of an entire function which is rapidly decreasing on the set $\{\mathbf{y} \in \mathbb{C}^n : \|\operatorname{Im} \mathbf{y}\|_2 \leq m\}$ according to Proposition 6.6.7. The second to last equality is a consequence of the inverse Fourier transform. ■

Definition 6.7.9

Let Ω be an open subset of $\mathbb{R}^n$. The differential operator $L(\mathbf{x}, D)u = \sum_{|\alpha| \leq m} a_\alpha D^\alpha u$, where the $a_\alpha \in C^\infty(\Omega)$, is **hypoelliptic** if $L(\mathbf{x}, D)u \in C^\infty(\Omega)$ implies that $u \in C^\infty(\Omega)$.

Recall that we say that $u \in \mathcal{D}'(\Omega)$ is of class $C^\infty(\Omega)$ and write $u \in C^\infty(\Omega)$ if there exists a function $f \in C^\infty(\Omega)$ such that $u(\phi) = \int_\Omega f(x)\phi(x) dx$ for all $\phi \in \mathcal{D}(\Omega)$.

The next lemma will be used to give a sufficient condition to determine if a differential operator is hypoelliptic.

Lemma 6.7.10

If $f \in \mathcal{D}'(\mathbb{R}^n)$ is given by $f \in C^\infty(\mathbb{R}^n \setminus \{\mathbf{0}\})$ and $v \in \mathcal{E}'(\mathbb{R}^n)$, then $f * v \in C^\infty(\mathbb{R}^n \setminus \operatorname{supp} v)$.

Proof.

Consider $\mathbf{z} \notin \operatorname{supp} v$ and choose $\epsilon > 0$ such that $B_\epsilon(\mathbf{z}) \cap \operatorname{supp} v = \emptyset$. Choose $\psi \in \mathcal{D}(\mathbb{R}^n)$ such that $\psi(\mathbf{x}) = 1$ for $\mathbf{x} \in B_{\epsilon/4}(\mathbf{0})$ and $\operatorname{supp} \psi \subset B_{\epsilon/2}(\mathbf{0})$.

Then $f * v = (\psi f) * v + ((1 - \psi)f) * v$ where $((1 - \psi)f) * v \in C^\infty(\mathbb{R}^n)$ according to item (3) of Proposition 6.5.5 and Proposition 6.5.7 because $(1 - \psi)f \in C^\infty(\mathbb{R}^n)$. Moreover,

$$\text{supp}((\psi f) * v) \subset D = \{\mathbf{x} + \mathbf{y} : \mathbf{x} \in \text{supp } \psi, \mathbf{y} \in \text{supp } v\}$$

and $D \cap B_{\epsilon/2}(\mathbf{z}) = \emptyset$. Hence $((\phi f) * v)(\phi) = 0$ for all $\phi \in \mathcal{D}(B_{\epsilon/2}(\mathbf{z}))$. Thus $f * v = ((1 - \phi)f) * v$ on $\mathcal{D}(B_{\epsilon/2}(\mathbf{z}))$ and so $f * v \in C^\infty(B_{\epsilon/2}(\mathbf{z}))$.

Since $\mathbf{z} \notin \text{supp } v$ is arbitrary, we get $f * v \in C^\infty(\mathbb{R}^n \setminus \text{supp } v)$. ■

Theorem 6.7.11

Consider the differential operator

$$L(\mathbf{x}, D)u = \sum_{|\alpha| \leq m} a_\alpha D^\alpha u,$$

where $a_\alpha \in \mathbb{R}$ for all multi-indices α . The following sentences are equivalent.

1. There exists a fundamental solution of $L(\mathbf{x}, D)$ of class C^∞ on $\mathbb{R}^n \setminus \{\mathbf{0}\}$.
2. Every fundamental solution of $L(\mathbf{x}, D)$ is of class C^∞ on $\mathbb{R}^n \setminus \{\mathbf{0}\}$.
3. $L(\mathbf{x}, D)$ is hypoelliptic.

Proof.

(3) $\Rightarrow$ (2)) If F is a fundamental solution, then $L(\mathbf{x}, D)F = \delta \in C^\infty(\mathbb{R}^n \setminus \{\mathbf{0}\})$ and (3) implies that $F \in C^\infty(\mathbb{R}^n \setminus \{\mathbf{0}\})$.

(2) $\Rightarrow$ (1)) This is trivial because of Theorem 6.7.8.

(1) $\Rightarrow$ (3)) Let $F \in C^\infty(\mathbb{R}^n \setminus \{\mathbf{0}\})$ be a fundamental solution of $L(\mathbf{x}, D)$. Suppose that $u \in \mathcal{D}'(\Omega)$ is such that $L(\mathbf{x}, D)u \in C^\infty(\Omega)$.

Given $\mathbf{z} \in \Omega$ and $\epsilon > 0$ such that $B_\epsilon(\mathbf{z}) \subset \Omega$, we prove that $u \in C^\infty(B_{\epsilon/2}(\mathbf{z}))$. Choose $\phi \in \mathcal{D}(\mathbb{R}^n)$ such that $\text{supp } \phi \subset B_\epsilon(\mathbf{z})$ and $\phi(\mathbf{x}) = 1$ for all $\mathbf{x} \in B_{\epsilon/2}(\mathbf{z})$.

We have

$$L(\mathbf{x}, D)(\phi u) = \phi L(\mathbf{x}, D)u + v,$$

where v is an expression involving the partial derivatives of ϕ (but not ϕ itself) and the partial derivatives of u of order less than m .

It follows from item (1) of Proposition 6.5.5 that $F * (\phi L(\mathbf{x}, D)u) \in C^\infty(\mathbb{R}^n)$ because $\phi L(\mathbf{x}, D)u \in \mathcal{D}(\mathbb{R}^n)$.

We have that $v(\phi) = 0$ for all $\phi \in \mathcal{D}(\mathbb{R}^n)$ with $\text{supp } \phi \subset \overline{B_{\epsilon/2}(\mathbf{z})} \cup (\mathbb{R}^n \setminus B_\epsilon(\mathbf{z}))$ because $\phi(\mathbf{x}) = 0$ for $\mathbf{x} \notin B_\epsilon(\mathbf{z})$ and $\phi(\mathbf{x}) = 1$ for $\mathbf{x} \in B_{\epsilon/2}(\mathbf{z})$. Hence, from Lemma 6.7.10, we have that $F * v \in C^\infty(B_{\epsilon/2}(\mathbf{z}))$ because $\text{supp } v \subset \overline{B_\epsilon(\mathbf{z})} \setminus B_{\epsilon/2}(\mathbf{z})$.

Therefore,

$$\phi u = \delta * (\phi u) = (L(\mathbf{x}, D)F) * (\phi u) = F * (L(\mathbf{x}, D)(\phi u))$$

$$= F * (\phi L(\mathbf{x}, D)u) + F * v \in C^\infty(B_{\epsilon/2}(\mathbf{z}))$$

and, since $\phi(\mathbf{x}) = 1$ for $\mathbf{x} \in B_{\epsilon/2}(\mathbf{z})$, we get that $u \in C^\infty(B_{\epsilon/2}(\mathbf{z}))$. ■

6.8 Exercises

We only suggest elementary problems about distribution and the Fourier transform of functions. For more advanced problems; in particular concerning the Fourier transform of distributions, we suggest [27, 31].

The remark made in the exercise section of Chapter 13 is worth repeating here. Namely, some of the problems below refer to weak solutions of partial differential equations which for us means solutions in the sense of distribution or strong solutions. The expression “weak solutions” is reserved to solutions of variational problems that we will study in Chapter 12.

Suggested exercises:

- In [16]: numbers 1 in Section 3.6 ; all the numbers in Section 1.9.
- In [22]: numbers 3 to 5, 7 to 9 in Sections 2.3.
- In [38]: all the numbers in Sections 12.1 and 12.4 (Warning: the definition of Fourier transform in [38] is slightly different than the definition given above).

Chapter 7

Wave Equation in One-Dimension

This chapter is devoted to the study of the **wave equation**

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

The constant c is called the **(propagation) speed**. We will justify shortly why c is given this name. We will also assume that the solution of the wave equation satisfies some initial and boundary conditions according to the domain of the wave equation.

7.1 D'Alembert Method

7.1.1 Over the Real Line

Consider the wave equation

$$\frac{\partial^2 u}{\partial t^2} - c^2 \frac{\partial^2 u}{\partial x^2} = 0 \quad , \quad t > 0 \text{ and } x \in \mathbb{R} , \quad (7.1.1)$$

with the initial conditions

$$u(x, 0) = f(x) \quad \text{and} \quad \frac{\partial u}{\partial t}(x, 0) = g(x) \quad , \quad x \in \mathbb{R} , \quad (7.1.2)$$

for two functions $f : \mathbb{R} \rightarrow \mathbb{R}$ and $g : \mathbb{R} \rightarrow \mathbb{R}$. The smoothness of the functions f and g will be specified later.

As we have seen in Section 2.3.2, the wave equation is an hyperbolic partial differential equation. We may use a change of variables to reduce the wave equation to the standard form for an hyperbolic equation. We now repeat this reduction process. The wave equation is of the form

$$L(x, t, D)u = a_{2,0} \frac{\partial^2 u}{\partial t^2} + 2a_{1,1} \frac{\partial^2 u}{\partial x \partial t} + a_{0,2} \frac{\partial^2 u}{\partial x^2} = 0$$

with $a_{2,0} = 1$, $a_{1,1} = 0$ and $a_{0,2} = -c^2$. The principal symbol is

$$p(x, t, \xi_1, \xi_2) = \begin{pmatrix} \xi_1 & \xi_2 \end{pmatrix} \begin{pmatrix} a_{2,0} & a_{1,1} \\ a_{1,1} & a_{0,2} \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}.$$

The wave equation is hyperbolic on $\mathbb{R}^2$ because $\det \begin{pmatrix} a_{2,0} & a_{1,1} \\ a_{1,1} & a_{0,2} \end{pmatrix} = -c^2 < 0$. The solutions of the differential equations

$$\frac{dx}{dt} = \frac{a_{1,1} + \sqrt{a_{1,1}^2 - a_{2,0}a_{0,2}}}{a_{2,0}} = c \quad \text{and} \quad \frac{dx}{dt} = \frac{a_{1,1} - \sqrt{a_{1,1}^2 - a_{2,0}a_{0,2}}}{a_{2,0}} = -c$$

are respectively $x = ct + \eta$ and $x = -ct + \xi$. The level curves of $\eta(x, t) = x - ct$ and $\xi(x, t) = x + ct$ are characteristic curves. We use the change of coordinates $\eta = x - ct$ and $\xi = x + ct$ to reduce the wave equation.

Since

$$\frac{\partial u}{\partial x} = \frac{\partial u}{\partial \eta} \frac{\partial \eta}{\partial x} + \frac{\partial u}{\partial \xi} \frac{\partial \xi}{\partial x} = \frac{\partial u}{\partial \eta} + \frac{\partial u}{\partial \xi} \quad \text{and} \quad \frac{\partial u}{\partial t} = \frac{\partial u}{\partial \eta} \frac{\partial \eta}{\partial t} + \frac{\partial u}{\partial \xi} \frac{\partial \xi}{\partial t} = -c \frac{\partial u}{\partial \eta} + c \frac{\partial u}{\partial \xi},$$

we get

$$\begin{aligned} \frac{\partial^2 u}{\partial x^2} &= \frac{\partial}{\partial x} \left(\frac{\partial u}{\partial x} \right) = \frac{\partial}{\partial \eta} \left(\frac{\partial u}{\partial \eta} + \frac{\partial u}{\partial \xi} \right) \frac{\partial \eta}{\partial x} + \frac{\partial}{\partial \xi} \left(\frac{\partial u}{\partial \eta} + \frac{\partial u}{\partial \xi} \right) \frac{\partial \xi}{\partial x} \\ &= \frac{\partial^2 u}{\partial \eta^2} \frac{\partial \eta}{\partial x} + \frac{\partial^2 u}{\partial \eta \partial \xi} \frac{\partial \eta}{\partial x} + \frac{\partial^2 u}{\partial \xi \partial \eta} \frac{\partial \xi}{\partial x} + \frac{\partial^2 u}{\partial \xi^2} \frac{\partial \xi}{\partial x} = \frac{\partial^2 u}{\partial \eta^2} + 2 \frac{\partial^2 u}{\partial \eta \partial \xi} + \frac{\partial^2 u}{\partial \xi^2} \end{aligned}$$

and

$$\begin{aligned} \frac{\partial^2 u}{\partial t^2} &= \frac{\partial}{\partial t} \left(\frac{\partial u}{\partial t} \right) = \frac{\partial}{\partial \eta} \left(-c \frac{\partial u}{\partial \eta} + c \frac{\partial u}{\partial \xi} \right) \frac{\partial \eta}{\partial t} + \frac{\partial}{\partial \xi} \left(-c \frac{\partial u}{\partial \eta} + c \frac{\partial u}{\partial \xi} \right) \frac{\partial \xi}{\partial t} \\ &= -c \left(-c \frac{\partial^2 u}{\partial \eta^2} + c \frac{\partial^2 u}{\partial \eta \partial \xi} \right) + c \left(-c \frac{\partial^2 u}{\partial \xi \partial \eta} + c \frac{\partial^2 u}{\partial \xi^2} \right) = c^2 \left(\frac{\partial^2 u}{\partial \eta^2} - 2 \frac{\partial^2 u}{\partial \eta \partial \xi} + \frac{\partial^2 u}{\partial \xi^2} \right). \end{aligned}$$

Hence, (7.1.1) becomes

$$0 = \frac{\partial^2 u}{\partial t^2} - c^2 \frac{\partial^2 u}{\partial x^2} = c^2 \left(\frac{\partial^2 u}{\partial \eta^2} - 2 \frac{\partial^2 u}{\partial \eta \partial \xi} + \frac{\partial^2 u}{\partial \xi^2} \right) - c^2 \left(\frac{\partial^2 u}{\partial \eta^2} + 2 \frac{\partial^2 u}{\partial \eta \partial \xi} + \frac{\partial^2 u}{\partial \xi^2} \right) = -4c^2 \frac{\partial^2 u}{\partial \eta \partial \xi}.$$

We have to solve the partial differential equation

$$\frac{\partial^2 u}{\partial \eta \partial \xi} = 0. \quad (7.1.3)$$

Integrating (7.1.3) with respect to ξ yields $\frac{\partial u}{\partial \eta} = H(\eta)$ for some function $H : \mathbb{R} \rightarrow \mathbb{R}$. Integrating $\frac{\partial u}{\partial \eta} = H(\eta)$ with respect to η yields $u(\eta, \xi) = \int H(\eta) d\eta + G(\xi)$ for some function $G : \mathbb{R} \rightarrow \mathbb{R}$. If we introduce the function F defined by $F(\eta) = \int H(\eta) d\eta$, we get the solution

$$u(\eta, \xi) = F(\eta) + G(\xi)$$

for (7.1.3). In terms of x and t , we get the solution

$$u(x, t) = F(x - ct) + G(x + ct) \quad (7.1.4)$$

for (7.1.1).

We now consider the initial conditions in (7.1.2). The functions F and G satisfy the equations

$$f(x) = F(x) + G(x) \quad \text{and} \quad g(x) = -cF'(x) + cG'(x) .$$

If we assume that g is locally integrable, we may write

$$\int_0^x g(s) \, ds = -cF(x) + cG(x) + cF(0) - cG(0) .$$

We end up with two linearly independent equations for F and G .

$$F(x) + G(x) = f(x) \quad (7.1.5)$$

and

$$F(x) - G(x) = F(0) - G(0) - \frac{1}{c} \int_0^x g(s) \, ds . \quad (7.1.6)$$

Adding (7.1.5) and (7.1.6), and dividing by 2 yield

$$F(x) = \frac{1}{2}f(x) - \frac{1}{2c} \int_0^x g(s) \, ds + \frac{1}{2}(F(0) - G(0)) .$$

Subtracting (7.1.6) from (7.1.5), and dividing by 2 yield

$$G(x) = \frac{1}{2}f(x) + \frac{1}{2c} \int_0^x g(s) \, ds - \frac{1}{2}(F(0) - G(0)) .$$

We finally get

$$u(x, t) = F(x - ct) + G(x + ct) = \frac{1}{2}(f(x - ct) + f(x + ct)) + \frac{1}{2c} \int_{x-ct}^{x+ct} g(s) \, ds . \quad (7.1.7)$$

If $f \in C^2(\mathbb{R})$ and $g \in C^1(\mathbb{R})$, then (7.1.7) provides a solution of class $C^2(\mathbb{R})$ for (7.1.1) with the initial conditions (7.1.2).

If $g = 0$, then the solution of the wave equation is $u(x, t) = \frac{1}{2}(f(x - ct) + f(x + ct))$. We have illustrated in Figure 7.1 the behaviour of this solution with respect to time.

The value of u at a fixed point x and time t is determined by the values of f at $x \pm ct$ and the values of g between $x - ct$ and $x + ct$. For this reason, the interval $[x - ct, x + ct]$ is called the **domain of determinacy** for u at (x, t) . Conversely, the values of f and g on the interval $[a, b]$ influence only the values of $u(x, t)$ for $a - ct \leq x \leq b + ct$ and $t > 0$. The region $\{(x, t) : t > 0 \text{ and } a - ct \leq x \leq b + ct\}$ is called the **range of influence** of the interval $[a, b]$. These definition are illustrated in Figure 7.2. A discontinuity of the initial condition

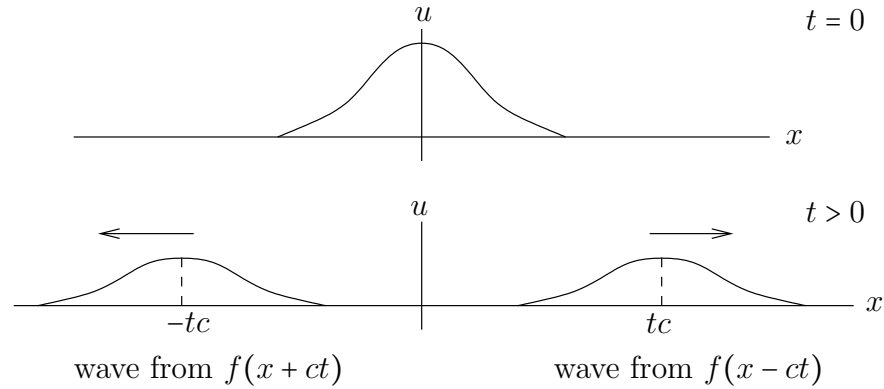


Figure 7.1: There are two waves emanating from the initial deformation and moving at speed c in opposite directions.

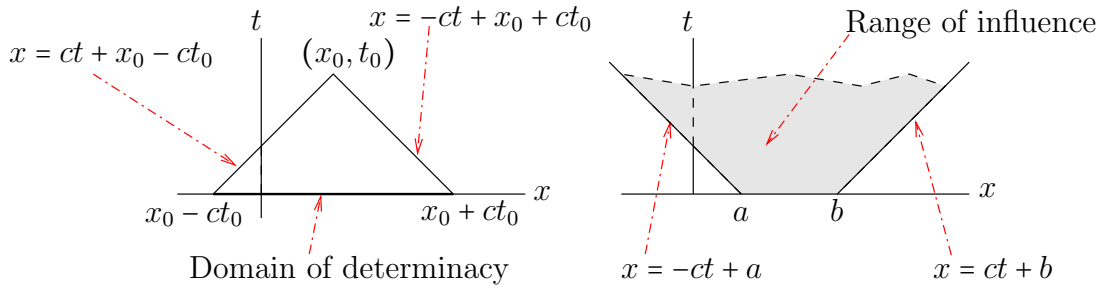


Figure 7.2: Domain of determinacy and range of influence for the wave equation.

on u at $(x_0, 0)$ is instantaneously propagated along the characteristic lines emanating from $(x_0, 0)$

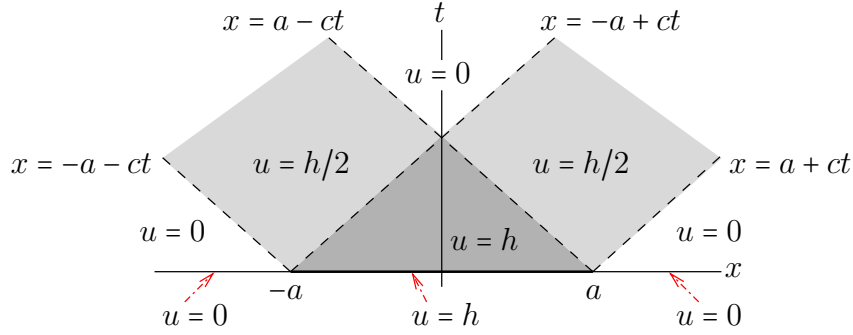
Example 7.1.1

Sketch the solution u of the wave equation $u_{tt} = c^2 u_{xx}$ with the initial conditions

$$u_t(x, 0) = g(x) = 0 \quad \text{and} \quad u(x, 0) = f(x) = \begin{cases} h & \text{if } |x| \leq a \\ 0 & \text{if } |x| > a \end{cases}$$

for $x \in \mathbb{R}$. We assume that a , c and h are arbitrary positive constants.

We have from (7.1.7) that $u(x, t) = f(x - ct) + f(x + ct)$. Hence, we get the following sketch for the solution.



♣

7.1.2 Over a Semi-Finite Interval

In this section, we consider the wave equation

$$\frac{\partial^2 u}{\partial t^2} - c^2 \frac{\partial^2 u}{\partial x^2} = 0 \quad , \quad t > 0 \text{ and } x > 0 \quad ,$$

with the boundary condition

$$u(0, t) = 0 \quad , \quad t > 0 \quad , \quad (7.1.8)$$

and the initial conditions

$$u(x, 0) = f(x) \quad \text{and} \quad \frac{\partial u}{\partial t}(x, 0) = g(x) \quad , \quad x > 0 \quad .$$

As in the previous section, we may use the change of coordinates $\eta = x - ct$ and $\xi = x + ct$ to reduce the wave equation to (7.1.3). Integrating this equation yields

$$u(\eta, \xi) = F(\eta) + G(\xi) \quad (7.1.9)$$

for some functions F and G . In terms of x and t , we get the solution

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

as long as $x - ct > 0$. With the help of the initial conditions as in the previous section, we get

$$u(x, t) = \frac{1}{2} (f(x - ct) + f(x + ct)) + \frac{1}{2c} \int_{x-ct}^{x+ct} g(s) ds \quad (7.1.10)$$

for $x - ct > 0$.

We now consider what happens when $x - ct < 0$. We use the boundary condition (7.1.8) to obtain $0 = F(-ct) + G(ct)$ for $t > 0$. Thus, $F(w) = -G(-w)$ for $w < 0$. We can now proceed as we did for the case $x - ct > 0$.

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

$$\begin{aligned}
&= -\left(\frac{1}{2}f(ct-x) + \frac{1}{2c} \int_0^{ct-x} g(s) ds - \frac{1}{2}(F(0) - G(0))\right) \\
&\quad + \left(\frac{1}{2}f(x+ct) + \frac{1}{2c} \int_0^{x+ct} g(s) ds - \frac{1}{2}(F(0) - G(0))\right) \\
&= \frac{1}{2}(f(x+ct) - f(ct-x)) + \frac{1}{2c} \int_{ct-x}^{ct+x} g(s) ds
\end{aligned} \tag{7.1.11}$$

for $x - ct < 0$.

For $x - ct > 0$, The value of u at a fixed point x and time t is determined by the values of f at $x \pm ct$ and the values of g between $x - ct$ and $x + ct$. For $x - ct < 0$, The value of u at a fixed point x and time t is determined by the values of f at $ct \pm x$ and the values of g between $ct - x$ and $ct + x$. The interval $[ct - x, ct + x]$ is the **domain of determinacy** for u at (x, t) when $x - ct < 0$.

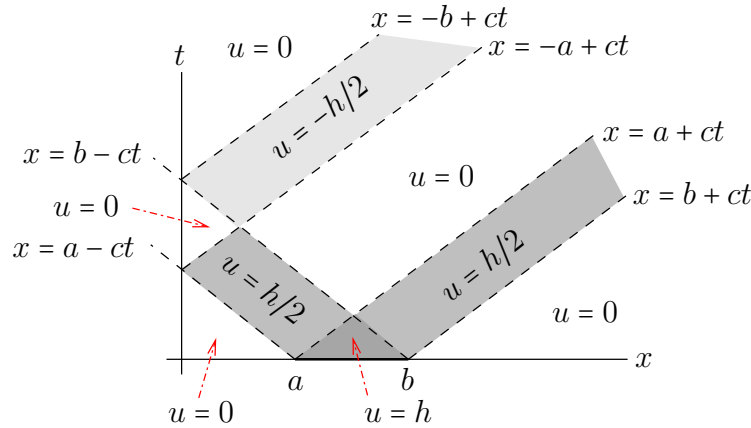
Example 7.1.2

Sketch the solution of the wave equation $u_{tt} = c^2 u_{xx}$ for $x \geq 0$ and $t \in \mathbb{R}$ with the initial conditions

$$u(x, 0) = \begin{cases} 0 & \text{if } 0 < x < a \text{ or } x > b \\ h & \text{if } a \leq x \leq b \end{cases}$$

and $u_t(x, 0) = 0$ for $x \geq 0$, and the boundary condition $u(0, t) = 0$ for $t \geq 0$.

Using (7.1.10) and (7.1.11), we get the following sketch for the solution.



♣

7.1.3 Over a Finite Interval

We first deduce a graphical method to evaluate the solution u of the wave equation

$$\frac{\partial^2 u}{\partial t^2} - c^2 \frac{\partial^2 u}{\partial x^2} = 0 \quad , \quad t > 0 \text{ and } 0 < x < L ,$$

with the boundary conditions

$$u(0, t) = A(t) \quad \text{and} \quad u(L, t) = B(t) \quad , \quad t > 0 ,$$

and the initial conditions

$$u(x, 0) = f(x) \quad \text{and} \quad \frac{\partial u}{\partial t}(x, 0) = g(x) \quad , \quad 0 < x < L .$$

The function A , B , f and g are given.

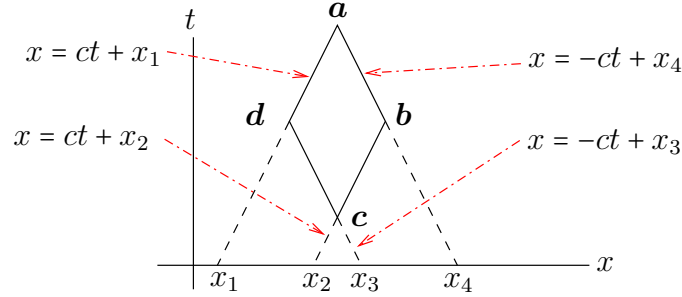


Figure 7.3: Graphical Method to evaluate the solution of the wave equation in one dimension.

The equation (7.1.4) is still true for $0 \leq x - ct < x + ct \leq L$. Consider Figure 7.3. We have that F is constant along the lines $x = ct + x_1$ and $x = ct + x_2$, and G is constant along the lines $x = -ct + x_3$ and $x = -ct + x_4$. Hence,

$$\begin{aligned} u(\mathbf{a}) + u(\mathbf{c}) &= F(x_1) + G(x_4) + F(x_2) + G(x_3) \\ &= F(x_1) + G(x_3) + F(x_2) + G(x_4) = u(\mathbf{d}) + u(\mathbf{b}) . \end{aligned}$$

Thus,

$$u(\mathbf{a}) = u(\mathbf{d}) + u(\mathbf{b}) - u(\mathbf{c}) . \quad (7.1.12)$$

Example 7.1.3

Consider Figure 7.4. The values of u inside the region R_1 are given by (7.1.7). To find the values of u at $\mathbf{a}$ in R_3 , we use (7.1.12). $u(\mathbf{b})$ is given by the boundary condition on $x = L$, $u(\mathbf{c})$ and $u(\mathbf{d})$ are given by the values of u on the boundary of R_1 . Proceeding recursively, we may evaluate u in any region R_i .

Is the solution that we have defined continuous? Inside the region R_i , u will always be continuous if we assume that $f \in C([0, L])$ and $g \in L^1([0, L])$. We only have to consider u along the characteristic lines. Referring to Figure 7.4, we see that

$$u(\mathbf{a}) \rightarrow u(\mathbf{d}) + B(0) - f(L) \quad (7.1.13)$$

as $\mathbf{a}$ approaches the line $x = -ct + L$. If $B(0) = f(L)$, then u will be continuous along $x = -ct + L$ delimiting R_3 . Similarly, for u to be continuous along the line $x = ct$ delimiting the region R_1 , we need $A(0) = f(0)$. We have that $A, B \in C([0, \infty[)$, $g \in L^1([0, L])$ and $f \in C([0, L])$ with $B(0) = f(L)$ and $A(0) = f(0)$ are sufficient to guarantee that u is continuous on the domain $\{(x, t) : t \geq 0, 0 \leq x \leq L\}$.

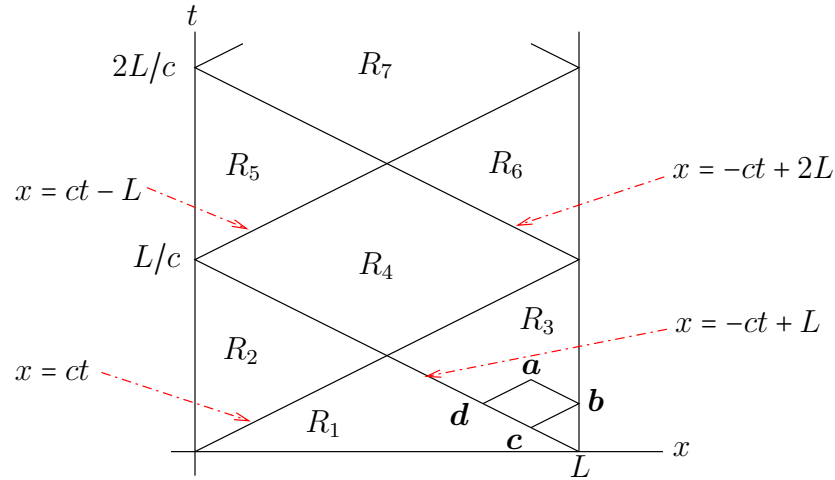


Figure 7.4: D'Alembert method to solve the wave equation with boundary and initial conditions

The solution u will be of class C^1 inside the R_i 's if $f \in C^1(]0, L[)$ and $g \in C(]0, L[)$. The solution u will even be of class C^2 inside the R_i 's if $f \in C^2(]0, L[)$ and $g \in C^1(]0, L[)$. The issue with the smoothness of the solution is along the lines $x = ct$ and $x = -ct + L$, and their reflections on the boundary lines $x = 0$ and $x = L$.

If $A, B \in C^2(]0, \infty[) \cap C([0, \infty[)$, $g \in C^1(]0, L[) \cap C([0, L])$, $f \in C^2(]0, L[) \cap C([0, L])$, $A(0) = f(0)$, $A'(0) = g(0)$, $A''(0) = c^2 f''(0)$, $B(0) = f(L)$, $B'(0) = g(L)$ and $B''(0) = c^2 f''(L)$, then the solution u will be of class C^2 on all the domain. The conditions on the second order derivatives are necessary to satisfies the wave equation on the characteristic lines. The conditions above are the **consistency conditions** to get a classical solution of the wave equation. ♣

7.2 Non-Homogeneous Wave Equation

Consider the **non-homogeneous wave equation**

$$\frac{\partial^2 u}{\partial t^2} - c^2 \frac{\partial^2 u}{\partial x^2} = F(x, t) \quad , \quad t > 0 \text{ and } x \in \mathbb{R} \quad , \quad (7.2.1)$$

with the initial conditions

$$u(x, 0) = f(x) \quad \text{and} \quad \frac{\partial u}{\partial t}(x, 0) = g(x) \quad , \quad x \in \mathbb{R} \quad . \quad (7.2.2)$$

From now on, we assume that F , f and g are sufficiently differentiable.

To find a solution of this problem, we use Green's Integral Formula

$$\iint_K \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial t} \right) dx dt = \int_{\partial K} P dx + Q dt \quad , \quad (7.2.3)$$

where P and Q are continuous differentiable functions in an open set containing $K \in \mathbb{R}^2$, and ∂K is positively oriented. The region $K = K_{x_0, t_0}$ that we consider is drawn in Figure 7.5.

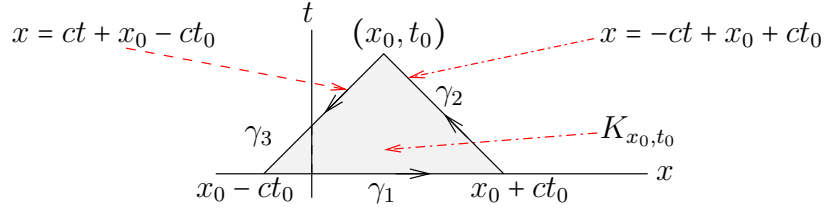


Figure 7.5: Domain of determinacy used to find the solution of the non-homogeneous wave equation (7.2.1) with the initial conditions (7.2.2)

From (7.2.1), we have

$$-\iint_{K_{x_0, t_0}} F(x, t) \, dx \, dt = \iint_{K_{x_0, t_0}} \left(c^2 \frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial t^2} \right) \, dx \, dt = \int_{\Gamma} \frac{\partial u}{\partial t} \, dx + c^2 \frac{\partial u}{\partial x} \, dt ,$$

where we have used Green's Integral Formula (7.2.3) with $P = \frac{\partial u}{\partial t}$ and $Q = c^2 \frac{\partial u}{\partial x}$, and $\Gamma = \gamma_1 + \gamma_2 + \gamma_3$ is the positively oriented boundary of K_{x_0, t_0} as drawn in Figure 7.5.

On γ_1 , we use the parametric representation $(x, t) = (x, 0)$ for $x_0 - ct_0 \leq x \leq x_0 + ct_0$ to get

$$\int_{\gamma_1} \frac{\partial u}{\partial t} \, dx + c^2 \frac{\partial u}{\partial x} \, dt = \int_{x_0 - ct_0}^{x_0 + ct_0} g(x) \, dx$$

because $\frac{\partial u}{\partial t}(x, 0) = g(x)$ for $x \in \mathbb{R}$. On γ_2 , we use the parametric representation $(x, t) = (-ct + x_0 + ct_0, t)$ for $0 \leq t \leq t_0$ to get

$$\begin{aligned} \int_{\gamma_2} \frac{\partial u}{\partial t} \, dx + c^2 \frac{\partial u}{\partial x} \, dt &= -c \int_0^{t_0} \left(\frac{\partial u}{\partial t} - c \frac{\partial u}{\partial x} \right) \, dt = -c \int_0^{t_0} \frac{d}{dt} u(-ct + x_0 + ct_0, t) \, dt \\ &= -c (u(x_0, t_0) - u(x_0 + ct_0, 0)) = -c (u(x_0, t_0) - f(x_0 + ct_0)) \end{aligned}$$

because $u(x, 0) = f(x)$ for $x \in \mathbb{R}$. On γ_3 , we use the parametric representation $(x, t) = (c(t_0 - t) + x_0 - ct_0, t_0 - t)$ for $0 \leq t \leq t_0$ to get

$$\begin{aligned} \int_{\gamma_3} \frac{\partial u}{\partial t} \, dx + c^2 \frac{\partial u}{\partial x} \, dt &= -c \int_0^{t_0} \left(\frac{\partial u}{\partial t} + c \frac{\partial u}{\partial x} \right) \, dt = c \int_0^{t_0} \frac{d}{dt} u(c(t_0 - t) + x_0 - ct_0, t_0 - t) \, dt \\ &= c (u(x_0 - ct_0, 0) - u(x_0, t_0)) = c (f(x_0 - ct_0) - u(x_0, t_0)) \end{aligned}$$

because $u(x, 0) = f(x)$ for $x \in \mathbb{R}$.

Hence,

$$-\iint_{K_{x_0, t_0}} F(x, t) \, dx \, dt = \int_{\gamma_1} \frac{\partial u}{\partial t} \, dx + c^2 \frac{\partial u}{\partial x} \, dt + \int_{\gamma_2} \frac{\partial u}{\partial t} \, dx + c^2 \frac{\partial u}{\partial x} \, dt + \int_{\gamma_3} \frac{\partial u}{\partial t} \, dx + c^2 \frac{\partial u}{\partial x} \, dt$$

$$= \int_{x_0-ct_0}^{x_0+ct_0} g(x) dx + c(f(x_0-ct_0) + f(x_0+ct_0)) - 2cu(x_0, t_0) .$$

If we solved for $u(x_0, t_0)$, we get

$$u(x_0, t_0) = \frac{1}{2} (f(x_0-ct_0) + f(x_0+ct_0)) + \frac{1}{2c} \int_{x_0-ct_0}^{x_0+ct_0} g(x) dx + \frac{1}{2c} \iint_{K_{x_0, t_0}} F(x, t) dx dt$$

This suggest the solution $u = u_1 + u_2$ for (7.2.1) and (7.2.2), where

$$u_1(x, t) = \frac{1}{2} (f(x-ct) + f(x+ct)) + \frac{1}{2c} \int_{x-ct}^{x+ct} g(s) ds$$

is a solution of the homogeneous wave equation (7.1.1) and (7.1.2), and

$$u_2(x, t) = \frac{1}{2c} \iint_{K_{x,t}} F(z, s) dz ds = \frac{1}{2c} \int_0^t \int_{x-c(t-s)}^{x+c(t-s)} F(z, s) dz ds .$$

To prove that u is effectively the solution of our non-homogeneous wave equation, we need to prove that u_2 satisfies (7.2.1) with the initial condition $u(x, 0) = 0$ and $\frac{\partial u}{\partial t}(x, 0) = 0$ for $x \in \mathbb{R}$. To prove this, we need to use the Fundamental Theorem of Calculus several times. We obviously have $u_2(x, 0) = 0$. We also have that

$$\begin{aligned} \frac{\partial u_2}{\partial t}(x, t) &= \frac{1}{2c} \int_x^x F(z, t) dz + \frac{1}{2} \int_0^t (F(x+c(t-s), s) + F(x-c(t-s), s)) ds \\ &= \frac{1}{2} \int_0^t (F(x+c(t-s), s) + F(x-c(t-s), s)) ds . \end{aligned}$$

Thus, $\frac{\partial u_2}{\partial t}(x, 0) = 0$ as desired. Furthermore,

$$\frac{\partial^2 u_2}{\partial t^2}(x, t) = F(x, s) + \frac{c}{2} \int_0^t \left(\frac{\partial F}{\partial x}(x+c(t-s), s) - \frac{\partial F}{\partial x}(x-c(t-s), s) \right) ds . \quad (7.2.4)$$

Similarly,

$$\frac{\partial u_2}{\partial x}(x, t) = \frac{1}{2c} \int_0^t (F(x+c(t-s), s) - F(x-c(t-s), s)) ds$$

and

$$\frac{\partial^2 u_2}{\partial x^2}(x, t) = \frac{1}{2c} \int_0^t \left(\frac{\partial F}{\partial x}(x+c(t-s), s) - \frac{\partial F}{\partial x}(x-c(t-s), s) \right) ds . \quad (7.2.5)$$

(7.2.1) follows from (7.2.4) and (7.2.5).

Remark 7.2.1

We will introduce in Section 8.6 another method, the Duhamel principle, to solve non-homogeneous wave equations. ♠

Example 7.2.2

Solve the non-homogeneous wave equation

$$\frac{\partial^2 u}{\partial t^2} - 4 \frac{\partial^2 u}{\partial x^2} = e^x + \sin(t) \quad , \quad x \in \mathbb{R} \text{ and } t > 0 ,$$

with the initial conditions $u(x, 0) = 0$ and $\frac{\partial u}{\partial t}(x, 0) = \frac{1}{1+x^2}$ for $x \in \mathbb{R}$.

The solution is given by

$$u(x, t) = \frac{1}{2} (f(x-ct) + f(x+ct)) + \frac{1}{2c} \int_{x-ct}^{x+ct} g(s) ds + \frac{1}{2c} \int_0^t \int_{x-c(t-s)}^{x+c(t-s)} F(z, s) dz ds ,$$

where $c = 2$, $f(x) = 0$ for all $x \in \mathbb{R}$, $g(x) = 1/(1+x^2)$ for all $x \in \mathbb{R}$, and $F(x, t) = e^x + \sin(t)$. We get

$$\begin{aligned} u(x, t) &= \frac{1}{4} \int_{x-2t}^{x+2t} \frac{1}{1+s^2} ds + \frac{1}{4} \int_0^t \int_{x-2(t-s)}^{x+2(t-s)} (e^z + \sin(s)) dz ds \\ &= \frac{1}{4} \arctan(s) \Big|_{x-2t}^{x+2t} + \frac{1}{4} \int_0^t (e^{x+2(t-s)} - e^{x-2(t-s)} + 4(t-s)\sin(s)) ds \\ &= \frac{1}{4} (\arctan(x+2t) - \arctan(x-2t)) + \frac{1}{8} (e^{x+2t} - e^{x-2t}) - \frac{e^x}{4} + t - \sin(t) . \end{aligned}$$

♣

Remark 7.2.3

For the previous example, there is a more direct approach to solve the non-homogeneous wave equation.

It is easy to see that $u_1(x, t) = -\sin(t)$ is a solution of

$$\frac{\partial^2 u}{\partial t^2} - 4 \frac{\partial^2 u}{\partial x^2} = \sin(t) \quad , \quad x \in \mathbb{R} \text{ and } t > 0 ,$$

and $u_2(x, t) = -e^x/4$ is a solution of

$$\frac{\partial^2 u}{\partial t^2} - 4 \frac{\partial^2 u}{\partial x^2} = e^x \quad , \quad x \in \mathbb{R} \text{ and } t > 0 .$$

The solution of the non-homogeneous problem of the previous example is therefore $u = u_1 + u_2 + u_3$, where u_3 is a solution of the homogeneous wave equation

$$\frac{\partial^2 u}{\partial t^2} - 4 \frac{\partial^2 u}{\partial x^2} = 0 \quad , \quad x \in \mathbb{R} \text{ and } t > 0 ,$$

with the initial conditions

$$u(x, 0) = 0 - u_1(x, 0) - u_2(x, 0) = \frac{e^x}{4}$$

and

$$\frac{\partial u}{\partial t}(x, 0) = \frac{1}{1+x^2} - \frac{\partial u_1}{\partial t}(x, 0) - \frac{\partial u_2}{\partial t}(x, 0) = \frac{1}{1+x^2} + 1 .$$

We find

$$u_3(x, t) = \frac{1}{2} \left(\frac{e^{x-2t}}{4} + \frac{e^{x+2t}}{4} \right) + \frac{1}{4} \int_{x-2t}^{x+2t} \left(\frac{1}{1+s^2} + 1 \right) ds$$

$$\begin{aligned}
&= \frac{e^{x-2t}}{8} + \frac{e^{x+2t}}{8} + \frac{1}{4} (\arctan(s) + s) \Big|_{x-2t}^{x+2t} \\
&= \frac{e^{x-2t}}{8} + \frac{e^{x+2t}}{8} + \frac{1}{4} (\arctan(x+2t) - \arctan(x-2t)) + t .
\end{aligned}$$

♠

7.3 Strong Solution of the Wave equation

Consider the hyperbolic linear partial differential equation of order two

$$L(\eta, \xi, D)u = \frac{\partial^2 u}{\partial \eta \partial \xi} + A_{1,0} \frac{\partial u}{\partial \eta} + A_{0,1} \frac{\partial u}{\partial \xi} + A_{0,0}u = 0 \quad (7.3.1)$$

which is the general form of a linear homogeneous hyperbolic partial differential equation of order two after reduction using a change of variables provided by level curves which are characteristic curves.

The adjoint of $L(\eta, \xi, D)$ is defined by

$$L^*(\eta, \xi, D)\phi = \frac{\partial^2 \phi}{\partial \eta \partial \xi} - \frac{\partial}{\partial \eta}(A_{1,0}\phi) - \frac{\partial}{\partial \xi}(A_{0,1}\phi) + A_{0,0}\phi$$

for $\phi \in \mathcal{D}(\mathbb{R}^2)$.

$u : \Omega \rightarrow \mathbb{R}$ is a strong solution of $L(\eta, \xi, D)u = 0$ over an open set $\Omega \subset \mathbb{R}^2$ if

$$\int_{\Omega} u(\eta, \xi) L^*(\eta, \xi, D)\phi(\eta, \xi) d\eta d\xi = 0 \quad (7.3.2)$$

for all $\phi \in \mathcal{D}(\Omega)$. This is satisfied if $u \in C^2(\Omega)$ is a solution of (7.3.1) as can be shown using integration by parts.

Suppose that Ω is traversed by a characteristic curve Γ along which the smoothness of u is broken. Since the characteristic curves are the level curves of the change of variables used to get the reduced form of the hyperbolic equation, we may assume that the characteristic curve Γ traversing Ω is the line $\xi = 0$ as in Figure 7.6.

We have $\Omega = \Omega_+ \cup \Omega_0 \cup \Omega_-$, where $\Omega_+ = \{(\eta, \xi) \in \Omega : \xi > 0\}$, $\Omega_- = \{(\eta, \xi) \in \Omega : \xi < 0\}$ and $\Omega_0 = \{(\eta, \xi) \in \Omega : \xi = 0\}$.

Suppose that $u : \Omega \rightarrow \mathbb{R}$ is a strong solution which also satisfies (7.3.1) in the classical sense on Ω_+ and Ω_- . The fact that u is a strong solution imposes some extra conditions on u along Ω_0 .

We may assume that $u(\eta, \xi) = 0$ for $(\eta, \xi) \notin \Omega$. We may also assume that $\phi \in \mathcal{D}(\Omega)$ is in fact in $\mathcal{D}(\mathbb{R}^2)$ and has a compact support in Ω . This is something that we will often do to simplify the discussion when we have to work on an open subset of $\mathbb{R}^n$.

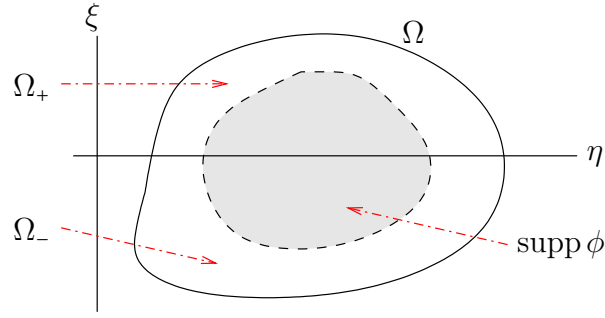


Figure 7.6: The Transmission condition for a strong solution across a characteristic line represented by the η axis.

Using integration by parts, we get for $\phi \in \mathcal{D}(\Omega)$ that

$$\begin{aligned}
 & \int_{\Omega} u(\eta, \xi) L^*(\eta, \xi, D) \phi(\eta, \xi) d\eta d\xi \\
 &= \int_{\Omega_+} u(\eta, \xi) L^*(\eta, \xi, D) \phi(\eta, \xi) d\eta d\xi + \int_{\Omega_-} u(\eta, \xi) L^*(\eta, \xi, D) \phi(\eta, \xi) d\eta d\xi \\
 &= \int_0^{+\infty} \int_{-\infty}^{+\infty} u \left(\frac{\partial^2 \phi}{\partial \eta \partial \xi} - \frac{\partial}{\partial \eta} (A_{1,0} \phi) - \frac{\partial}{\partial \xi} (A_{0,1} \phi) + A_{0,0} \phi \right) d\eta d\xi \\
 &\quad + \int_{-\infty}^0 \int_{-\infty}^{+\infty} u \left(\frac{\partial^2 \phi}{\partial \eta \partial \xi} - \frac{\partial}{\partial \eta} (A_{1,0} \phi) - \frac{\partial}{\partial \xi} (A_{0,1} \phi) + A_{0,0} \phi \right) d\eta d\xi \\
 &= \int_0^{+\infty} \left(u \frac{\partial \phi}{\partial \xi} \Big|_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} \frac{\partial u}{\partial \eta} \frac{\partial \phi}{\partial \xi} d\eta \right) d\xi - \int_0^{+\infty} \left(u A_{1,0} \phi \Big|_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} \frac{\partial u}{\partial \eta} A_{1,0} \phi d\eta \right) d\xi \\
 &\quad + \int_0^{+\infty} \int_{-\infty}^{+\infty} u \left(-\frac{\partial}{\partial \xi} (A_{0,1} \phi) + A_{0,0} \phi \right) d\eta d\xi \\
 &\quad + \int_{-\infty}^0 \left(u \frac{\partial \phi}{\partial \xi} \Big|_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} \frac{\partial u}{\partial \eta} \frac{\partial \phi}{\partial \xi} d\eta \right) d\xi - \int_{-\infty}^0 \left(u A_{1,0} \phi \Big|_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} \frac{\partial u}{\partial \eta} A_{1,0} \phi d\eta \right) d\xi \\
 &\quad + \int_{-\infty}^0 \int_{-\infty}^{+\infty} u \left(-\frac{\partial}{\partial \xi} (A_{0,1} \phi) + A_{0,0} \phi \right) d\eta d\xi \\
 &= \int_{-\infty}^{+\infty} \int_0^{+\infty} \left(-\frac{\partial u}{\partial \eta} \frac{\partial \phi}{\partial \xi} + \frac{\partial u}{\partial \eta} A_{1,0} \phi - u \frac{\partial}{\partial \xi} (A_{0,1} \phi) + u A_{0,0} \phi \right) d\xi d\eta \\
 &\quad + \int_{-\infty}^{+\infty} \int_{-\infty}^0 \left(-\frac{\partial u}{\partial \eta} \frac{\partial \phi}{\partial \xi} + \frac{\partial u}{\partial \eta} A_{1,0} \phi - u \frac{\partial}{\partial \xi} (A_{0,1} \phi) + u A_{0,0} \phi \right) d\xi d\eta \\
 &= \int_{-\infty}^{+\infty} \left(-\frac{\partial u}{\partial \eta} \phi \Big|_0^{+\infty} + \int_0^{+\infty} \frac{\partial^2 u}{\partial \eta \partial \xi} \phi d\xi \right) d\eta - \int_{-\infty}^{+\infty} \left(u A_{0,1} \phi \Big|_0^{+\infty} - \int_0^{+\infty} \frac{\partial u}{\partial \xi} A_{0,1} \phi d\xi \right) d\eta \\
 &\quad + \int_0^{+\infty} \int_{-\infty}^{+\infty} \left(\frac{\partial u}{\partial \eta} A_{1,0} \phi + u A_{0,0} \phi \right) d\eta d\xi \\
 &\quad + \int_{-\infty}^{+\infty} \left(-\frac{\partial u}{\partial \eta} \phi \Big|_{-\infty}^0 + \int_{-\infty}^0 \frac{\partial^2 u}{\partial \eta \partial \xi} \phi d\xi \right) d\eta - \int_{-\infty}^{+\infty} \left(u A_{0,1} \phi \Big|_{-\infty}^0 - \int_{-\infty}^0 \frac{\partial u}{\partial \xi} A_{0,1} \phi d\xi \right) d\eta
 \end{aligned}$$

$$\begin{aligned}
& + \int_{-\infty}^0 \int_{-\infty}^{+\infty} \left(\frac{\partial u}{\partial \eta} A_{1,0} \phi + u A_{0,0} \phi \right) d\eta d\xi \\
& = \int_{\Omega_+} (L(\eta, \xi, D)u(\eta, \xi)) \phi(\eta, \xi) d\eta d\xi + \int_{\Omega_-} (L(\eta, \xi, D)u(\eta, \xi)) \phi(\eta, \xi) dx dt \\
& + \int_{\Omega_0} \left(u_+(\eta, 0) A_{0,1}(\eta, 0) \phi(\eta, 0) + \frac{\partial u_+}{\partial \eta}(\eta, 0) \phi(\eta, 0) \right) d\eta \\
& - \int_{\Omega_0} \left(u_-(\eta, 0) A_{0,1}(\eta, 0) \phi(\eta, 0) + \frac{\partial u_-}{\partial \eta}(\eta, 0) \phi(\eta, 0) \right) d\eta \\
& = \int_{\Omega_0} (u_+(\eta, 0) - u_-(\eta, 0)) A_{0,1}(\eta, 0) \phi(\eta, 0) d\eta + \int_{\Omega_0} \left(\frac{\partial u_+}{\partial \eta}(\eta, 0) - \frac{\partial u_-}{\partial \eta}(\eta, 0) \right) \phi(\eta, 0) d\eta
\end{aligned}$$

where $u_{\pm}(\eta, 0) = \lim_{\xi \rightarrow 0^{\pm}} u(\eta, \xi)$ and $\frac{\partial u_{\pm}}{\partial \eta}(\eta, 0) = \lim_{\xi \rightarrow 0^{\pm}} \frac{\partial u}{\partial \eta}(\eta, \xi)$. We assume that these limits exist.

To get $\int_R u(\eta, \xi) L^*(\eta, \xi, D) \phi(\eta, \xi) d\eta d\xi = 0$ for all $\phi \in \mathcal{D}(\Omega)$, we must have

$$\int_{\Omega_0} (u_+(\eta, 0) - u_-(\eta, 0)) A_{0,1}(\eta, 0) \phi(\eta, 0) d\eta + \int_{\Omega_0} \left(\frac{\partial u_+}{\partial \eta}(\eta, 0) - \frac{\partial u_-}{\partial \eta}(\eta, 0) \right) \phi(\eta, 0) d\eta = 0$$

for all $\phi \in \mathcal{D}(\Omega)$. This is the **transmission condition** to get a strong solution.

If we consider the wave equation $L(\eta, \xi, D)u = \frac{\partial^2 u}{\partial \eta \partial \xi} = 0$, the transmission condition for u to be a strong solution is

$$\int_{\Omega_0} \left(\frac{\partial u_+}{\partial \eta}(\eta, 0) - \frac{\partial u_-}{\partial \eta}(\eta, 0) \right) \phi(\eta, 0) d\eta = 0 \quad (7.3.3)$$

for all $\phi \in \mathcal{D}(\Omega)$. Since this is true for all test functions $\phi \in \mathcal{D}(\Omega)$, we must have

$$\frac{\partial u_+}{\partial \eta}(\eta, 0) - \frac{\partial u_-}{\partial \eta}(\eta, 0) = 0$$

for almost every $\eta \in \mathbb{R}$. If u and $\frac{\partial u}{\partial \eta}$ can be continuously extended to the ξ axis, then this shows that $u_+(\eta, 0) - u_-(\eta, 0)$ is constant for $\eta \in \mathbb{R}$.

Example 7.3.1

In Example 7.1.3, if the functions f , g , A and B are C^∞ . then the solution u constructed in the example is a strong solution because the transmission condition is satisfied. We have a constant jump discontinuity along the characteristic lines $x = ct$ and $x = -ct + L$, and their reflections through the boundary $x = 0$ and $x = L$. For instance, it follows from (7.1.13) that $\lim_{\substack{a \rightarrow d \\ a \in E_3}} u(a) = u(d) + B(0) - f(L)$ where the region E_3 is defined in Figure 7.4. ♣

7.4 Separation of variables

In this section, we use the method of separation of variables to solve the wave equation with some initial and boundary conditions.

The method of separation of variables can generally be used to find strong solutions for the wave equation. We will come back on this subject later on. The next theorem states that if the initial conditions are smooth enough, then the method of separation of variables yields a classical solution.

Theorem 7.4.1

Suppose that $f \in C^3([0, L])$, $g \in C^2([0, L])$, $f(0) = f(L) = 0$, $f''(0) = f''(L) = 0$ and $g(0) = g(L) = 0$. Let

$$u(x, t) = \sum_{n=1}^{\infty} \left(a_n \cos\left(\frac{n\pi t}{L}\right) + b_n \sin\left(\frac{n\pi t}{L}\right) \right) \sin\left(\frac{n\pi x}{L}\right) \quad (7.4.1)$$

with

$$a_n = \frac{2}{L} \int_0^L f(x) \sin\left(\frac{n\pi x}{L}\right) dx \quad \text{and} \quad b_n = \frac{2}{n\pi} \int_0^L g(x) \sin\left(\frac{n\pi x}{L}\right) dx .$$

Then u is a classical solution of the wave equation

$$\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} = 0 \quad , \quad t > 0 \text{ and } 0 < x < L , \quad (7.4.2)$$

with the boundary condition

$$u(0, t) = 0 \quad \text{and} \quad u(L, t) = 0 \quad , \quad t > 0 ,$$

and the initial condition

$$u(x, 0) = f(x) \quad \text{and} \quad \frac{\partial u}{\partial t}(x, 0) = g(x) \quad , \quad 0 < x < L .$$

Proof.

To prove the theorem, it is enough to show that

$$\begin{aligned} u(x, t) &= \sum_{n=1}^{\infty} \left(a_n \cos\left(\frac{n\pi t}{L}\right) + b_n \sin\left(\frac{n\pi t}{L}\right) \right) \sin\left(\frac{n\pi x}{L}\right) , \\ \frac{\partial u}{\partial t}(x, t) &= \sum_{n=1}^{\infty} \frac{n\pi}{L} \left(-a_n \sin\left(\frac{n\pi t}{L}\right) + b_n \cos\left(\frac{n\pi t}{L}\right) \right) \sin\left(\frac{n\pi x}{L}\right) , \\ \frac{\partial^2 u}{\partial t^2}(x, t) &= - \sum_{n=1}^{\infty} \frac{n^2 \pi^2}{L^2} \left(a_n \cos\left(\frac{n\pi t}{L}\right) + b_n \sin\left(\frac{n\pi t}{L}\right) \right) \sin\left(\frac{n\pi x}{L}\right) , \\ \frac{\partial u}{\partial x}(x, t) &= \sum_{n=1}^{\infty} \frac{n\pi}{L} \left(a_n \cos\left(\frac{n\pi t}{L}\right) + b_n \sin\left(\frac{n\pi t}{L}\right) \right) \cos\left(\frac{n\pi x}{L}\right) , \\ \frac{\partial^2 u}{\partial x^2}(x, t) &= - \sum_{n=1}^{\infty} \frac{n^2 \pi^2}{L^2} \left(a_n \cos\left(\frac{n\pi t}{L}\right) + b_n \sin\left(\frac{n\pi t}{L}\right) \right) \sin\left(\frac{n\pi x}{L}\right) \end{aligned}$$

and

$$\frac{\partial^2 u}{\partial t \partial x}(x, t) = \sum_{n=1}^{\infty} \frac{n^2 \pi^2}{L^2} \left(-a_n \sin\left(\frac{n\pi t}{L}\right) + b_n \cos\left(\frac{n\pi t}{L}\right) \right) \cos\left(\frac{n\pi x}{L}\right)$$

converge uniformly on $[0, L] \times [0, T]$ whatever $T > 0$.

Assuming the uniform convergence of the series above, we have that u is continuous on $[0, L] \times [0, \infty[$ because it is continuous on each set $[0, L] \times [0, T]$ being the uniform limit in (7.4.1) of continuous functions. We can also interchange the summation in (7.4.1) and the derivatives $\frac{\partial}{\partial t}$, $\frac{\partial^2}{\partial t^2}$, $\frac{\partial}{\partial x}$, $\frac{\partial^2}{\partial x^2}$ and $\frac{\partial^2}{\partial x \partial t}$. Since each term of the series in (7.4.1) is twice differentiable, it follows that the same is true for u . Thus $u \in C^2([0, L] \times [0, \infty[)$. Moreover, u satisfies (7.4.2) because each term of the series in (7.4.1) satisfies (7.4.2). The boundary conditions are also satisfied by each term of the series in (7.4.1). The initial conditions are satisfied because the a_n and $n\pi b_n/L$ are the coefficients for the Fourier sine series of f and g respectively on the interval $[0, L]$.

To prove the convergence of all the series above, it is enough to prove that $\sum_{n=1}^{\infty} n^2 |a_n|$ and $\sum_{n=1}^{\infty} n^2 |b_n|$ converge.

Using integration by parts, we find

$$a_n = \frac{2}{L} \int_0^L f(x) \sin\left(\frac{n\pi x}{L}\right) dx = -\frac{2}{L} \left(\frac{L}{n\pi}\right)^3 \int_0^L \frac{d^3 f}{dx^3}(x) \cos\left(\frac{n\pi x}{L}\right) dx.$$

Thus, for $N > 0$, we have

$$\begin{aligned} 0 &\leq \frac{2}{L} \int_0^L \left(\frac{d^3 f}{dx^3}(x) + \left(\frac{\pi}{L}\right)^3 \sum_{n=1}^N n^3 a_n \cos\left(\frac{n\pi x}{L}\right) \right)^2 dx \\ &= \frac{2}{L} \int_0^L \left(\frac{d^3 f}{dx^3}(x) \right)^2 dx + \frac{4}{L} \left(\frac{\pi}{L}\right)^3 \sum_{n=1}^N n^3 a_n \underbrace{\int_0^L \frac{d^3 f}{dx^3}(x) \cos\left(\frac{n\pi x}{L}\right) dx}_{=-(a_n L/2)(n\pi/L)^3} \\ &\quad + \left(\frac{\pi}{L}\right)^6 \sum_{n=1}^N \sum_{m=1}^N n^3 a_n m^3 a_m \underbrace{\int_0^L \cos\left(\frac{n\pi x}{L}\right) \cos\left(\frac{m\pi x}{L}\right) dx}_{= \begin{cases} 0 & \text{if } n \neq m \\ 1/2 & \text{if } n = m \end{cases}} \\ &= \frac{2}{L} \int_0^L \left(\frac{d^3 f}{dx^3}(x) \right)^2 dx - 2 \left(\frac{\pi}{L}\right)^6 \sum_{n=1}^N n^6 a_n^2 + \left(\frac{\pi}{L}\right)^6 \sum_{n=1}^N n^6 a_n^2. \end{aligned}$$

It follows that

$$S_N \equiv \sum_{n=1}^N n^6 a_n^2 \leq \frac{2}{L} \left(\frac{L}{\pi}\right)^6 \int_0^L \left(\frac{d^3 f}{dx^3}(x) \right)^2 dx < \infty$$

for all $N > 0$. Since $\{S_N\}_{N=1}^{\infty}$ is a bounded increasing sequence, it converges. Namely, $\sum_{n=1}^{\infty} n^6 a_n^2$ is a convergent series of positive terms.

Using integration by parts, we find

$$b_n = \frac{2}{n\pi} \int_0^L g(x) \sin\left(\frac{n\pi x}{L}\right) dx = -\frac{2}{L} \left(\frac{L}{n\pi}\right)^3 \int_0^L \frac{d^2 g}{dx^2}(x) \sin\left(\frac{n\pi x}{L}\right) dx .$$

A reasoning similar to the previous one shows that $\sum_{n=1}^{\infty} n^6 b_n^2$ is a convergent series of positive terms.

Since

$$n^2 |a_n| = \frac{1}{2} \left(\frac{1}{n^2} + n^6 a_n^2 \right) - \frac{1}{2} \left(\frac{1}{n} - n^3 |a_n| \right)^2 \leq \frac{1}{2} \left(\frac{1}{n^2} + n^6 a_n^2 \right)$$

for all $n > 0$, we can use the comparison test for series of positive terms to conclude that $\sum_{n=1}^{\infty} n^2 |a_n|$ converges. Similarly, we have that $n^2 |b_n| \leq \frac{1}{2} \left(\frac{1}{n^2} + n^6 b_n^2 \right)$ and $\sum_{n=1}^{\infty} n^2 |b_n|$ converges. ■

Theorems like the previous one can be given for different boundary conditions. We will not provide an encyclopedic list of the results.

Remark 7.4.2

The function $u_n(x, t) = \left(a_n \cos\left(\frac{n\pi t}{L}\right) + b_n \sin\left(\frac{n\pi t}{L}\right) \right) \sin\left(\frac{n\pi x}{L}\right)$ is called the **n^{th} -normal mode** or **n^{th} -harmonic**. The first-harmonic is called the **fundamental harmonic**. The frequency of the fundamental harmonic is $1/(2L)$. All the other harmonics have frequencies which are integer multiples of $1/(2L)$. This is the reason why a well tuned instrument can produce good quality tones. ♠

Example 7.4.3

For our first example, we find the deflection $u(x, t)$ of a vibrating string of length $L = \pi$ attached at both ends. We assume that the tension T and density ρ of the string are such that $c^2 = T/\rho = 1$. The initial deflection is $f(x) = 0.1 \sin(x)$ and the initial velocity is $g(x) = -0.2 \sin(x)$.

The solution is given by (7.4.1) with $L = \pi$,

$$a_n = \frac{2}{\pi} \int_0^{\pi} f(x) \sin(nx) dx = \begin{cases} 0.1 & \text{if } n = 1 \\ 0 & \text{if } n > 1 \end{cases} \quad (7.4.3)$$

and

$$b_n = \frac{2}{n\pi} \int_0^{\pi} g(x) \sin(nx) dx = \begin{cases} -0.2 & \text{if } n = 1 \\ 0 & \text{if } n > 1 \end{cases} \quad (7.4.4)$$

Thus,

$$u(x, t) = (0, 1 \cos(t) - 0.2 \sin(t)) \sin(x) .$$

There was no need to compute the integrals to find the values of a_n and b_n . From $f(x) = u(x, 0)$, we have

$$0.1 \sin(x) = \sum_{n=1}^{\infty} a_n \sin(nx) .$$

By uniqueness of the Fourier series, we get (7.4.3). Likewise, from $g(x) = \frac{\partial u}{\partial t}(x, 0)$, we have

$$-0.2 \sin(x) = \sum_{n=1}^{\infty} n b_n \sin(nx) .$$

This time, by uniqueness of the Fourier series, we get (7.4.4). ♣

In the following examples, we will find the solutions starting from separation of variables. This allows us to deal with boundary conditions different than those in Theorem 7.4.1 if needed. The presentation will be a little bit formal. We will not study the convergence of the series solutions.

Example 7.4.4

We again consider a special case of Example 7.1.3.

$$\frac{\partial^2 u}{\partial t^2} - c^2 \frac{\partial^2 u}{\partial x^2} = 0 \quad , \quad t > 0 \text{ and } 0 < x < L , \quad (7.4.5)$$

with the boundary conditions $u(0, t) = 0$ and $u(L, t) = 0$ for $t > 0$, and the initial conditions $u(x, 0) = f(x)$ and $\frac{\partial u}{\partial t}(x, 0) = g(x)$ for $0 < x < L$.

Separation of Variables: If we substitute $u(x, t) = F(x)G(t)$ in (7.4.5), we get

$$F(x) \frac{d^2 G}{dt^2}(t) = c^2 \frac{d^2 F}{dx^2}(x) G(t) .$$

Thus, after dividing both sides by $c^2 F(x)G(t)$, we get

$$\frac{1}{c^2 G(t)} \frac{d^2 G}{dt^2}(t) = \frac{1}{F(x)} \frac{d^2 F}{dx^2}(x)$$

for $t > 0$ and $0 < x < L$. Since the right hand side is independent of t and the left hand side is independent of x , we get

$$\frac{1}{c^2 G(t)} \frac{d^2 G}{dt^2}(t) = \frac{1}{F(x)} \frac{d^2 F}{dx^2}(x) = k$$

for $t > 0$ and $0 < x < L$, and some constant k . We end up with two ordinary differential equations.

$$\frac{d^2 F}{dx^2}(x) - k F(x) = 0 \quad \text{and} \quad \frac{d^2 G}{dt^2}(t) - c^2 k G(t) = 0 . \quad (7.4.6)$$

The first ordinary differential equation in (7.4.6) satisfies two boundary conditions. From $u(0, t) = 0$, we get $F(0)G(t) = 0$. Since we assume that G is non-null, we get $F(0) = 0$.

Similarly, from $u(L, t) = 0$, we get $F(L)G(t) = 0$. Again, since we assume that G is non-null, we get $F(L) = 0$. The boundary conditions for the first ordinary differential equation in (7.4.6) are $F(0) = F(L) = 0$.

Simple Functions: We first consider the boundary value problem

$$\frac{d^2 F}{dx^2}(x) - kF(x) = 0 \quad \text{with} \quad F(0) = F(L) = 0. \quad (7.4.7)$$

The form of the general solution of the boundary value problem above is determined by the roots of the characteristic equation $\lambda^2 - k = 0$.

If $k > 0$, the roots of the characteristic equation are $\pm\sqrt{k}$. Since the roots are real, the solution of the ordinary differential equation is of the form $F(x) = Ae^{\sqrt{k}x} + Be^{-\sqrt{k}x}$. However, $F(0) = 0$ implies that $A + B = 0$ and $F(L) = 0$ implies $Ae^{L\sqrt{k}} + Be^{-L\sqrt{k}} = 0$. The only solution of the system

$$\begin{pmatrix} 1 & 1 \\ e^{L\sqrt{k}} & e^{-L\sqrt{k}} \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

is $A = B = 0$ because

$$\det \begin{pmatrix} 1 & 1 \\ e^{L\sqrt{k}} & e^{-L\sqrt{k}} \end{pmatrix} = e^{-L\sqrt{k}} - e^{L\sqrt{k}} = e^{-L\sqrt{k}} (1 - e^{2L\sqrt{k}}) \neq 0$$

for $k \neq 0$. Therefore, the null solution is the only solution of the boundary value problem (7.4.7) for $k > 0$.

If $k = 0$, the solution of the ordinary differential equation is $F(x) = Bx + A$. However, $F(0) = 0$ implies that $A = 0$ and $F(L) = 0$ implies that $BL + A = 0$. Thus, $A = B = 0$. The boundary value problem (7.4.7) has no non-null solutions for $k = 0$.

If $k < 0$, the roots of the characteristic equation are $\pm i\sqrt{-k}$. Since the roots are complex, the general solution of the ordinary differential equation is of the form $F(x) = A \cos(\sqrt{-k}x) + B \sin(\sqrt{-k}x)$. Hence, $F(0) = 0$ implies that $A = 0$ and $F(L) = 0$ implies that $B \sin(L\sqrt{-k}) = 0$. If we take $B = 0$, we get the null solution. We must therefore have $\sin(L\sqrt{-k}) = 0$ with $k < 0$. This implies that $k = k_n = -(n\pi/L)^2$ for $n > 0$. The boundary value problem (7.4.7) has non-null solutions only for $k = k_n = -(n\pi/L)^2 < 0$ with $n > 0$, and the solutions associated to k_n is of the form

$$F(x) = F_n(x) \equiv B_n \sin\left(\frac{n\pi x}{L}\right).$$

For $k = k_n = -(n\pi/L)^2$, the second ordinary differential equation in (7.4.6) becomes

$$\frac{d^2 G}{dt^2}(t) + \left(\frac{cn\pi}{L}\right)^2 G(t) = 0$$

for $n > 0$. For each n , this is a second order ordinary differential equation with the characteristic equation $\lambda^2 + (cn\pi/L)^2 = 0$. The two roots of this equation are the complex numbers $\lambda_{\pm} = \pm(cn\pi/L)i$. The general solution of this ordinary differential equation is therefore

$$G(t) = G_n(t) \equiv C_n \cos\left(\frac{cn\pi t}{L}\right) + D_n \sin\left(\frac{cn\pi t}{L}\right).$$

We have found the simple functions

$$u_n(x, t) \equiv F_n(x)G_n(t) = \left(a_n \cos\left(\frac{cn\pi t}{L}\right) + b_n \sin\left(\frac{cn\pi t}{L}\right) \right) \sin\left(\frac{n\pi x}{L}\right)$$

for $n > 0$ that satisfy the partial differential equation in (7.4.5) and the boundary conditions. The constant a_n is the product of the constants B_n and C_n , and b_n is the product of B_n and D_n . We now need to determine the values of a_n and b_n to satisfy the initial conditions.

Initial condition: We seek a solution of the form

$$u(x, t) = \sum_{n=1}^{\infty} u_n(x, t) = \sum_{n=1}^{\infty} \left(a_n \cos\left(\frac{cn\pi t}{L}\right) + b_n \sin\left(\frac{cn\pi t}{L}\right) \right) \sin\left(\frac{n\pi x}{L}\right) .$$

From $u(x, 0) = f(x)$, we get

$$f(x) = \sum_{n=1}^{\infty} a_n \sin\left(\frac{n\pi x}{L}\right)$$

for $0 < x < L$. This is the Fourier sine series of f . The coefficients of this series are given by

$$a_n = \frac{2}{L} \int_0^L f(x) \sin\left(\frac{n\pi x}{L}\right) dx .$$

From $\frac{\partial u}{\partial t}(x, 0) = g(x)$, we get

$$g(x) = \sum_{n=1}^{\infty} \frac{cn\pi b_n}{L} \sin\left(\frac{n\pi x}{L}\right)$$

for $0 < x < L$. This is the Fourier sine series of g . The formula to compute the coefficients of this Fourier series yields

$$b_n = \frac{2}{cn\pi} \int_0^L g(x) \cos\left(\frac{n\pi x}{L}\right) dx .$$

To show that we effectively get a solution of the wave equation, we should prove that the series converges as we have done for Theorem 7.4.1 if we seek a classical solution. This is left to the reader. ♣

Example 7.4.5

Solve the wave equation

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2} \quad , \quad 0 < x < L \text{ and } t > 0 , \quad (7.4.8)$$

with the boundary conditions $\frac{\partial u}{\partial x}(0, t) = \frac{\partial u}{\partial x}(L, t) = 0$ for $t > 0$, and the initial conditions $u(x, 0) = f(x)$ and $\frac{\partial u}{\partial t}(x, 0) = g(x)$ for $0 < x < L$.

Separation of Variables: If we substitute $u(x, t) = F(x)G(t)$ in (7.4.8), we get

$$F(x) \frac{d^2 G}{dt^2}(t) = c^2 \frac{d^2 F}{dx^2}(x) G(t) .$$

After dividing both sides by $c^2 F(x)G(t)$, we get

$$\frac{1}{c^2 G(t)} \frac{d^2 G}{dt^2}(t) = \frac{1}{F(x)} \frac{d^2 F}{dx^2}(x)$$

for $t > 0$ and $0 < x < L$. Since the right hand side is independent of t and the left hand side is independent of x , we get

$$\frac{1}{c^2 G(t)} \frac{d^2 G}{dt^2}(t) = \frac{1}{F(x)} \frac{d^2 F}{dx^2}(x) = k$$

for $t > 0$ and $0 < x < L$, and some constant k . We end up with two ordinary differential equations.

$$\frac{d^2 F}{dx^2}(x) - kF(x) = 0 \quad \text{and} \quad \frac{d^2 G}{dt^2}(t) - c^2 k G(t) = 0 . \quad (7.4.9)$$

The first ordinary differential equation in (7.4.9) satisfies two boundary conditions. From $\frac{\partial u}{\partial x}(0, t) = 0$, we get $F'(0)G(t) = 0$. Since we assume that G is non-null, we get $F'(0) = 0$. Similarly, from $\frac{\partial u}{\partial x}(L, t) = 0$, we get $F'(L)G(t) = 0$. Again, since we assume that G is non-null, we get $F'(L) = 0$. The boundary conditions for the first ordinary differential equation in (7.4.9) are $F'(0) = F'(L) = 0$.

Simple Functions: We first consider the boundary value problem

$$\frac{d^2 F}{dx^2}(x) - kF(x) = 0 \quad \text{with} \quad F'(0) = F'(L) = 0 . \quad (7.4.10)$$

The form of the general solution of the ordinary differential equation above is determined by the roots of the characteristic equation $\lambda^2 - k = 0$.

If $k > 0$, the roots of the characteristic equation are $\pm\sqrt{k}$. Since the roots are real, the solution of the ordinary differential equation is of the form $F(x) = Ae^{\sqrt{k}x} + Be^{-\sqrt{k}x}$. However, $F'(0) = 0$ implies that $A\sqrt{k} - B\sqrt{k} = 0$ and $F'(L) = 0$ implies $A\sqrt{k}e^{L\sqrt{k}} - B\sqrt{k}e^{-L\sqrt{k}} = 0$. The only solution of the system

$$\begin{pmatrix} \sqrt{k} & -\sqrt{k} \\ \sqrt{k}e^{L\sqrt{k}} & -\sqrt{k}e^{-L\sqrt{k}} \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

is $A = B = 0$ because

$$\det \begin{pmatrix} \sqrt{k} & -\sqrt{k} \\ \sqrt{k}e^{L\sqrt{k}} & -\sqrt{k}e^{-L\sqrt{k}} \end{pmatrix} = -k(e^{-L\sqrt{k}} - e^{L\sqrt{k}}) \neq 0$$

for $k \neq 0$. Therefore, the null solution is the only solution of the boundary value problem (7.4.10) for $k > 0$.

If $k = 0$, the solution of the ordinary differential equation is $F(x) = B_0x + A_0$. However, $F'(0) = 0$ implies that $B_0 = 0$. $F'(L) = 0$ does not give more information. Thus $B_0 = 0$ and A_0 is free. The boundary value problem (7.4.10) has a non-null solution for $k = k_0 \equiv 0$ which is $F(x) = F_0(x) \equiv A_0$ for all $x \in [0, L]$, where A_0 can be any constant.

If $k < 0$, the roots of the characteristic equation are $\pm i\sqrt{-k}$. Since the roots are complex, the solution of the ordinary differential equation is of the form $F(x) = A \cos(\sqrt{-k}x) + B \sin(\sqrt{-k}x)$. However, $F'(0) = 0$ implies that $B\sqrt{-k} = 0$. So $B = 0$ because $k \neq 0$. Then $F'(L) = 0$ implies that $A\sqrt{-k} \sin(L\sqrt{-k}) = 0$. If we take $A = 0$, we get the null solution. We must therefore have $\sin(L\sqrt{-k}) = 0$ because $k < 0$. This implies that $k = k_n = -(n\pi/L)^2$ for $n > 0$. The boundary value problem (7.4.10) has non-null solutions only for $k = k_n \equiv -(n\pi/L)^2 < 0$ with $n > 0$, and the solution associated to each k_n is of the form

$$F(x) = F_n(x) \equiv A_n \cos\left(\frac{n\pi x}{L}\right).$$

For $k = 0$, the second ordinary differential equation in (7.4.9) becomes $g''(t) = 0$. This equation has the general solution $G(t) = G_0(t) \equiv D_0t + C_0$.

For $k = k_n = -(n\pi/L)^2$, the second ordinary differential equation in (7.4.9) becomes

$$\frac{d^2G}{dt^2}(t) + \left(\frac{cn\pi}{L}\right)^2 G(t) = 0$$

for $n > 0$. For each n , this is a second order ordinary differential equation with the characteristic equation $\lambda^2 + (cn\pi/L)^2 = 0$. The two roots of this equation are the complex numbers $\lambda_{\pm} = \pm(cn\pi/L)i$. The general solution of this ordinary differential equation is therefore

$$G(t) = G_n(t) \equiv C_n \cos\left(\frac{cn\pi t}{L}\right) + D_n \sin\left(\frac{cn\pi t}{L}\right).$$

We have found the simple functions

$$u_0(x, t) \equiv F_0(x)G_0(t) = b_0t + a_0$$

and

$$u_n(x, t) \equiv F_n(x)G_n(t) = \left(a_n \cos\left(\frac{cn\pi t}{L}\right) + b_n \sin\left(\frac{cn\pi t}{L}\right)\right) \cos\left(\frac{n\pi x}{L}\right)$$

for $n > 0$. The constant a_n is the product of the constants A_n and C_n , and b_n is the product of A_n and D_n .

Initial condition: We seek a solution of the form

$$u(x, t) = \sum_{n=0}^{\infty} u_n(x, t) = b_0t + a_0 + \sum_{n=1}^{\infty} \left(a_n \cos\left(\frac{cn\pi t}{L}\right) + b_n \sin\left(\frac{cn\pi t}{L}\right)\right) \cos\left(\frac{n\pi x}{L}\right).$$

From $u(x, 0) = f(x)$, we get

$$f(x) = a_0 + \sum_{n=1}^{\infty} a_n \cos\left(\frac{n\pi x}{L}\right)$$

for $0 < x < L$. This is the Fourier cosine series of f . The coefficients of this series are given by

$$a_0 = \frac{1}{L} \int_0^L f(x) dx \quad \text{and} \quad a_n = \frac{2}{L} \int_0^L f(x) \cos\left(\frac{n\pi x}{L}\right) dx$$

for $n > 0$.

From $\frac{\partial u}{\partial t}(x, 0) = g(x)$, we get

$$g(x) = b_0 + \sum_{n=1}^{\infty} \frac{cn\pi b_n}{L} \cos\left(\frac{n\pi x}{L}\right)$$

for $0 < x < L$. This is the Fourier cosine series of g . The coefficients of this series are given by

$$b_0 = \frac{1}{L} \int_0^L g(x) dx \quad \text{and} \quad b_n = \frac{2}{cn\pi} \int_0^L g(x) \cos\left(\frac{n\pi x}{L}\right) dx$$

for $n > 0$. ♣

Example 7.4.6

Solve the wave equation

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2} \quad , \quad t > 0 \text{ and } 0 < x < 5 \quad , \quad (7.4.11)$$

with the boundary conditions $u(0, t) = 0$ and $\frac{\partial u}{\partial x}(5, t) = 0$ for $t > 0$, and the initial conditions $u(x, 0) = 2 \sin\left(\frac{7\pi x}{10}\right)$ and $\frac{\partial u}{\partial t}(x, 0) = 5 \sin\left(\frac{11\pi x}{10}\right)$ for $0 < x < 5$.

Separation of Variables: If we substitute $u(x, t) = F(x)G(t)$ in (7.4.11), we get

$$F(x) \frac{d^2 G}{dt^2}(t) = \frac{d^2 F}{dx^2}(x) G(t) \quad .$$

Thus, after dividing both sides by $F(x)G(t)$, we get

$$\frac{1}{G(t)} \frac{d^2 G}{dt^2}(t) = \frac{1}{F(x)} \frac{d^2 F}{dx^2}(x) \quad , \quad t > 0 \text{ and } 0 < x < 5 \quad .$$

Since the right hand side is independent of t and the left hand side is independent of x , we get

$$\frac{1}{G(t)} \frac{d^2 G}{dt^2}(t) = \frac{1}{F(x)} \frac{d^2 F}{dx^2}(x) = k$$

for $t > 0$ and $0 < x < 5$, and some constant k . We end up with two ordinary differential equations.

$$\frac{d^2 F}{dx^2}(x) - kF(x) = 0 \quad \text{and} \quad \frac{d^2 G}{dt^2}(t) - kG(t) = 0 \quad . \quad (7.4.12)$$

The first ordinary differential equation in (7.4.12) satisfies two boundary conditions. From $u(0, t) = 0$, we get $F(0)G(t) = 0$. Since we assume that G is non-null, we get $F(0) = 0$. Similarly, from $\frac{\partial u}{\partial x}(5, t) = 0$, we get $\frac{dF}{dx}(5)G(t) = 0$. Again, since we assume that G is non-null, we get $\frac{dF}{dx}(5) = 0$. The boundary conditions for the first ordinary differential equation in (7.4.12) are $F(0) = \frac{dF}{dx}(5) = 0$.

Simple Functions: We first consider the boundary value problem

$$\frac{d^2 F}{dx^2}(x) - kF(x) = 0 \quad \text{with} \quad F(0) = \frac{dF}{dx}(5) = 0. \quad (7.4.13)$$

The form of the general solution of the ordinary differential equation above is determined by the roots of the characteristic equation $\lambda^2 - k = 0$.

If $k > 0$, the roots of the characteristic equation are $\pm\sqrt{k}$. Since the roots are real, the solution of the ordinary differential equation is of the form $F(x) = Ae^{\sqrt{k}x} + Be^{-\sqrt{k}x}$. However, $F(0) = 0$ implies that $A + B = 0$ and $\frac{dF}{dx}(5) = 0$ implies $A\sqrt{k}e^{5\sqrt{k}} - B\sqrt{k}e^{-5\sqrt{k}} = 0$. The only solution of the system

$$\begin{pmatrix} 1 & 1 \\ \sqrt{k}e^{5\sqrt{k}} & -\sqrt{k}e^{-5\sqrt{k}} \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

is $A = B = 0$ because

$$\det \begin{pmatrix} 1 & 1 \\ \sqrt{k}e^{5\sqrt{k}} & -\sqrt{k}e^{-5\sqrt{k}} \end{pmatrix} = -\sqrt{k}(e^{-5\sqrt{k}} + e^{5\sqrt{k}}) \neq 0$$

for $k \neq 0$. Therefore, the null solution is the only solution of the boundary value problem (7.4.13) for $k > 0$.

If $k = 0$, the solution of the ordinary differential equation is $F(x) = Ax + B$. However, $F(0) = 0$ implies that $B = 0$ and $\frac{dF}{dx}(5) = 0$ implies that $A = 0$. Thus $A = B = 0$ and again the null solution is the only solution of the boundary value problem (7.4.13) for $k = 0$.

If $k < 0$, the roots of the characteristic equation are $\pm i\sqrt{-k}$. Since the roots are complex, the solution of the ordinary differential equation is of the form $F(x) = A \cos(\sqrt{-k}x) + B \sin(\sqrt{-k}x)$. Hence, $F(0) = 0$ implies that $A = 0$ and $\frac{dF}{dx}(5) = 0$ implies that $B\sqrt{-k} \cos(5\sqrt{-k}) = 0$. If we take $B = 0$, we get the null solution. We must therefore have $\cos(5\sqrt{-k}) = 0$. This implies that $5\sqrt{-k} = (2n-1)\pi/2$ for $n > 0$. Thus $k = k_n \equiv -((2n-1)\pi/10)^2$ for $n > 0$. The boundary value problem (7.4.13) has non-null solutions only for these values of k , and the solution associated to each k_n is of the form

$$F(x) = F_n(x) \equiv B_n \sin\left(\frac{(2n-1)\pi x}{10}\right).$$

The second ordinary differential equation in (7.4.12) with $k = k_n = -((2n-1)\pi/10)^2$ becomes

$$\frac{d^2 G}{dt^2}(t) + \left(\frac{(2n-1)\pi}{10}\right)^2 G(t) = 0$$

for $n > 0$. For each n , the general solution of this second order ordinary differential equation is

$$G(t) = G_n(t) \equiv C_n \cos\left(\frac{(2n-1)\pi t}{10}\right) + D_n \sin\left(\frac{(2n-1)\pi t}{10}\right) .$$

We have found the simple functions

$$u_n(x, t) \equiv F_n(x)G_n(t) = \left(a_n \cos\left(\frac{(2n-1)\pi t}{10}\right) + b_n \sin\left(\frac{(2n-1)\pi t}{10}\right)\right) \sin\left(\frac{(2n-1)\pi x}{10}\right)$$

for $n > 0$. The constant a_n comes from the product of the constants B_n and C_n , and b_n comes from the product of B_n and D_n .

Initial condition: We seek a solution of the form

$$u(x, t) = \sum_{n=1}^{\infty} u_n(x, t) = \sum_{n=1}^{\infty} \left(a_n \cos\left(\frac{(2n-1)\pi t}{10}\right) + b_n \sin\left(\frac{(2n-1)\pi t}{10}\right)\right) \sin\left(\frac{(2n-1)\pi x}{10}\right) .$$

From $u(x, 0) = 2 \sin\left(\frac{7\pi x}{10}\right)$, we get

$$2 \sin\left(\frac{7\pi x}{10}\right) = \sum_{n=1}^{\infty} a_n \sin\left(\frac{(2n-1)\pi x}{10}\right)$$

for $0 < x < 5$. Thus

$$a_n = \begin{cases} 2 & \text{if } n = 4 \\ 0 & \text{otherwise} \end{cases}$$

From $\frac{\partial u}{\partial t}(x, 0) = 5 \sin\left(\frac{11\pi x}{10}\right)$, we get

$$5 \sin\left(\frac{11\pi x}{10}\right) = \sum_{n=1}^{\infty} \left(\frac{(2n-1)\pi}{10}\right) b_n \sin\left(\frac{(2n-1)\pi x}{10}\right)$$

for $0 < x < 5$. Thus

$$b_n = \begin{cases} \frac{50}{11\pi} & \text{if } n = 6 \\ 0 & \text{otherwise} \end{cases}$$

The solution is

$$u(x, t) = 2 \cos\left(\frac{7\pi t}{10}\right) \sin\left(\frac{7\pi x}{10}\right) + \frac{50}{11\pi} \sin\left(\frac{11\pi t}{10}\right) \sin\left(\frac{11\pi x}{10}\right) .$$

♣

Example 7.4.7

Solve the wave equation

$$\frac{\partial^2 u}{\partial t^2} = 9 \frac{\partial^2 u}{\partial x^2} \quad , \quad t > 0 \text{ and } 0 < x < \pi/2 \quad , \quad (7.4.14)$$

with the boundary conditions $u(0, t) = 0$ and $\frac{\partial u}{\partial x}\left(\frac{\pi}{2}, t\right) = 0$ for $t > 0$, and the initial conditions $u(x, 0) = 3 \sin(7x) + 5 \sin(13x)$ and $\frac{\partial u}{\partial t}(x, 0) = 3 \sin(7x)$ for $0 < x < \pi/2$.

Separation of Variables: If we substitute $u(x, t) = F(x)G(t)$ in (7.4.14), we get

$$F(x) \frac{d^2 G}{dt^2}(t) = 9 \frac{d^2 F}{dx^2}(x) G(t) .$$

After dividing both sides by $9 F(x)G(t)$, we get

$$\frac{1}{9 G(t)} \frac{d^2 G}{dt^2}(t) = \frac{1}{F(x)} \frac{d^2 F}{dx^2}(x)$$

for $t > 0$ and $0 < x < \pi/2$. Since the right hand side is independent of t and the left hand side is independent of x , we get

$$\frac{1}{9 G(t)} \frac{d^2 G}{dt^2}(t) = \frac{1}{F(x)} \frac{d^2 F}{dx^2}(x) = k$$

for $t > 0$ and $0 < x < \pi/2$, and some constant k . We end up with two ordinary differential equations.

$$\frac{d^2 F}{dx^2}(x) - k F(x) = 0 \quad \text{and} \quad \frac{d^2 G}{dt^2}(t) - 9 k G(t) = 0 . \quad (7.4.15)$$

The first ordinary differential equation in (7.4.15) satisfies two boundary conditions. From $u(0, t) = 0$, we get $F(0)G(t) = 0$. Since we assume that G is non-null, we get $F(0) = 0$. Similarly, from $\frac{\partial u}{\partial x}\left(\frac{\pi}{2}, t\right) = 0$, we get $\frac{dF}{dx}\left(\frac{\pi}{2}\right)G(t) = 0$. Again, since we assume that G is non-null, we get $\frac{dF}{dx}\left(\frac{\pi}{2}\right) = 0$. The boundary conditions for the first ordinary differential equation in (7.4.15) are $F(0) = \frac{dF}{dx}\left(\frac{\pi}{2}\right) = 0$.

Simple Functions: We first consider the boundary value problem

$$\frac{d^2 F}{dx^2}(x) - k F(x) = 0 \quad \text{with} \quad F(0) = \frac{dF}{dx}\left(\frac{\pi}{2}\right) = 0 . \quad (7.4.16)$$

The form of the general solution of the ordinary differential equation above is determined by the roots of the characteristic equation $\lambda^2 - k = 0$.

If $k > 0$, the roots of the characteristic equation are $\pm\sqrt{k}$. Since the roots are real, the solution of the ordinary differential equation is of the form $F(x) = Ae^{\sqrt{k}x} + Be^{-\sqrt{k}x}$. However, $F(0) = 0$ implies that $A + B = 0$ and $\frac{dF}{dx}\left(\frac{\pi}{2}\right) = 0$ implies $A\sqrt{k}e^{\pi\sqrt{k}/2} - B\sqrt{k}e^{-\pi\sqrt{k}/2} = 0$. The only solution of the system

$$\begin{pmatrix} 1 & 1 \\ \sqrt{k}e^{\pi\sqrt{k}/2} & -\sqrt{k}e^{-\pi\sqrt{k}/2} \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

is $A = B = 0$ because

$$\det \begin{pmatrix} 1 & 1 \\ \sqrt{k}e^{\pi\sqrt{k}/2} & -\sqrt{k}e^{-\pi\sqrt{k}/2} \end{pmatrix} = -\sqrt{k} \left(e^{-\pi\sqrt{k}/2} + e^{\pi\sqrt{k}/2} \right) \neq 0$$

for $k \neq 0$. Therefore, the null solution is the only solution of the boundary value problem (7.4.16) for $k > 0$.

If $k = 0$, the solution of the ordinary differential equation is $F(x) = Ax + B$. However, $F(0) = 0$ implies that $B = 0$ and $\frac{dF}{dx} \left(\frac{\pi}{2} \right) = 0$ implies that $A = 0$. Thus $A = B = 0$ and again the null solution is the only solution of the boundary value problem (7.4.16) for $k = 0$.

If $k < 0$, the roots of the characteristic equation are $\pm i\sqrt{-k}$. Since the roots are complex, the solution of the ordinary differential equation is of the form $F(x) = A \cos(\sqrt{-k}x) + B \sin(\sqrt{-k}x)$. Moreover, $F(0) = 0$ implies that $A = 0$ and $\frac{dF}{dx} \left(\frac{\pi}{2} \right) = 0$ implies that $B\sqrt{-k} \cos(\pi\sqrt{-k}/2) = 0$. If we take $B = 0$, we get the null solution. We must therefore have $\cos(\pi\sqrt{-k}/2) = 0$. This implies that $\pi\sqrt{-k}/2 = (2n-1)\pi/2$ for $n \in \mathbb{N}$. Thus $k = k_n \equiv -(2n-1)^2$ for $n > 0$. The boundary value problem (7.4.16) has non-null solutions only for these values of k , and the solution associated to each k_n is of the form

$$F(x) = F_n(x) \equiv B_n \sin((2n-1)x) .$$

The second ordinary differential equation in (7.4.15) with $k = k_n = -(2n-1)^2$ becomes

$$\frac{d^2 G}{dt^2}(t) + (3(2n-1))^2 G(t) = 0$$

for $n > 0$. For each n , the general solution of this second order ordinary differential equation is

$$G(t) = G_n(t) \equiv C_n \cos(3(2n-1)t) + D_n \sin(3(2n-1)t) .$$

We have found the simple functions

$$u_n(x, t) \equiv F_n(x)G_n(t) = (a_n \cos(3(2n-1)t) + b_n \sin(3(2n-1)t)) \sin((2n-1)x)$$

for $n > 0$. The constant a_n comes from the product of the constants B_n and C_n , and b_n comes from the product of B_n and D_n .

Initial condition: We seek a solution of the form

$$u(x, t) = \sum_{n=1}^{\infty} u_n(x, t) = \sum_{n=1}^{\infty} (a_n \cos(3(2n-1)t) + b_n \sin(3(2n-1)t)) \sin((2n-1)x) .$$

From $u(x, 0) = 3 \sin(7x) + 5 \sin(13x)$, we get

$$3 \sin(7x) + 5 \sin(13x) = \sum_{n=1}^{\infty} a_n \sin((2n-1)x)$$

for $0 < x < \pi/2$. Thus

$$a_n = \begin{cases} 3 & \text{if } n = 4 \\ 5 & \text{if } n = 7 \\ 0 & \text{otherwise} \end{cases}$$

From $\frac{\partial u}{\partial t}(x, 0) = 3 \sin(7x)$, we get

$$3 \sin(7x) = \sum_{n=1}^{\infty} 3(2n-1)b_n \sin((2n-1)x)$$

for $0 < x < \pi/2$. Thus

$$b_n = \begin{cases} 1/7 & \text{if } n = 4 \\ 0 & \text{otherwise} \end{cases}$$

The solution is

$$u(x, t) = \left(3 \cos(21t) + \frac{1}{7} \sin(21t) \right) \sin(7x) + 5 \cos(39t) \sin(13x) .$$

♣

In the next example, we consider a non-homogeneous wave equation.

Example 7.4.8

Solve the non-homogeneous wave equation

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2} = h(x, t) \equiv 2Lx - x^2 \quad , \quad 0 < x < L \text{ and } t > 0 , \quad (7.4.17)$$

with the boundary conditions $\frac{\partial u}{\partial x}(0, t) = \frac{\partial u}{\partial x}(L, t) = 0$ for $t > 0$, and the initial conditions $u(x, 0) = x^2(3L - 2x)$ and $\frac{\partial u}{\partial t}(x, 0) = \sin^2(\pi x/L)$ for $0 < x < L$.

The solution of the homogeneous wave equation in Example 7.4.5 suggest to seek a solution of the form

$$u(x, t) = T_0(t) + \sum_{n=1}^{\infty} T_n(t) \cos\left(\frac{n\pi x}{L}\right) . \quad (7.4.18)$$

Suppose that the Fourier cosine series of h with respect to the variable x is

$$h(x, t) = h_0(t) + \sum_{n=1}^{\infty} h_n(t) \cos\left(\frac{n\pi x}{L}\right) . \quad (7.4.19)$$

If we substitute (7.4.18) and (7.4.19) in (7.4.17), assuming that we may interchange derivatives and infinite summations, we get

$$T_0''(t) + \sum_{n=1}^{\infty} \left(T_n''(t) + \left(\frac{cn\pi}{L} \right)^2 T_n(t) \right) \cos\left(\frac{n\pi x}{L}\right) = h_0(t) + \sum_{n=1}^{\infty} h_n(t) \cos\left(\frac{n\pi x}{L}\right)$$

for $0 < x < L$. Thus,

$$T_0''(t) = h_0(t) \quad \text{and} \quad T_n''(t) + \left(\frac{cn\pi}{L}\right)^2 T_n(t) = h_n(t)$$

for $t > 0$ and $n > 0$.

We first have to find the Fourier cosine series of $h(x, t) = 2Lx - x^2$ for $0 < x < L$. We have

$$a_0 = \frac{1}{L} \int_0^L (2Lx - x^2) dx = \frac{1}{L} \left(Lx^2 - \frac{x^3}{3} \right) \Big|_0^L = \frac{2L^2}{3}$$

and, using integration by parts, we have

$$\begin{aligned} a_n &= \frac{2}{L} \int_0^L (2Lx - x^2) \cos\left(\frac{n\pi x}{L}\right) dx \\ &= \left(\frac{2}{n\pi} (2Lx - x^2) \sin\left(\frac{n\pi x}{L}\right) \right) \Big|_0^L - \frac{4}{n\pi} \int_0^L (L - x) \sin\left(\frac{n\pi x}{L}\right) dx \\ &= \left(\frac{4L}{n^2\pi^2} (L - x) \cos\left(\frac{n\pi x}{L}\right) \right) \Big|_0^L + \frac{4L}{n^2\pi^2} \int_0^L \cos\left(\frac{n\pi x}{L}\right) dx \\ &= -\frac{4L^2}{n^2\pi^2} + \frac{4L^2}{n^3\pi^3} \sin\left(\frac{n\pi x}{L}\right) \Big|_0^L = -\frac{4L^2}{n^2\pi^2}. \end{aligned}$$

Thus

$$h(x, t) = \frac{2L^2}{3} - \sum_{n=1}^{\infty} \frac{4L^2}{n^2\pi^2} \cos\left(\frac{n\pi x}{L}\right).$$

To determine the value of T_n for $n \geq 0$, we have to solve

$$T_0''(t) = \frac{2L^2}{3} \quad \text{and} \quad T_n''(t) + \left(\frac{cn\pi}{L}\right)^2 T_n(t) = -\frac{4L^2}{n^2\pi^2}$$

for $t > 0$ and $n > 0$. We find

$$T_0(t) = \frac{L^2}{3} t^2 + B_0 t + A_0 \quad \text{and} \quad T_n(t) = A_n \cos\left(\frac{cn\pi t}{L}\right) + B_n \sin\left(\frac{cn\pi t}{L}\right) - \frac{4L^4}{c^2 n^4 \pi^4}$$

for $t > 0$ and $n > 0$.

The solution is

$$u(x, t) = \left(\frac{L^2}{3} t^2 + B_0 t + A_0 \right) + \sum_{n=1}^{\infty} \left(A_n \cos\left(\frac{cn\pi t}{L}\right) + B_n \sin\left(\frac{cn\pi t}{L}\right) - \frac{4L^4}{c^2 n^4 \pi^4} \right) \cos\left(\frac{n\pi x}{L}\right)$$

for $0 < x < L$ and $t > 0$.

From $u(x, 0) = x^2(3L - 2x)$, we get

$$u(x, 0) = A_0 + \sum_{n=1}^{\infty} \left(A_n - \frac{4L^4}{c^2 n^4 \pi^4} \right) \cos\left(\frac{n\pi x}{L}\right) = x^2(3L - 2x).$$

This is the Fourier cosine series of $x^2(3L - 2x)$. Thus,

$$A_0 = \frac{1}{L} \int_0^L x^2(3L - 2x) dx = \frac{1}{L} \left(Lx^3 - \frac{1}{2}x^4 \right) \Big|_0^L = \frac{L^3}{2}$$

and, using integration by parts,

$$\begin{aligned} A_n - \frac{4L^4}{c^2 n^4 \pi^4} &= \frac{2}{L} \int_0^L x^2(3L - 2x) \cos\left(\frac{n\pi x}{L}\right) dx \\ &= \left(\frac{2}{n\pi} x^2(3L - 2x) \sin\left(\frac{n\pi x}{L}\right) \right) \Big|_0^L - \frac{12}{n\pi} \int_0^L (Lx - x^2) \sin\left(\frac{n\pi x}{L}\right) dx \\ &= \left(\frac{12L}{n^2 \pi^2} (Lx - x^2) \cos\left(\frac{n\pi x}{L}\right) \right) \Big|_0^L - \frac{12L}{n^2 \pi^2} \int_0^L (L - 2x) \cos\left(\frac{n\pi x}{L}\right) dx \\ &= - \left(\frac{12L^2}{n^3 \pi^3} (L - 2x) \sin\left(\frac{n\pi x}{L}\right) \right) \Big|_0^L - \frac{24L^2}{n^3 \pi^3} \int_0^L \sin\left(\frac{n\pi x}{L}\right) dx \\ &= \frac{24L^3}{n^4 \pi^4} \cos\left(\frac{n\pi x}{L}\right) \Big|_0^L = \frac{24L^3}{n^4 \pi^4} ((-1)^n - 1) . \end{aligned}$$

From

$$\frac{\partial u}{\partial t}(x, 0) = \sin^2\left(\frac{\pi x}{L}\right) = \frac{1}{2} - \frac{1}{2} \cos\left(\frac{2\pi x}{L}\right) ,$$

we get

$$B_0 + \sum_{n=1}^{\infty} B_n \frac{cn\pi}{L} \cos\left(\frac{n\pi x}{L}\right) = \frac{1}{2} - \frac{1}{2} \cos\left(\frac{2\pi x}{L}\right) .$$

This is the Fourier cosine series of $\frac{1}{2} - \frac{1}{2} \cos\left(\frac{2\pi x}{L}\right)$. By uniqueness of the Fourier series, we have

$$B_n = \begin{cases} \frac{1}{2} & \text{if } n = 0 \\ -\frac{1}{2} & \text{if } n = 2 \\ 0 & \text{otherwise} \end{cases}$$

The solution of the non-homogeneous wave equation is

$$\begin{aligned} u(x, t) &= \left(\frac{L^2}{3} t^2 + \frac{t}{2} + \frac{L^3}{2} \right) + \sum_{\substack{n=1 \\ n \neq 2}}^{\infty} \left(\frac{24L^3}{n^4 \pi^4} ((-1)^n - 1) - \frac{4L^4}{c^2 n^4 \pi^4} \right) \cos\left(\frac{n\pi x}{L}\right) \\ &\quad - \left(\frac{1}{2} \sin\left(\frac{2c\pi t}{L}\right) + \frac{L^4}{4c^2 \pi^4} \right) \cos\left(\frac{2\pi x}{L}\right) . \end{aligned}$$

We could show that u and its derivatives of order up to 2 converges absolutely and locally uniformly for $0 < x < L$ and $t > 0$. ♣

Remark 7.4.9

Using separation of variables, we find that the solutions of the wave equation in (7.4.2) are linear combinations of simple functions of the form

$$u(x, t) = (A \cos(ckt) + B \sin(ckt))(C \cos(kx) + D \sin(kx)) .$$



7.5 Exercises

Suggested exercises:

- In [16]: all the numbers in Section 2.4.
- In [22]: all the numbers in Sections 3.1.
- In [24]: all the numbers in Sections 4.6; numbers 6.20, 6.24 and 6.26 in Section 6.7.
- In [38]: all numbers in Sections 2.1, 3.2 and 3.4; numbers 1, 2, 4 and 5 in Section 4.1; number 2 in Section 4.2; number 13 in Section 4.3; numbers 4 to 6, 9 and 13 in Section 5.6.

Chapter 8

Wave Equation in $\mathbb{R}^n$

8.1 The Cauchy Problem

Consider the Cauchy problem provided by the **wave equation**

$$L(\mathbf{x}, t, D) = \frac{\partial^2 u}{\partial t^2} - c^2 \Delta u = 0 \quad , \quad (\mathbf{x}, t) \in \mathbb{R}^n \times \mathbb{R} \quad , \quad (8.1.1)$$

with the initial condition

$$u(\mathbf{x}, 0) = f(\mathbf{x}) \quad \text{and} \quad \frac{\partial u}{\partial t}(\mathbf{x}, 0) = g(\mathbf{x}) \quad , \quad \mathbf{x} \in \mathbb{R}^n \quad . \quad (8.1.2)$$

This is an hyperbolic equation because the principal symbol of $L(\mathbf{x}, t, D)$ is

$$Q((\mathbf{x}, t), (\boldsymbol{\xi}, r)) = r^2 - c^2 \sum_{i=1}^n \xi_i^2 = \begin{pmatrix} \xi_1 & \dots & \xi_n & r \end{pmatrix} \begin{pmatrix} -c^2 & 0 & \dots & 0 & 0 \\ 0 & -c^2 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & -c^2 & 0 \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \\ r \end{pmatrix} \quad ,$$

where the matrix has all eigenvalues but one equal to $-c^2 < 0$. The other eigenvalue is $1 > 0$.

Definition 8.1.1

An hyper-surface $S \subset \mathbb{R}^n \times \mathbb{R}$ is called **space-like** if its normal vector

$$\nu(\mathbf{x}, t) = (v(\mathbf{x}, t), w(\mathbf{x}, t)) \quad ,$$

where $v : S \rightarrow \mathbb{R}^n$ and $w : S \rightarrow \mathbb{R}$, satisfies $|w(\mathbf{x}, t)| > \|v(\mathbf{x}, t)\|$ for all $(\mathbf{x}, t) \in S$.

As in Figure 8.1, if the surface is $S = \{(\mathbf{x}, t) : t = \phi(\mathbf{x}) \text{ and } \mathbf{x} \in \mathbb{R}^n\}$ for some smooth function $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$, then S is space-like if the normal to S at a point $(\mathbf{x}_0, t_0) \in S$ is inside

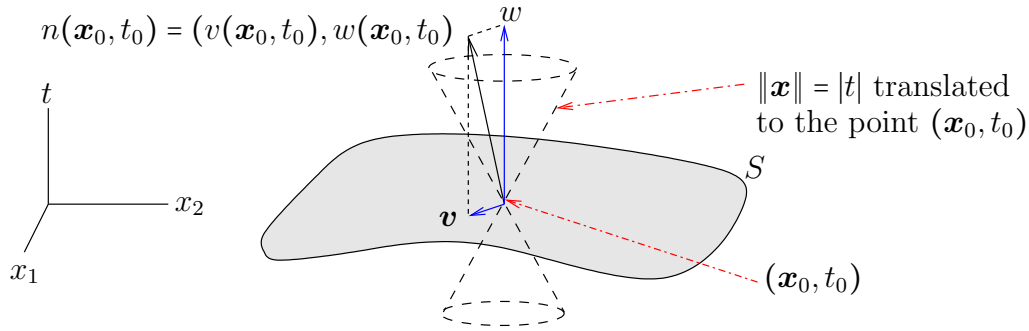


Figure 8.1: Graphical interpretation of space-like surface for the wave equation.

the cone $\|\mathbf{x}\| = |t|$. In particular, if S is the hyperplane $t = 0$, then S is space-like because the outward normal to S is $\nu(\mathbf{x}, 0) = (\mathbf{0}, -1)$ for all $(\mathbf{x}, 0)$.

The condition $\|v(\mathbf{x}, t)\| < |w(\mathbf{x}, t)|$ is necessary to get a smooth change of coordinates

$$\begin{aligned} \mathbb{R}^n \times \mathbb{R} &\rightarrow \mathbb{R}^n \times \mathbb{R} \\ (\mathbf{x}, t) &\mapsto (\mathbf{x}, t - \phi(\mathbf{x})) \end{aligned}$$

that preserve the (strict) hyperbolicity and reduce the hyper-surface S to the surface $t = 0$.

Remark 8.1.2

The hyperplane $t = 0$ is not a characteristic surface because $Q((\mathbf{x}, 0), \nu(\mathbf{x}, 0)) = 1$ for all $\mathbf{x} \in \mathbb{R}^n$. ♠

Since $L(\mathbf{x}, t, D)$ commutes with the application

$$\begin{aligned} \mathbb{R}^n \times \mathbb{R} &\rightarrow \mathbb{R}^n \times \mathbb{R} \\ (\mathbf{x}, t) &\mapsto (\mathbf{x}, -t) \end{aligned}$$

we only have to consider the wave equation for $t \geq 0$.

8.2 Uniqueness of the Solution

We start by proving the uniqueness of the solution of the wave equation. The proof is interesting because it introduces the concept of energy functions.

Theorem 8.2.1

Let

$$R = \{(\mathbf{x}, t) \in \mathbb{R}^n \times \mathbb{R} : 0 \leq t \leq T\}$$

and

$$B = \{(\mathbf{x}, 0) \in \mathbb{R}^n \times \mathbb{R} : \|\mathbf{x} - \mathbf{x}_0\| < c|t_0|\}$$

for some point $(\mathbf{x}_0, t_0)$ with $t_0 < T$ (Figure 8.2). Suppose that $u \in C^2(R)$ satisfies $L(\mathbf{x}, t, D)u = 0$ on R , $u = 0$ on B and $\frac{\partial u}{\partial t} = 0$ on B . Then $u = 0$ on

$$\Omega = \{(\mathbf{x}, t) \in \mathbb{R}^n \times \mathbb{R} : \|\mathbf{x} - \mathbf{x}_0\| < c|t - t_0|, 0 \leq t \leq t_0\}.$$

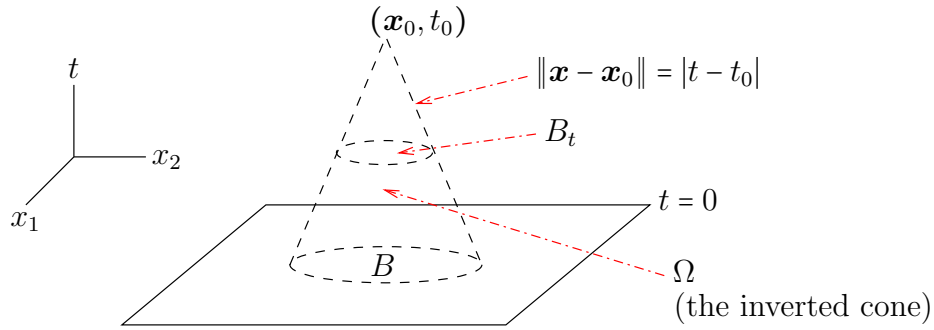


Figure 8.2: This figure is used in the proof of Theorem 8.2.1 which implies the uniqueness of the solution for the wave equation.

Proof.

With the substitution $t = s/\sqrt{c}$, we get $\frac{\partial^2 u}{\partial s^2} - \Delta_{\mathbf{x}} u = 0$. Thus, we may assume that $L(\mathbf{x}, t, D) = \frac{\partial^2 u}{\partial t^2} - \Delta_{\mathbf{x}} u = 0$. In other words, we may assume that $c = 1$. In this proof, we refer to Figure 8.2. Let

$$B_t = \{\mathbf{x} : \|\mathbf{x} - \mathbf{x}_0\| \leq |t - t_0|\}$$

for $0 \leq t \leq t_0$, and

$$E(t) = \frac{1}{2} \int_{B_t} \left(\left(\frac{\partial u}{\partial t} \right)^2 + \|\nabla_{\mathbf{x}} u\|^2 \right) d\mathbf{x} = \frac{1}{2} \int_{B_t} \left(\left(\frac{\partial u}{\partial t} \right)^2 + \sum_{j=1}^n \left(\frac{\partial u}{\partial x_j} \right)^2 \right) d\mathbf{x}.$$

$E(t)$ is the **energy** of the wave in B_t at time t . We have

$$\frac{dE}{dt} = \int_{B_t} \left(\frac{\partial^2 u}{\partial t^2} \frac{\partial u}{\partial t} + \sum_{j=1}^n \frac{\partial^2 u}{\partial x_j \partial t} \frac{\partial u}{\partial x_j} \right) d\mathbf{x} - \frac{1}{2} \int_{\partial B_t} \left(\left(\frac{\partial u}{\partial t} \right)^2 + \|\nabla_{\mathbf{x}} u\|^2 \right) dS_{\mathbf{x}},$$

where we have used Remark 8.2.2 below with $F(\mathbf{x}, t) = \left(\frac{\partial u}{\partial t} \right)^2 + \sum_{j=1}^n \left(\frac{\partial u}{\partial x_j} \right)^2$.

Since

$$\frac{\partial}{\partial x_j} \left(\frac{\partial u}{\partial x_j} \frac{\partial u}{\partial t} \right) = \frac{\partial^2 u}{\partial x_j \partial t} \frac{\partial u}{\partial x_j} + \frac{\partial^2 u}{\partial x_j^2} \frac{\partial u}{\partial t} ,$$

we get

$$\int_{B_t} \left(\frac{\partial^2 u}{\partial t^2} \frac{\partial u}{\partial t} + \sum_{j=1}^n \frac{\partial^2 u}{\partial x_j \partial t} \frac{\partial u}{\partial x_j} \right) d\mathbf{x} = \int_{B_t} \left(\sum_{j=1}^n \frac{\partial}{\partial x_j} \left(\frac{\partial u}{\partial x_j} \frac{\partial u}{\partial t} \right) - \sum_{j=1}^n \frac{\partial^2 u}{\partial x_j^2} \frac{\partial u}{\partial t} + \frac{\partial^2 u}{\partial t^2} \frac{\partial u}{\partial t} \right) d\mathbf{x} .$$

Using the divergence theorem in the hyperplane t constant, we get

$$\int_{B_t} \sum_{j=1}^n \frac{\partial}{\partial x_j} \left(\frac{\partial u}{\partial x_j} \frac{\partial u}{\partial t} \right) d\mathbf{x} = \int_{B_t} \operatorname{div}_{\mathbf{x}} \left(\frac{\partial u}{\partial t} \nabla_{\mathbf{x}} u \right) d\mathbf{x} = \int_{\partial B_t} \frac{\partial u}{\partial t} (\nabla_{\mathbf{x}} u \cdot \nu) dS_{\mathbf{x}} ,$$

where $\nu(\mathbf{x})$ is the unit normal to the surface ∂B_t at $\mathbf{x} \in \partial B_t$ in the hyperplane t constant which is pointing outside B_t . Moreover,

$$\int_{B_t} \left(- \sum_{j=1}^n \frac{\partial^2 u}{\partial x_j^2} \frac{\partial u}{\partial t} + \frac{\partial^2 u}{\partial t^2} \frac{\partial u}{\partial t} \right) d\mathbf{x} = \int_{B_t} \left(\frac{\partial^2 u}{\partial t^2} - \sum_{j=1}^n \frac{\partial^2 u}{\partial x_j^2} \right) \frac{\partial u}{\partial t} d\mathbf{x} = 0$$

because $L(\mathbf{x}, t, D)u = \frac{\partial^2 u}{\partial t^2} - \Delta_{\mathbf{x}} u = 0$. Thus

$$\int_{B_t} \left(\frac{\partial^2 u}{\partial t^2} \frac{\partial u}{\partial t} + \sum_{j=1}^n \frac{\partial^2 u}{\partial x_j \partial t} \frac{\partial u}{\partial x_j} \right) d\mathbf{x} = \int_{\partial B_t} \frac{\partial u}{\partial t} (\nabla_{\mathbf{x}} u \cdot \nu) dS_{\mathbf{x}} .$$

We end up with

$$\frac{dE}{dt} = \int_{\partial B_t} \frac{\partial u}{\partial t} (\nabla_{\mathbf{x}} u \cdot \nu) dS_{\mathbf{x}} - \frac{1}{2} \int_{\partial B_t} \left(\left(\frac{\partial u}{\partial t} \right)^2 + \|\nabla_{\mathbf{x}} u\|^2 \right) dS_{\mathbf{x}}$$

for $0 < t < t_0$. From Schwarz inequality, we get

$$\left| \frac{\partial u}{\partial t} (\nabla_{\mathbf{x}} u \cdot \nu) \right| \leq \left| \frac{\partial u}{\partial t} \right| \|\nabla_{\mathbf{x}} u\| \|\nu\| = \left| \frac{\partial u}{\partial t} \right| \|\nabla_{\mathbf{x}} u\| \leq \frac{1}{2} \left(\left(\frac{\partial u}{\partial t} \right)^2 + \|\nabla_{\mathbf{x}} u\|^2 \right) ,$$

where we have use the inequality $2ab \leq a^2 + b^2$ for the last inequality. Therefore, $\frac{dE}{dt} \leq 0$ for $0 < t < t_0$.

Since $u(\mathbf{x}, 0) = 0$ for all $\mathbf{x} \in B$ by hypothesis, we get $\frac{\partial u}{\partial x_j}(\mathbf{x}, 0) = 0$ for all $\mathbf{x} \in B$. Moreover, since $\frac{\partial u}{\partial t}(\mathbf{x}, 0) = 0$ for all $\mathbf{x} \in B$ by hypothesis, we get $E(0) = 0$.

Finally, since $E(t) \geq 0$ for $0 \leq t \leq t_0$ by definition, $E(0) = 0$, and $\frac{dE}{dt} \leq 0$ for $0 < t < t_0$, we get that $E(t) = 0$ for $0 \leq t \leq t_0$. This implies that $\nabla_{\mathbf{x}, t} u = 0$ on Ω . Therefore u is constant on Ω . Since $u(\mathbf{x}, 0) = 0$ for all $\mathbf{x} \in B$, we get that $u = 0$ on Ω . ■

Remark 8.2.2

Suppose that $F : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$ is a function of class C^1 and

$$G(t) = \int_{B_t} F(\mathbf{x}, t) d\mathbf{x} ,$$

where B_t is defined in the proof of the previous proposition. Then

$$G(t) = \int_0^{t_0-t} \int_{\|\mathbf{x}\|=1} F(\mathbf{x}_0 + r\mathbf{x}, t) dS_{\mathbf{x}} dr$$

and

$$\begin{aligned} \frac{dG}{dt}(t) &= \int_0^{t_0-t} \int_{\|\mathbf{x}\|=1} \frac{\partial F}{\partial t}(\mathbf{x}_0 + r\mathbf{x}, t) dS_{\mathbf{x}} dr - \int_{\|\mathbf{x}\|=1} F(\mathbf{x}_0 + (t_0-t)\mathbf{x}, t) dS_{\mathbf{x}} \\ &= \int_{B_t} \frac{\partial F}{\partial t}(\mathbf{x}, t) d\mathbf{x} - \int_{\partial B_t} F(\mathbf{x}, t) dS_{\mathbf{x}} . \end{aligned}$$

♠

Corollary 8.2.3

If u and v in $C^2(\mathbb{R}^n \times \mathbb{R})$ are two solutions of (8.1.1) satisfying the Cauchy (or initial) conditions (8.1.2), then $u = v$.

Proof.

Apply the previous proposition to $u - v$. ■

8.3 Laplace Operator

Before proving the existence of solutions for the Cauchy problem provided by the wave equation, we need to study the **Laplace operator** better known as the Laplacian.

Theorem 8.3.1

A linear differential operator $L(\mathbf{x}, D)$ commutes with Euclidean translations and rotations if and only if $L(\mathbf{x}, D)$ is a polynomial in Δ .

Proof.

$\Rightarrow$) Suppose that

$$L(\mathbf{x}, D) = \sum_{|\alpha| \leq k} a_{\alpha}(\mathbf{x}) D^{\alpha}$$

for some positive integer k , where $D^{\alpha} = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \dots \partial x_n^{\alpha_n}}$ for all multi-indices α as usual.

$L(\mathbf{x}, D)$ commutes with translations means that $(L(\mathbf{x}, D)u)(\mathbf{x} + \mathbf{y}) = L(\mathbf{x}, D)(u(\mathbf{x} + \mathbf{y}))$ for all $u \in C^k(\mathbb{R}^n)$, and for all $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$. Namely,

$$\sum_{|\alpha| \leq k} a_\alpha(\mathbf{x} + \mathbf{y}) D^\alpha u(\mathbf{x} + \mathbf{y}) = \sum_{|\alpha| \leq k} a_\alpha(\mathbf{x}) D^\alpha u(\mathbf{x} + \mathbf{y}) \quad (8.3.1)$$

for all $u \in C^k(\mathbb{R}^n)$, and for all $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$. If $u(\mathbf{x}) = \mathbf{x}^\alpha$ for some multi-indices α such that $|\alpha| \leq k$, then $D^\beta u(\mathbf{x}) = \frac{\alpha!}{(\alpha - \beta)!} \mathbf{x}^{\alpha - \beta}$ for $\beta \leq \alpha$ and 0 otherwise. In particular, $D^\alpha u(\mathbf{x}) = \alpha!$. With $\mathbf{y} = -\mathbf{x}$, (8.3.1) becomes

$$\alpha! a_\alpha(\mathbf{0}) = \alpha! a_\alpha(\mathbf{x}) .$$

Since $\mathbf{x}$ is arbitrary, a_α is a constant. Since the multi-index α is arbitrary, $L(D)$ has constant coefficients. Since $L(D)$ obviously commutes with translations when the coefficients are constant, we have shown that $L(D)$ commutes with translations if and only if the coefficients a_α are constant.

We have for all $u \in C^k(\mathbb{R}^n)$ that

$$(L(\mathbf{x}, D)u)^\wedge(\mathbf{y}) = p(\mathbf{y}) \hat{u}(\mathbf{y}) ,$$

where

$$p(\mathbf{y}) = \sum_{|\alpha| \leq k} a_\alpha (i\mathbf{y})^\alpha .$$

For R a rotation, we have

$$\begin{aligned} \hat{u}(R\mathbf{y}) &= \int_{\mathbb{R}^n} e^{-i(R\mathbf{y}) \cdot \mathbf{x}} u(\mathbf{x}) d\mathbf{x} = \int_{\mathbb{R}^n} e^{-i\mathbf{y} \cdot (R^\top \mathbf{x})} u(\mathbf{x}) d\mathbf{x} \\ &= \int_{\mathbb{R}^n} e^{-i\mathbf{y} \cdot \mathbf{z}} u(R\mathbf{z}) |\det R| d\mathbf{z} = \int_{\mathbb{R}^n} e^{-i\mathbf{y} \cdot \mathbf{z}} u(R\mathbf{z}) d\mathbf{z} = (u \circ R)^\wedge(\mathbf{y}) , \end{aligned}$$

where we have use the fact that $|\det R| = 1$ for rotations. Hence, for R a rotation,

$$(L(\mathbf{x}, D)(u \circ R))^\wedge(\mathbf{y}) = p(\mathbf{y})(u \circ R)^\wedge(\mathbf{y}) = p(\mathbf{y}) \hat{u}(R\mathbf{y})$$

and

$$((L(\mathbf{x}, D)(u)) \circ R)^\wedge(\mathbf{y}) = (L(\mathbf{x}, D)u)^\wedge(R\mathbf{y}) = p(R\mathbf{y}) \hat{u}(R\mathbf{y}) .$$

Therefore, $L(\mathbf{x}, D)$ commutes with the rotations if and only if p is radial; namely, $p(\mathbf{y}) = p(R\mathbf{y})$ for all $\mathbf{y}$ and all rotation R or, in other words, p depends only of the radius $r = \|\mathbf{y}\|_2$.

If $p(\mathbf{y}) = \sum_{|\alpha| \leq k} a_\alpha (i\mathbf{y})^\alpha$ is radial, then $p_j(\mathbf{y}) = \sum_{|\alpha|=j} a_\alpha (i\mathbf{y})^\alpha$ is radial for $0 \leq j \leq k$ by homogeneity because rotations preserve the degree of p_j . Thus p_j is an homogeneous polynomial of degree j in $\|\mathbf{y}\|_2$; namely, $p_j(\mathbf{y}) = b_j (i\|\mathbf{y}\|_2)^j$ for some constant b_j . However, since p is a polynomial, $b_j = 0$ for j odd.

We have found that

$$p(\mathbf{y}) = \sum_{j \leq k/2} b_{2j} (i\|\mathbf{y}\|_2)^{2j} = \sum_{j \leq k/2} b_{2j} ((iy_1)^2 + (iy_2)^2 + \dots + (iy_n)^2)^j .$$

Hence, k must be even and

$$L(\mathbf{x}, D) = \sum_{j \leq k/2} b_{2j} \Delta^j .$$

$\Leftarrow$) If $L(\mathbf{x}, D) = \sum_{j=0}^k a_j \Delta^j$, then $L(\mathbf{x}, D)$ commutes with translations because it has constant coefficients. Since $p(\mathbf{y}) = \sum_{j=0}^k a_j (i\|\mathbf{y}\|_2^2)^j$ is radial, then $L(\mathbf{x}, D)$ commutes with rotations. ■

Proposition 8.3.2

Consider $\phi \in C^2(\mathbb{R})$. If $f(\mathbf{x}) = \phi(\|\mathbf{x}\|_2)$ for $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$, then

$$\Delta f(\mathbf{x}) = \phi''(r) + \frac{n-1}{r} \phi'(r)$$

for $r = \|\mathbf{x}\|_2$.

Proof.

Let $r = \|\mathbf{x}\|_2$. Since $\frac{\partial r}{\partial x_j} = \frac{x_j}{r}$ for $1 \leq j \leq n$, we get

$$\begin{aligned} \Delta_{\mathbf{x}} f(\mathbf{x}) &= \sum_{j=1}^n \frac{\partial}{\partial x_j} \left(\frac{\partial}{\partial x_j} \phi(r) \right) = \sum_{j=1}^n \frac{\partial}{\partial x_j} \left(\phi'(r) \frac{\partial r}{\partial x_j} \right) = \sum_{j=1}^n \frac{\partial}{\partial x_j} \left(\phi'(r) \frac{x_j}{r} \right) \\ &= \sum_{j=1}^n \left(\left(\frac{\partial}{\partial x_j} \phi'(r) \right) \frac{x_j}{r} + \phi'(r) \frac{\partial}{\partial x_j} \left(\frac{x_j}{r} \right) \right) = \sum_{j=1}^n \left(\phi''(r) \left(\frac{x_j}{r} \right)^2 + \frac{1}{r} \phi'(r) - \frac{x_j^2}{r^3} \phi'(r) \right) \\ &= \phi''(r) \left(\frac{1}{r^2} \sum_{j=1}^n x_j^2 \right) + \frac{n}{r} \phi'(r) - \left(\frac{1}{r^3} \sum_{j=1}^n x_j^2 \right) \phi'(r) = \phi''(r) + \left(\frac{n}{r} - \frac{1}{r} \right) \phi'(r) . \quad \blacksquare \end{aligned}$$

Corollary 8.3.3

Suppose that $f(\mathbf{x}) = \phi(\|\mathbf{x}\|)$ satisfies $\Delta f = 0$ on $\mathbb{R}^n \setminus \{\mathbf{0}\}$. Then

$$\phi(r) = \begin{cases} a + br^{2-n} & \text{if } n \neq 2 \\ a + b \ln(r) & \text{if } n = 2 \end{cases}$$

for $r = \|\mathbf{x}\|$.

Proof.

From $\Delta f = 0$, we get $\phi''(r) + \frac{n-1}{r}\phi'(r) = 0$. With $w(r) = \phi'(r)$, this last equation becomes the separable first order ordinary differential equation $w'(r) + \frac{n-1}{r}w(r) = 0$ whose solution is $w(r) = Cr^{1-n}$ if $n \neq 1$, where C is an arbitrary constant.

For $n \neq 2$, we get $\phi(r) = br^{2-n} + a$, where $b = C/(2-n)$ and a is a constant. For $n = 2$, we get $\phi(r) = b \ln(r) + a$, where $b = C$ and a is a constant. ■

8.4 Existence of the Solution

In the next subsections, we turn our attention to the existence of a solution for the Cauchy problem (8.1.1) with the initial conditions (8.1.2). We follow the approach presented in [10].

8.4.1 One Dimensional Space

We have seen in Section 7.1 that the solution of the Cauchy problem (8.1.1) with the initial conditions (8.1.2) is given by

$$u(x, t) = \frac{1}{2} (f(x - ct) + f(x + ct)) + \frac{1}{2c} \int_{x-ct}^{x+ct} g(s) ds$$

if $f \in C^2(\mathbb{R})$ and $g \in C^1(\mathbb{R})$. This solution can be written as

$$u(x, t) = \frac{1}{2} (f_{ct}(x) + f_{-ct}(x)) + \frac{1}{2c} (g * \mathcal{X}_{[-ct, ct]})(x), \quad (8.4.1)$$

where $\mathcal{X}_{[a, b]}$ is the characteristic function defined by

$$\mathcal{X}_{[a, b]}(x) = \begin{cases} 1 & \text{if } x \in [a, b] \\ 0 & \text{otherwise} \end{cases}$$

If f and g are locally integrable functions, then (8.4.1) is a strong solution of the Cauchy problem.

8.4.2 Higher Odd Dimensional Spaces

The **spherical mean** of a continuous function $h : \mathbb{R}^n \rightarrow \mathbb{R}$ is defined by the surface integral

$$M_h(\mathbf{x}, r) = \frac{1}{\omega_n} \int_{\|\mathbf{y}\|=1} h(\mathbf{x} + r\mathbf{y}) dS_{\mathbf{y}}, \quad (8.4.2)$$

where ω_n is the area of a sphere of radius 1 in $\mathbb{R}^n$. For now, we do not need to assume that n is odd.

The spherical mean of h has the following properties.

1. $M_h(\mathbf{x}, r) = M_h(\mathbf{x}, -r)$ for $r \in \mathbb{R}$ and $\mathbf{x} \in \mathbb{R}^n$.
2. $M_h \in C^k(\mathbb{R}^n \times \mathbb{R})$ if $h \in C^k(\mathbb{R}^n)$.
3. $M_h(\mathbf{x}, 0) = h(\mathbf{x})$ for $\mathbf{x} \in \mathbb{R}^n$ if $h \in C(\mathbb{R}^n)$.

For $r \neq 0$, the change of variables $\mathbf{z} = \mathbf{x} + r\mathbf{y}$ transform the spherical mean formula into

$$M_h(\mathbf{x}, r) = \frac{1}{r^{n-1}\omega_n} \int_{\|\mathbf{x}-\mathbf{z}\|=r} h(\mathbf{z}) dS. \quad (8.4.3)$$

In $\mathbb{R}^2$, the surface integrals above are computed using polar coordinates. For (8.4.2), we use $\mathbf{y} = (\cos(\theta), \sin(\theta))$ with $0 \leq \theta < 2\pi$. Thus $dS_{\mathbf{y}} = d\theta$. For (8.4.3), we use $\mathbf{z} = \mathbf{x} + r(\cos(\theta), \sin(\theta))$ with $0 \leq \theta < 2\pi$. Thus $dS = r d\theta$.

In $\mathbb{R}^3$, the surface integrals above are computed using spherical coordinates. For (8.4.2), we use $\mathbf{y} = (\cos(\theta)\sin(\phi), \sin(\theta)\sin(\phi), \cos(\phi))$ with $0 \leq \theta < 2\pi$ and $0 \leq \phi < \pi$. Thus $dS_{\mathbf{y}} = \sin(\phi) d\phi d\theta$. For (8.4.3), we use $\mathbf{z} = \mathbf{x} + r(\cos(\theta)\sin(\phi), \sin(\theta)\sin(\phi), \cos(\phi))$ with $0 \leq \theta < 2\pi$ and $0 \leq \phi < \pi$. Thus $dS = r^2 \sin(\phi) d\phi d\theta$.

Integrating $e^{-\pi\|\mathbf{x}\|_2^2}$ over $\mathbb{R}^n$ using spherical coordinates¹, the reader can verify that $\omega_n = 2\pi^{n/2}/\Gamma(n/2)$ for $n > 0$. It is a nice calculus exercise whose solution can be found in [10]). In particular, $\omega_2 = 2\pi$ and $\omega_3 = 4\pi$.

Remark 8.4.1

This remark is for those familiar with the theory to measure and integration. For any $\mathbf{y} \in \mathbb{R}^n$ such that $\|\mathbf{y}\| = r > 0$, we have that

$$M_h(\mathbf{x}, r) = \frac{1}{\omega_n} \int_{T \in O(n)} h(\mathbf{x} + T(\mathbf{y})) d\mu(T)$$

for $\mathbf{x} \in \mathbb{R}^n$, where $\mu(T)$ is the Haar measure on the group $O(n)$ of orthogonal transformations on $\mathbb{R}^n$.

Hence, if $h \in C^2(\mathbb{R}^n)$,

$$\begin{aligned} \Delta_{\mathbf{x}} M_h(\mathbf{x}, r) &= \frac{1}{\omega_n} \int_{T \in O(n)} \Delta_{\mathbf{x}} h(\mathbf{x} + T(\mathbf{y})) d\mu(T) \\ &= \frac{1}{\omega_n} \int_{T \in O(n)} \Delta_{\mathbf{y}} h(\mathbf{x} + T(\mathbf{y})) d\mu(T) = \Delta_{\mathbf{y}} M_h(\mathbf{x}, r) \end{aligned} \quad (8.4.4)$$

with $r = \|\mathbf{y}\|$ because Δ is invariant under orthogonal transformations (rotations, reflections and their combinations) and translations; so

$$\Delta_{\mathbf{x}} h(\mathbf{x} + T(\mathbf{y})) = \Delta h(\mathbf{x} + T(\mathbf{y})) = \Delta_{\mathbf{y}} h(\mathbf{x} + T(\mathbf{y})).$$

Suppose that $f \in C^2(\mathbb{R}^n)$ satisfies $f(\mathbf{x}) = g(r)$ with $r = \|\mathbf{x}\|$ for all $\mathbf{x} \in \mathbb{R}^n$. Using the spherical coordinates $\mathbf{x} = r\mathbf{u}$ with $\|\mathbf{u}\| = 1$ and $r > 0$, it follows from Proposition 8.3.2 that

$$\Delta f = \sum_{i=1}^n \frac{\partial^2 f}{\partial x_i^2} = \frac{\partial^2 g}{\partial r^2} + \left(\frac{n-1}{r} \right) \frac{\partial g}{\partial r}.$$

¹We call spherical coordinates in $\mathbb{R}^n$ any coordinates of the form $\mathbf{x} = r\mathbf{y}$ with $\|\mathbf{y}\| = 1$.

If we apply this result to $g(r) = M_h(\mathbf{x}, r)$, we get from (8.4.4) that

$$\Delta_{\mathbf{x}} M_h(\mathbf{x}, r) = \Delta_{\mathbf{y}} M_h(\mathbf{x}, r) = \left(\frac{\partial^2}{\partial r^2} + \left(\frac{n-1}{r} \right) \frac{\partial}{\partial r} \right) M_h(\mathbf{x}, r)$$

for $r = \|\mathbf{y}\|$. This result can be proved using elementary calculus as it is shown in the next proposition. $\spadesuit$

Proposition 8.4.2

If $h \in C^2(\mathbb{R}^n)$, then

$$\Delta_{\mathbf{x}} M_h(\mathbf{x}, r) = \left(\frac{\partial^2}{\partial r^2} + \left(\frac{n-1}{r} \right) \frac{\partial}{\partial r} \right) M_h(\mathbf{x}, r) .$$

Proof.

Since M_h is even with respect to r , we may assume that $r > 0$.

We have

$$\frac{\partial M_h}{\partial r}(\mathbf{x}, r) = \frac{1}{\omega_n} \int_{\|\mathbf{y}\|=1} \nabla h(\mathbf{x} + r\mathbf{y}) \cdot \mathbf{y} \, dS_{\mathbf{y}} .$$

At the point $\mathbf{y}$ on the sphere of radius 1 centred at the origin, the vector $\mathbf{y}$ is a unit vector perpendicular to the sphere and pointing outside of the sphere. We may therefore use the divergence theorem (in $\mathbf{y}$ coordinates) to conclude that

$$\begin{aligned} \frac{\partial M_h}{\partial r}(\mathbf{x}, r) &= \frac{1}{\omega_n} \int_{\|\mathbf{y}\| \leq 1} r \Delta h(\mathbf{x} + r\mathbf{y}) \, d\mathbf{y} = \frac{1}{r^{n-1} \omega_n} \int_{\|\mathbf{y}\| \leq r} \Delta h(\mathbf{x} + \mathbf{y}) \, d\mathbf{y} \\ &= \frac{1}{r^{n-1} \omega_n} \int_0^r \left(\int_{\|\mathbf{y}\|=1} \Delta h(\mathbf{x} + \rho\mathbf{y}) \rho^{n-1} \, dS_{\mathbf{y}} \right) d\rho \end{aligned}$$

because $\operatorname{div}_{\mathbf{y}}(\nabla h(\mathbf{x} + r\mathbf{y})) = r \Delta h(\mathbf{x} + r\mathbf{y})$. If we derive both sides of the equality

$$r^{n-1} \frac{\partial M_h}{\partial r}(\mathbf{x}, r) = \frac{1}{\omega_n} \int_0^r \left(\int_{\|\mathbf{y}\|=1} \Delta h(\mathbf{x} + \rho\mathbf{y}) \rho^{n-1} \, dS_{\mathbf{y}} \right) d\rho$$

with respect to r , we get

$$(n-1)r^{n-2} \frac{\partial M_h}{\partial r}(\mathbf{x}, r) + r^{n-1} \frac{\partial^2 M_h}{\partial r^2}(\mathbf{x}, r) = \frac{1}{\omega_n} \int_{\|\mathbf{y}\|=1} \Delta h(\mathbf{x} + r\mathbf{y}) r^{n-1} \, dS_{\mathbf{y}} = r^{n-1} \Delta_{\mathbf{x}} M_h(\mathbf{x}, r) .$$

We get the conclusion of the proposition by dividing both side of the previous equality by r^{n-1} . $\blacksquare$

Corollary 8.4.3

If $u \in C^2(\mathbb{R}^n \times \mathbb{R})$, let

$$M_u((\mathbf{x}, t), r) = \frac{1}{\omega_n} \int_{\|\mathbf{y}\|=1} u(\mathbf{x} + r\mathbf{y}, t) dS_{\mathbf{y}} = \frac{1}{r^{n-1}\omega_n} \int_{\|\mathbf{y}-\mathbf{x}\|=r} u(\mathbf{y}, t) dS_{\mathbf{y}}$$

for $(\mathbf{x}, t) \in \mathbb{R}^n \times \mathbb{R}$. This is the spherical mean of the function $h(\mathbf{x}) = u(\mathbf{x}, t)$ for $t \in \mathbb{R}$ fixed. Then u satisfies $L(\mathbf{x}, t, D)u = 0$ on $\mathbb{R}^n$ if and only if

$$\frac{\partial^2}{\partial t^2}(M_u((\mathbf{x}, t), r)) - c^2 \left(\frac{\partial^2}{\partial r^2} + \left(\frac{n-1}{r} \right) \frac{\partial}{\partial r} \right) M_u((\mathbf{x}, t), r) = 0 \quad (8.4.5)$$

for $(\mathbf{x}, t) \in \mathbb{R}^n \times \mathbb{R}$ and $r > 0$.

Proof.

The conclusion of the corollary follows from

$$\begin{aligned} M_{L(\mathbf{x}, t, D)u}((\mathbf{x}, t), r) &= M_{\frac{\partial^2 u}{\partial t^2}}((\mathbf{x}, t), r) - c^2 M_{\Delta u}((\mathbf{x}, t), r) \\ &= \frac{\partial^2}{\partial t^2} M_u((\mathbf{x}, t), r) - c^2 \left(\frac{\partial^2}{\partial r^2} + \left(\frac{n-1}{r} \right) \frac{\partial}{\partial r} \right) M_u((\mathbf{x}, t), r) \end{aligned}$$

and the fact that, for a continuous function $h : \mathbb{R}^n \rightarrow \mathbb{R}$, $M_h(\mathbf{x}, r) = 0$ for all $\mathbf{x} \in \mathbb{R}^n$ and $r \in \mathbb{R}$ if and only if $h(\mathbf{x}) = 0$ for all $\mathbf{x} \in \mathbb{R}^n$. ■

Remark 8.4.4

We could use the previous Corollary to find a solution u of (8.1.1) and (8.1.2). We first find M_u satisfying (8.4.5) and the initial conditions $M_u((\mathbf{x}, 0), r) = M_f(\mathbf{x}, r)$ and $\frac{\partial M_u}{\partial t}((\mathbf{x}, 0), r) = M_g(\mathbf{x}, r)$. Then $u(\mathbf{x}, t) = \lim_{r \rightarrow 0} M_u((\mathbf{x}, t), r)$. We will not proceed this way to find the solution of (8.1.1) and (8.1.2). ♠

Lemma 8.4.5

If $h \in C^{k+2}(\mathbb{R})$ with $k > 0$, then

$$\frac{\partial^2}{\partial r^2} \left(\frac{1}{r} \frac{\partial}{\partial r} \right)^{k-1} (r^{2k-1} h(r)) = \left(\frac{1}{r} \frac{\partial}{\partial r} \right)^k \left(r^{2k} \frac{\partial h}{\partial r}(r) \right). \quad (8.4.6)$$

Proof.

A proof by induction on k shows that (8.4.6) is true for $h(r) = r^m$. By linearity, (8.4.6) is true for any polynomial. We also have that (8.4.6) is true at $r = r_0$ if h is such that $\frac{\partial^j h}{\partial r^j}(r_0) = 0$ for $0 \leq j \leq k+1$ because both sides of (8.4.6) are then null.

For $h \in C^{k+2}(\mathbb{R})$ given, choose $r_0 \in \mathbb{R}$. The Taylor's expansion of h about r_0 of order $k+1$ is of the form $h(r) = p(r) + q(r)$, where

$$p(r) = \sum_{j=0}^{k+1} \frac{1}{j!} \frac{\partial^j h}{\partial r^j}(r_0) (r - r_0)^j$$

for $r \in \mathbb{R}$, and

$$q(r) = \frac{1}{(k+1)!} \int_{r_0}^r (r-s)^{k+1} \frac{d^{k+2}h}{ds^{k+2}}(s) ds .$$

Hence, by linearity, (8.4.6) is true for h at $r = r_0$ because (8.4.6) is true for the polynomial p and (8.4.6) is true for the function q at $r = r_0$ since

$$\frac{\partial^j q}{\partial r^j}(r_0) = \frac{1}{(k+1-j)!} \int_{r_0}^r \frac{d^{k+2}h}{ds^{k+2}}(s) ds \Big|_{r=r_0} = 0$$

for $0 \leq j \leq k+1$. Since r_0 is arbitrary, (8.4.6) is true for all $r \in \mathbb{R}$. ■

Lemma 8.4.6

Given $h \in C^{k+2}(\mathbb{R})$ with $k > 0$, let $T_{h,k} : \mathbb{R} \rightarrow \mathbb{R}$ be the function defined by

$$T_{h,k}(r) = \left(\frac{1}{r} \frac{\partial}{\partial r} \right)^{k-1} (r^{2k-1} h(r)) \quad , \quad r > 0 .$$

Then,

$$\frac{\partial^2}{\partial r^2} T_{h,(n-1)/2}(r) = T_{h''+(n-1)/r, h', (n-1)/2}(r)$$

for $r > 0$ and $n > 1$ odd. In other words, $\frac{\partial^2}{\partial r^2}$ is “conjugate” to $\left(\frac{\partial^2}{\partial r^2} + \left(\frac{n-1}{r} \right) \frac{\partial}{\partial r} \right)$.

Moreover, $T_{h,k}(r) = \sum_{j=0}^{k-1} c_j r^{j+1} \frac{\partial^j h}{\partial r^j}(r)$ for some constants $c_0, c_1, \dots, c_{k-1}$. In particular, $c_0 = 1 \cdot 3 \cdot 5 \cdot \dots \cdot (2k-1)$.

Proof.

From the previous lemma, we have

$$\begin{aligned} \frac{\partial^2}{\partial r^2} T_{h,(n-1)/2}(r) &= \left(\frac{1}{r} \frac{\partial}{\partial r} \right)^{(n-1)/2} \left(r^{n-1} \frac{\partial h}{\partial r}(r) \right) \\ &= \left(\frac{1}{r} \frac{\partial}{\partial r} \right)^{(n-3)/2} \left((n-1) r^{n-3} \frac{\partial h}{\partial r}(r) + r^{n-2} \frac{\partial^2 h}{\partial r^2}(r) \right) \\ &= \left(\frac{1}{r} \frac{\partial}{\partial r} \right)^{(n-3)/2} \left(r^{n-2} \left(\left(\frac{n-1}{r} \right) \frac{\partial h}{\partial r}(r) + \frac{\partial^2 h}{\partial r^2}(r) \right) \right) = T_{h''+(n-1)/r, h', (n-1)/2}(r) . \end{aligned}$$

The second conclusion of the lemma can be proved by induction on k and is left to the reader. Note that $c_0 r = \left(\frac{1}{r} \frac{\partial}{\partial r}\right)^{k-1} (r^{2k-1})$ is the coefficient of $h(r)$ in the series expansion of $T_{h,k}(r)$ given in the statement of the lemma. ■

Theorem 8.4.7

If $n > 1$ is odd, then a solution of (8.1.1) with the initial conditions given in (8.1.2), where $f \in C^{(n+3)/2}(\mathbb{R}^n)$ and $g \in C^{(n+1)/2}(\mathbb{R}^n)$, is given by

$$u(\mathbf{x}, t) = \frac{1}{1 \cdot 3 \cdot 5 \cdot \dots \cdot (n-2) \omega_n} \left(\frac{\partial}{\partial t} \left(\frac{1}{t} \frac{\partial}{\partial t} \right)^{(n-3)/2} \left(t^{n-2} \int_{\|\mathbf{y}\|=1} f(\mathbf{x} + ct\mathbf{y}) dS_{\mathbf{y}} \right) + \left(\frac{1}{t} \frac{\partial}{\partial t} \right)^{(n-3)/2} \left(t^{n-2} \int_{\|\mathbf{y}\|=1} g(\mathbf{x} + ct\mathbf{y}) dS_{\mathbf{y}} \right) \right). \quad (8.4.7)$$

Proof.

We may assume that $c = 1$. The substitution $s = ct$ in (8.1.1) yields the equation

$$L(\mathbf{x}, s, D) = \frac{\partial^2 u}{\partial s^2} - \Delta u = 0 \quad (8.4.8)$$

for $(\mathbf{x}, s) \in \mathbb{R}^n \times \mathbb{R}$, and the initial conditions

$$u(\mathbf{x}, 0) = f(\mathbf{x}) \quad \text{and} \quad \frac{\partial u}{\partial s}(\mathbf{x}, 0) = \tilde{g}(\mathbf{x}) \equiv \frac{1}{c} g(\mathbf{x}) \quad (8.4.9)$$

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

We solve (8.4.8) with the initial conditions (8.4.9). For $h \in C^{(n+1)/2}(\mathbb{R}^n)$, let

$$v_h(\mathbf{x}, s) = T_{M_h(\mathbf{x}, \cdot), (n-1)/2}(s) .$$

We have

$$\begin{aligned} \Delta_{\mathbf{x}} v_h(\mathbf{x}, s) &= T_{\Delta_{\mathbf{x}} M_h(\mathbf{x}, \cdot), (n-1)/2}(s) = T_{(\partial^2/\partial s^2 + (n-1)/2 \partial/\partial s) M_h(\mathbf{x}, \cdot), (n-1)/2}(s) \\ &= \frac{\partial^2}{\partial s^2} T_{M_h(\mathbf{x}, \cdot), (n-1)/2}(s) = \frac{\partial^2}{\partial s^2} v_h(\mathbf{x}, s) , \end{aligned}$$

where the second equality comes from Proposition 8.4.2 and the last equality is a consequence of Lemma 8.4.6.

So $v_{\tilde{g}}$ is a solution of (8.4.8). We also have that $\frac{\partial v_f}{\partial s}$ is a solution (8.4.8) because v_f is a solution and $\frac{\partial}{\partial s}$ commutes with $\frac{\partial^2}{\partial s^2} - \Delta_{\mathbf{x}}$.

Thus

$$u(\mathbf{x}, s) = \frac{1}{1 \cdot 3 \cdot 5 \cdot \dots \cdot (n-2)} \left(\frac{\partial v_f}{\partial s} + v_{\tilde{g}} \right)(\mathbf{x}, s) \quad (8.4.10)$$

is a solution of (8.4.8).

Since

$$v_f(\mathbf{x}, s) = T_{M_f(\mathbf{x}, \cdot), (n-1)/2}(s) = c_0 s M_f(\mathbf{x}, s) + c_1 s^2 \frac{\partial M_f}{\partial s}(\mathbf{x}, s) + O(s^3)$$

and

$$v_{\tilde{g}}(\mathbf{x}, s) = T_{M_{\tilde{g}}(\mathbf{x}, \cdot), (n-1)/2}(s) = c_0 s M_{\tilde{g}}(\mathbf{x}, s) + O(s^2) ,$$

we get

$$\begin{aligned} u(\mathbf{x}, s) &= \frac{\partial}{\partial s} \left(s M_f(\mathbf{x}, s) + \frac{c_1}{c_0} s^2 \frac{\partial M_f}{\partial s}(\mathbf{x}, s) + O(s^3) \right) + s M_{\tilde{g}}(\mathbf{x}, s) + O(s^2) \\ &= M_f(\mathbf{x}, s) + \frac{c_0 + 2c_1}{c_0} s \frac{\partial M_f}{\partial s}(\mathbf{x}, s) + s M_{\tilde{g}}(\mathbf{x}, s) + O(s^2) \end{aligned}$$

Hence,

$$u(\mathbf{x}, 0) = M_f(\mathbf{x}, 0) = f(\mathbf{x})$$

and

$$\frac{\partial u}{\partial s}(\mathbf{x}, 0) = \frac{2(c_0 + c_1)}{c_0} \frac{\partial M_f}{\partial s}(\mathbf{x}, 0) + M_{\tilde{g}}(\mathbf{x}, 0) = \tilde{g}(\mathbf{x})$$

because $\frac{\partial M_f}{\partial s}(\mathbf{x}, 0)$ is even with respect to s , and so $\frac{\partial M_f}{\partial s}(\mathbf{x}, 0) = 0$.

The conclusion of the theorem is obtained after substituting $s = ct$ and $\tilde{g} = g/c$ into (8.4.10). $\blacksquare$

For $n = 3$, the solution of (8.1.1) with initial condition (8.1.2) is

$$u(\mathbf{x}, t) = \frac{1}{4\pi} \left(\frac{\partial}{\partial t} \left(t \int_{\|\mathbf{y}\|=1} f(\mathbf{x} + ct\mathbf{y}) dS_{\mathbf{y}} \right) + \left(t \int_{\|\mathbf{y}\|=1} g(\mathbf{x} + ct\mathbf{y}) dS_{\mathbf{y}} \right) \right) . \quad (8.4.11)$$

Remark 8.4.8

There is an easier proof of (8.4.11) than the general proof that has been given above. For $n = 3$, (8.4.5) becomes

$$\frac{\partial^2}{\partial t^2} (M_u((\mathbf{x}, t), r)) - c^2 \left(\frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} \right) M_u((\mathbf{x}, t), r) = 0$$

for $(\mathbf{x}, t) \in \mathbb{R}^3 \times \mathbb{R}$ and $r > 0$. This can be rewritten as

$$\frac{\partial^2}{\partial t^2} (r M_u((\mathbf{x}, t), r)) - c^2 \frac{\partial^2}{\partial r^2} (r M_u((\mathbf{x}, t), r)) = 0$$

for $(\mathbf{x}, t) \in \mathbb{R}^3 \times \mathbb{R}$ and $r > 0$.

It follows from Corollary 8.4.3 that u is a solution of (8.1.1) with the initial conditions (8.1.2) if and only if $v(r, t) = r M_u((\mathbf{x}, t), r)$ is a solution of the wave equation $\frac{\partial^2 v}{\partial t^2} -$

$c^2 \frac{\partial^2 v}{\partial r^2} = 0$ with the initial conditions $v(r, 0) = rM_u((\mathbf{x}, 0), r) = rM_f(\mathbf{x}, r)$ and $\frac{\partial v}{\partial t}(r, 0) = rM_{\frac{\partial u}{\partial t}}((\mathbf{x}, 0), r) = rM_g(\mathbf{x}, r)$, and the boundary condition $v(0, t) = \lim_{r \rightarrow 0} rM_u((\mathbf{x}, t), r) = 0$ because $\lim_{r \rightarrow 0} M_u((\mathbf{x}, t), r) = u(\mathbf{x}, t) \in \mathbb{R}$ as we have seen in Remark 8.4.4. The boundary condition is required because the origin is not uniquely represented in spherical coordinates.

As we have found for the wave equation in $\mathbb{R}$, the solution of this Cauchy problem is

$$v(r, t) = \frac{1}{2} (v(r + ct, 0) + v(r - ct, 0)) + \frac{1}{2c} \int_{r-ct}^{r+ct} \frac{\partial v}{\partial t}(s, 0) ds$$

for $r, t \geq 0$. Moreover, $v(r - ct, 0) = -v(ct - r, 0)$ and $\int_{r-ct}^0 \frac{\partial v}{\partial t}(s, 0) ds = \int_{ct-r}^0 \frac{\partial v}{\partial t}(s, 0) ds$ because $v(r, 0) = rM_f(\mathbf{x}, r)$ and $\frac{\partial v}{\partial t}(r, 0) = rM_g(\mathbf{x}, r)$ are odd functions with respect to r . Recall that $M_f(\mathbf{x}, r)$ and $M_g(\mathbf{x}, r)$ are even with respect to r . Hence

$$\begin{aligned} M_u((\mathbf{x}, t), r) &= \frac{v(r, t)}{r} = \frac{1}{2r} ((ct + r)M_f(\mathbf{x}, ct + r) - (ct - r)M_f(\mathbf{x}, ct - r)) \\ &\quad + \frac{1}{2cr} \int_{ct-r}^{ct+r} sM_g(\mathbf{x}, s) ds . \end{aligned}$$

We have

$$\begin{aligned} u(\mathbf{x}, t) &= \lim_{r \rightarrow 0} M_u((\mathbf{x}, t), r) \\ &= \lim_{r \rightarrow 0} \left(\frac{1}{2r} ((ct + r)M_f(\mathbf{x}, ct + r) - (ct - r)M_f(\mathbf{x}, ct - r)) + \frac{1}{2cr} \int_{ct-r}^{ct+r} sM_g(\mathbf{x}, s) ds \right) \end{aligned}$$

for $\mathbf{x} \in \mathbb{R}^3$ and $t \geq 0$. However,

$$\begin{aligned} \lim_{r \rightarrow 0} \frac{1}{2r} (rM_f(\mathbf{x}, ct + r) + rM_f(\mathbf{x}, ct - r)) &= \lim_{r \rightarrow 0} \frac{M_f(\mathbf{x}, ct + r) + M_f(\mathbf{x}, ct - r)}{2} = M_f(\mathbf{x}, ct) , \\ \lim_{r \rightarrow 0} \frac{1}{2r} (ctM_f(\mathbf{x}, ct + r) - ctM_f(\mathbf{x}, ct - r)) \\ &= t \lim_{r \rightarrow 0} \frac{1}{2r/c} (M_f(\mathbf{x}, c(t + r/c)) - M_f(\mathbf{x}, c(t - r/c))) = t \frac{\partial}{\partial t} M_f(\mathbf{x}, ct) \end{aligned}$$

and

$$\begin{aligned} \lim_{r \rightarrow 0} \frac{1}{2cr} \int_{ct-r}^{ct+r} sM_g(\mathbf{x}, s) ds &= \frac{1}{c} \left(\lim_{r \rightarrow 0} \frac{1}{2r} \left(\int_0^{ct+r} sM_g(\mathbf{x}, s) ds - \int_0^{ct-r} sM_g(\mathbf{x}, s) ds \right) \right) \\ &= \frac{1}{c} (ctM_g(\mathbf{x}, ct)) = tM_g(\mathbf{x}, ct) . \end{aligned}$$

We therefore find that

$$u(\mathbf{x}, t) = M_f(\mathbf{x}, ct) + t \frac{\partial}{\partial t} M_f(\mathbf{x}, ct) + tM_g(\mathbf{x}, ct) = \frac{\partial}{\partial t} (tM_f(\mathbf{x}, ct)) + tM_g(\mathbf{x}, ct)$$

for $\mathbf{x} \in \mathbb{R}^3$ and $t \geq 0$, which is (8.4.11). ♠

8.4.3 Even Dimensional Space

The solution of (8.1.1) with the initial conditions (8.1.2) when n is even is given by the **method of descent**. Namely, if u is a solution on $\mathbb{R}^{n+1} \times \mathbb{R}$ which is independent of x_{n+1} , then u is a solution on $\mathbb{R}^n \times \mathbb{R}$.

Theorem 8.4.9

If $n > 0$ is even, then a solution of (8.1.1) with the initial conditions (8.1.2), where $f \in C^{(n+4)/2}(\mathbb{R}^n)$ and $g \in C^{(n+2)/2}(\mathbb{R}^n)$, is given by

$$u(\mathbf{x}, t) = \frac{2}{1 \cdot 3 \cdot 5 \cdot \dots \cdot (n-1) \omega_{n+1}} \left(\frac{\partial}{\partial t} \left(\frac{1}{t} \frac{\partial}{\partial t} \right)^{(n-2)/2} \left(t^{n-1} \int_{\|\mathbf{y}\| \leq 1} \frac{f(\mathbf{x} + ct\mathbf{y})}{\sqrt{1 - \|\mathbf{y}\|^2}} d\mathbf{y} \right) + \left(\frac{1}{t} \frac{\partial}{\partial t} \right)^{(n-2)/2} \left(t^{n-1} \int_{\|\mathbf{y}\| \leq 1} \frac{g(\mathbf{x} + ct\mathbf{y})}{\sqrt{1 - \|\mathbf{y}\|^2}} d\mathbf{y} \right) \right). \quad (8.4.12)$$

Proof.

If we consider f and g as functions in $\mathbb{R}^{n+1}$ which are independent of x_{n+1} , then the solution of (8.1.1) with the initial conditions (8.1.2) is

$$u(\mathbf{x}, t) = \frac{1}{1 \cdot 3 \cdot 5 \cdot \dots \cdot (n-1) \omega_{n+1}} \left(\frac{\partial}{\partial t} \left(\frac{1}{t} \frac{\partial}{\partial t} \right)^{(n-2)/2} \left(t^{n-1} \int_{\|\mathbf{y}\|=1} f(\mathbf{x} + ct\mathbf{y}) dS_{\mathbf{y}} \right) + \left(\frac{1}{t} \frac{\partial}{\partial t} \right)^{(n-2)/2} \left(t^{n-1} \int_{\|\mathbf{y}\|=1} g(\mathbf{x} + ct\mathbf{y}) dS_{\mathbf{y}} \right) \right), \quad (8.4.13)$$

where $\mathbf{x}$ and $\mathbf{y}$ are in $\mathbb{R}^{n+1}$.

If we write $\mathbf{y} = (\tilde{\mathbf{y}}, y_{n+1})$ and $\mathbf{x} = (\tilde{\mathbf{x}}, x_{n+1})$, then

$$\int_{\|\mathbf{y}\|=1} f(\mathbf{x} + ct\mathbf{y}) dS_{\mathbf{y}} = \int_{\|\tilde{\mathbf{y}}\|^2 + |y_{n+1}|^2 = 1} f(\tilde{\mathbf{x}} + ct\tilde{\mathbf{y}}) dS_{\tilde{\mathbf{y}}}.$$

If we parameterize the upper part (i.e. $y_{n+1} > 0$) and lower part (i.e. $y_{n+1} < 0$) of the sphere of radius 1 centred at the origin in $\mathbb{R}^{n+1}$ with $y_{n+1} = \phi_{\pm}(\tilde{\mathbf{y}}) = \pm \sqrt{1 - \|\tilde{\mathbf{y}}\|^2}$ for $\|\tilde{\mathbf{y}}\| \leq 1$ (Figure 8.3), then the previous integral becomes

$$\int_{\|\mathbf{y}\|=1} f(\mathbf{x} + ct\mathbf{y}) dS_{\mathbf{y}} = 2 \int_{\|\tilde{\mathbf{y}}\| \leq 1} f(\tilde{\mathbf{x}} + ct\tilde{\mathbf{y}}) \sqrt{1 + \|\nabla \phi_{+}(\tilde{\mathbf{y}})\|^2} d\tilde{\mathbf{y}}.$$

We have used the fact that $f(\mathbf{x} + t\mathbf{y})$ is independent of $x_{n+1} + ty_{n+1}$ to replace the integral over the lower part of the sphere by the integral over the upper part. Finally, since

$$\nabla \phi_{+}(\tilde{\mathbf{y}}) = \frac{-1}{\sqrt{1 - \|\tilde{\mathbf{y}}\|^2}} \tilde{\mathbf{y}},$$

we get

$$\int_{\|\mathbf{y}\|=1} f(\mathbf{x} + ct\mathbf{y}) dS_{\mathbf{y}} = 2 \int_{\|\tilde{\mathbf{y}}\| \leq 1} \frac{f(\tilde{\mathbf{x}} + ct\tilde{\mathbf{y}})}{\sqrt{1 - \|\tilde{\mathbf{y}}\|^2}} d\tilde{\mathbf{y}}. \quad (8.4.14)$$

An identical reasoning yields

$$\int_{\|\mathbf{y}\|=1} g(\mathbf{x} + ct\mathbf{y}) dS_{\mathbf{y}} = 2 \int_{\|\tilde{\mathbf{y}}\|\leq 1} \frac{g(\tilde{\mathbf{x}} + ct\tilde{\mathbf{y}})}{\sqrt{1 - \|\tilde{\mathbf{y}}\|^2}} d\tilde{\mathbf{y}}. \quad (8.4.15)$$

We obtain (8.4.13) from (8.4.14) and (8.4.15). ■

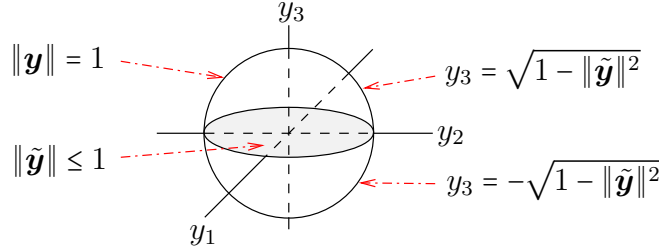


Figure 8.3: The implementation of the method of descent in the proof of Theorem 8.4.9 reduces the surface integral on the sphere of radius 1 to the integral over the disk $\|\tilde{\mathbf{y}}\| \leq 1$.

For $n = 2$, the solution of (8.1.1) with initial condition (8.1.2) is

$$u(\mathbf{x}, t) = \frac{1}{2\pi} \left(\frac{\partial}{\partial t} \left(t \int_{\|\mathbf{y}\|\leq 1} \frac{f(\mathbf{x} + ct\mathbf{y})}{\sqrt{1 - \|\mathbf{y}\|^2}} d\mathbf{y} \right) + t \int_{\|\mathbf{y}\|\leq 1} \frac{g(\mathbf{x} + ct\mathbf{y})}{\sqrt{1 - \|\mathbf{y}\|^2}} d\mathbf{y} \right).$$

8.5 Huygens Principle

Let u be the solution of (8.1.1) with the initial conditions (8.1.2).

When $n = 1$, the value of u at (x, t) is determined by the value of f at $x \pm ct$ and the values of g on the interval $[x - ct, x + ct]$. If g is not null on $\mathbb{R}$, then u will generally be non-null at (x, t) such that $[x - ct, x + ct] \cap \text{supp } g \neq \emptyset$. Physically, it means that, starting at the time $\tilde{t}$ when the listener at position x begins to hear a sound (i.e. $[x - c\tilde{t}, x + c\tilde{t}] \cap \text{supp } g \neq \emptyset$), he or she will keep hearing noises for ever because $[x - ct, x + ct] \cap \text{supp } g$ will be non-empty for $t > \tilde{t}$.

When $n = 2$ (or n even), we can draw the same conclusion than with $n = 1$. The value of u at $(\mathbf{x}, t)$ is determined by the value of f and g in the disk of radius ct centred at $\mathbf{x}$. If f and g are not null on $\mathbb{R}^2$, then u will generally be non-null at $(\mathbf{x}, t)$ such that $\{\mathbf{y} : \|\mathbf{y} - \mathbf{x}\| \leq ct\} \cap (\text{supp } f \cup \text{supp } g) \neq \emptyset$. Physically, it means that, starting at the time $\tilde{t}$ when the listener at position $\mathbf{x}$ begins to hear a sound (i.e. $\{\mathbf{y} : \|\mathbf{y} - \mathbf{x}\| \leq c\tilde{t}\} \cap (\text{supp } f \cup \text{supp } g) \neq \emptyset$), he or she will keep hearing noises for ever because $\{\mathbf{y} : \|\mathbf{y} - \mathbf{x}\| \leq ct\} \cap (\text{supp } f \cup \text{supp } g)$ will be non-empty for $t > \tilde{t}$ (Figure 8.4).

Finally, when $n = 3$ (or n odd and greater than 1, the value u at $(\mathbf{x}, t)$ is determined by the value of f and g on the surface of the ball of radius ct centres at $\mathbf{x}$; namely, $\{\mathbf{y} : \|\mathbf{y} - \mathbf{x}\| = ct\}$. If f and g are not null on $\mathbb{R}^3$, then u will generally be non-null at $(\mathbf{x}, t)$ such that $\{\mathbf{y} : \|\mathbf{y} - \mathbf{x}\| = ct\} \cap (\text{supp } f \cup \text{supp } g) \neq \emptyset$. Physically, it means that the listener at position $\mathbf{x}$ will hear the sound as long as $\{\mathbf{y} : \|\mathbf{y} - \mathbf{x}\| = ct\} \cap (\text{supp } f \cup \text{supp } g) \neq \emptyset$. If f and g have compact supports, the listener will hear the sound for only a finite among of time. He or she will not near the sound when t is large enough to have $\text{supp } f \cup \text{supp } g \subset B_{ct}(\mathbf{x})$ because the intersection of $\{\mathbf{y} : \|\mathbf{y} - \mathbf{x}\| = ct\}$ with $\text{supp } f \cup \text{supp } g$ will be empty.

We are lucky to live in a three-dimensional world though sometime we may wonder if we are not living in a two-dimensional world because of all the noise around us all the time.

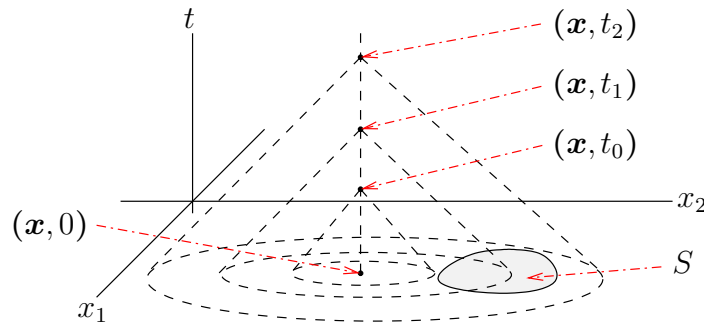


Figure 8.4: The figure illustrates the Huygens principle for n even. The region S in the x_1, x_2 plane is the union of the support of f and g , and $0 < t_1 < t_2 < t_3$.

8.6 Nonhomogeneous Wave Equations

Consider the Nonhomogeneous wave equation

$$L(\mathbf{x}, t, D) = \frac{\partial^2 u}{\partial t^2} - c^2 \Delta u = w(\mathbf{x}, t) \quad , \quad (\mathbf{x}, t) \in \mathbb{R}^n \times \mathbb{R} \quad ,$$

with the initial condition

$$u(\mathbf{x}, 0) = f(\mathbf{x}) \quad \text{and} \quad \frac{\partial u}{\partial t}(\mathbf{x}, 0) = g(\mathbf{x}) \quad , \quad \mathbf{x} \in \mathbb{R}^n \quad .$$

We seek a solution of the form $u = u_1 + u_2$, where u_1 is a solution of the (homogeneous) wave equation

$$L(\mathbf{x}, t, D) = \frac{\partial^2 u}{\partial t^2} - c^2 \Delta u = 0 \quad , \quad (\mathbf{x}, t) \in \mathbb{R}^n \times \mathbb{R} \quad ,$$

with the initial condition

$$u(\mathbf{x}, 0) = f(\mathbf{x}) \quad \text{and} \quad \frac{\partial u}{\partial t}(\mathbf{x}, 0) = g(\mathbf{x}) \quad , \quad \mathbf{x} \in \mathbb{R}^n \quad ,$$

and u_2 is the solution of the nonhomogeneous wave equation

$$L(\mathbf{x}, t, D) = \frac{\partial^2 u}{\partial t^2} - c^2 \Delta u = w(\mathbf{x}, t) \quad , \quad (\mathbf{x}, t) \in \mathbb{R}^n \times \mathbb{R} \quad , \quad (8.6.1)$$

with the initial condition

$$u(\mathbf{x}, 0) = 0 \quad \text{and} \quad \frac{\partial u}{\partial t}(\mathbf{x}, 0) = 0 \quad , \quad \mathbf{x} \in \mathbb{R}^n \quad . \quad (8.6.2)$$

The function u_1 is given by one of the formulae in (8.4.1), Theorem 8.4.7 and Theorem 8.4.9 according to the value of n . We now show how to find u_2 .

Theorem 8.6.1 (Duhamel's Principle)

Suppose that $w \in C^{\lfloor n/2 \rfloor + 1}(\mathbb{R}^n \times \mathbb{R})$. For $s \in \mathbb{R}$, let $v(\mathbf{x}, t, s)$ be the solution of

$$L(\mathbf{x}, t, D) = \frac{\partial^2 v}{\partial t^2} - c^2 \Delta v = 0 \quad , \quad (\mathbf{x}, t) \in \mathbb{R}^n \times \mathbb{R} \quad , \quad (8.6.3)$$

with the initial condition

$$v(\mathbf{x}, 0, s) = 0 \quad \text{and} \quad \frac{\partial v}{\partial t}(\mathbf{x}, 0, s) = w(\mathbf{x}, s) \quad , \quad \mathbf{x} \in \mathbb{R}^n \quad . \quad (8.6.4)$$

Then

$$u(\mathbf{x}, t) = \int_0^t v(\mathbf{x}, t-s, s) \, ds \quad , \quad (\mathbf{x}, t) \in \mathbb{R}^n \times \mathbb{R} \quad ,$$

is the solution of (8.6.1) satisfying (8.6.2).

Remark 8.6.2

Recall that $\lfloor n/2 \rfloor$ is the greatest integer smaller than or equal to $n/2$. For each s , the solution v of (8.6.3) satisfying (8.6.4) is given by one of the formulae in (8.4.1), (8.4.7) and (8.4.12) according to the value of n . ♠

Proof.

The initial conditions (8.6.2) are satisfied. We have $u(\mathbf{x}, 0) = 0$ for all $\mathbf{x} \in \mathbb{R}^n$. Moreover,

$$\frac{\partial u}{\partial t}(\mathbf{x}, t) = v(\mathbf{x}, 0, t) + \int_0^t \frac{\partial v}{\partial t}(\mathbf{x}, t-s, s) \, ds = \int_0^t \frac{\partial v}{\partial t}(\mathbf{x}, t-s, s) \, ds$$

because $v(\mathbf{x}, 0, t) = 0$ is one of the initial conditions that v must satisfy. Hence $\frac{\partial u}{\partial t}(\mathbf{x}, 0) = 0$ for all $\mathbf{x} \in \mathbb{R}^n$.

The function u satisfies (8.6.1) because

$$\begin{aligned} \frac{\partial^2 u}{\partial t^2}(\mathbf{x}, t) &= \frac{\partial v}{\partial t}(\mathbf{x}, 0, t) + \int_0^t \frac{\partial^2 v}{\partial t^2}(\mathbf{x}, t-s, s) \, ds = w(\mathbf{x}, t) + \int_0^t c^2 \Delta_{\mathbf{x}} v(\mathbf{x}, t-s, s) \, ds \\ &= w(\mathbf{x}, t) + c^2 \Delta_{\mathbf{x}} \int_0^t v(\mathbf{x}, t-s, s) \, ds = w(\mathbf{x}, t) + c^2 \Delta_{\mathbf{x}} u(\mathbf{x}, t) \quad . \end{aligned}$$

We may interchange integration and partial derivatives with respect to x_j , $1 \leq j \leq n$, because v is at least of class C^2 .

The condition $w \in C^{\lfloor n/2 \rfloor + 1}(\mathbb{R}^n \times \mathbb{R})$ is required to ensure that Theorems 8.4.7 and 8.4.9 could be used to find v . ■

8.7 Separation of Variables

8.7.1 Cartesian Coordinates

Let $D = \{(x, y) : 0 < x < a \text{ and } 0 < y < b\}$. Using separation of variables as we did for the wave equation in one-dimension, we find that the solutions of the wave equation

$$\frac{\partial^2 u}{\partial t^2} = c^2 \Delta u = c^2 \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right) \quad , \quad (x, y) \in D \text{ and } t > 0 \quad ,$$

with some boundary and initial conditions are linear combinations of simple functions of the form

$$u(x, y, t) = (A \cos(ckt) + B \sin(ckt))(C \cos(\mu x) + D \sin(\mu x))(E \cos(\nu y) + F \sin(\nu y))$$

with $k^2 = \mu^2 + \nu^2$.

Example 8.7.1

Let $D = \{(x, y) : 0 < x < 4, 0 < y < 2\}$ and $c^2 = 9$. Find the solution of the wave equation

$$\frac{\partial^2 u}{\partial t^2} = c^2 \Delta u = c^2 \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right) \quad , \quad (x, y) \in D \text{ and } t > 0 \quad ,$$

with the boundary conditions $u(x, 0, t) = u(x, 2, t) = 0$ for $0 < x < 4$ and $u(0, y, t) = u(4, y, t) = 0$ for $0 < y < 2$, and the initial conditions $u(x, y, 0) = (4x - x^2)(2y - y^2)$ and $\frac{\partial u}{\partial t}(x, y, 0) = 5 \sin\left(\frac{3\pi x}{4}\right) \sin\left(\frac{7\pi y}{2}\right)$ for $(x, y) \in D$.

The simple functions are of the form

$$u(x, y, t) = (A \cos(ckt) + B \sin(ckt))(C \cos(\mu x) + D \sin(\mu x))(E \cos(\nu y) + F \sin(\nu y))$$

with $k^2 = \nu^2 + \mu^2$.

We get from $u(x, 0, t) = 0$ that $E(A \cos(ckt) + B \sin(ckt))(C \cos(\mu x) + D \sin(\mu x)) = 0$ for $0 < x < 4$ and $t > 0$. Thus, $E = 0$. We then get from $u(x, 2, t) = 0$ that $F \sin(2\nu)(A \cos(ckt) + B \sin(ckt))(C \cos(\mu x) + D \sin(\mu x)) = 0$ for $0 < x < 4$ and $t > 0$. Thus, $F \sin(2\nu) = 0$. Since we are looking for non-null solutions, we must have $F \neq 0$ and $\sin(2\nu) = 0$. This implies that $\nu = n\pi/2$ for $n > 0$.

Since $E = 0$ and $F \neq 0$, we get from $u(0, y, t) = 0$ that $CF(A \cos(ckt) + B \sin(ckt)) \sin(n\pi y/2) = 0$ for $0 < y < 2$ and $t > 0$. Thus $C = 0$. We then get

from $u(4, y, t) = 0$ that $DF \sin(4\mu) (A \cos(ckt) + B \sin(ckt)) \sin(2\pi y/2) = 0$ for $0 < y < 2$ and $t > 0$. Thus $D \sin(4\mu) = 0$. Since we are looking for non-null solutions, we must have $D \neq 0$ and $\sin(4\mu) = 0$. This implies that $\mu = m\pi/4$ for $m > 0$.

We have found that the simple functions are of the form

$$u_{n,m}(x, y, t) = (A_{n,m} \cos(3k_{n,m}t) + B_{n,m} \sin(3k_{n,m}t)) \sin\left(\frac{m\pi x}{4}\right) \sin\left(\frac{n\pi y}{2}\right),$$

where $k_{n,m} = \sqrt{\left(\frac{m\pi}{4}\right)^2 + \left(\frac{n\pi}{2}\right)^2}$ and $n, m > 0$.

The solution is of the form

$$\begin{aligned} u(x, y) &= \sum_{n,m=1}^{\infty} u_{n,m}(x, y) \\ &= \sum_{n,m=1}^{\infty} (A_{n,m} \cos(3k_{n,m}t) + B_{n,m} \sin(3k_{n,m}t)) \sin\left(\frac{m\pi x}{4}\right) \sin\left(\frac{n\pi y}{2}\right). \end{aligned}$$

From $u(x, y, 0) = (4x - x^2)(2y - y^2)$, we get

$$\sum_{n,m=1}^{\infty} A_{n,m} \sin\left(\frac{m\pi x}{4}\right) \sin\left(\frac{n\pi y}{2}\right) = (4x - x^2)(2y - y^2).$$

This is the Fourier sine series in $L^2(D)$ of $(4x - x^2)(2y - y^2)$ ². Thus

$$\begin{aligned} A_{n,m} &= \frac{4}{2 \cdot 4} \int_0^4 \int_0^2 (4x - x^2)(2y - y^2) \sin\left(\frac{m\pi x}{4}\right) \sin\left(\frac{n\pi y}{2}\right) dy dx \\ &= \frac{1}{2} \left(\int_0^4 (4x - x^2) \sin\left(\frac{m\pi x}{4}\right) dx \right) \left(\int_0^2 (2y - y^2) \sin\left(\frac{n\pi y}{2}\right) dy \right) \\ &= \frac{1}{2} \left(\frac{128}{m^3 \pi^3} (1 - (-1)^m) \right) \left(\frac{16}{n^3 \pi^3} (1 - (-1)^n) \right) = \frac{1024}{n^3 m^3 \pi^6} (1 - (-1)^m) (1 - (-1)^n). \end{aligned}$$

Since

$$\frac{\partial u}{\partial t}(x, y, t) = \sum_{n,m=1}^{\infty} (-3k_{n,m} A_{n,m} \sin(3k_{n,m}t) + 3k_{n,m} B_{n,m} \cos(3k_{n,m}t)) \sin\left(\frac{m\pi x}{4}\right) \sin\left(\frac{n\pi y}{2}\right),$$

we get from $\frac{\partial u}{\partial t}(x, y, 0) = 5 \sin\left(\frac{3\pi x}{4}\right) \sin\left(\frac{7\pi y}{2}\right)$ that

$$\sum_{n,m=1}^{\infty} 3k_{n,m} B_{n,m} \sin\left(\frac{m\pi x}{4}\right) \sin\left(\frac{n\pi y}{2}\right) = 5 \sin\left(\frac{3\pi x}{4}\right) \sin\left(\frac{7\pi y}{2}\right).$$

Using the uniqueness of the sine Fourier series, we find

$$B_{n,m} = \begin{cases} 5/(3k_{n,m}) & \text{if } m = 3, n = 7 \\ 0 & \text{otherwise} \end{cases}$$

²We leave it to the reader to generalize the concept of classical Fourier series in $L^2(I)$, where I is an interval in $\mathbb{R}$, to $L^2(D)$, where D is a rectangle in $\mathbb{R}^2$, using Theorem 4.1.13.

The solution is

$$\begin{aligned}
 u(x, y, t) &= \sum_{n,m=1}^{\infty} \frac{1024}{n^3 m^3 \pi^6} (1 - (-1)^m)(1 - (-1)^n) \cos(3k_{n,m}t) \sin\left(\frac{m\pi x}{4}\right) \sin\left(\frac{n\pi y}{2}\right) \\
 &\quad + \frac{5}{3k_{7,3}} \sin(3k_{7,3}t) \sin\left(\frac{3\pi x}{4}\right) \sin\left(\frac{7\pi y}{2}\right) \\
 &= \sum_{p,q=1}^{\infty} \frac{4096}{(2q-1)^3(2p-1)^3\pi^6} \cos(3k_{2q-1,2p-1}t) \sin\left(\frac{(2p-1)\pi x}{4}\right) \sin\left(\frac{(2q-1)\pi y}{2}\right) \\
 &\quad + \frac{5}{3k_{7,3}} \sin(3k_{7,3}t) \sin\left(\frac{3\pi x}{4}\right) \sin\left(\frac{7\pi y}{2}\right).
 \end{aligned}$$

♣

Remark 8.7.2

If we consider that the wave equation in the previous example describes the membrane of a square drum, the function $u_{n,m}(x, y)$ is the n, m -harmonic. The first-harmonic is given by $n = m = 1$. The frequency of the first-harmonic is $3\sqrt{3}/4$. However, the frequencies of the other harmonics are $3\sqrt{m^2/4 + n^2/2}/2$ with $n + m > 2$ and $n, m > 0$. Hence, most of the other harmonics do not have frequencies which are integer multiples of $3\sqrt{3}/4$. This is the reason why we cannot get an harmonious square drum. ♠

8.7.2 Polar, Cylindrical and Spherical Coordinates

The polar coordinates are related to the Cartesian coordinates in the plane by the transformation

$$\begin{aligned}
 \mathbb{R}^2 &\rightarrow \mathbb{R}^2 \\
 (r, \theta) &\mapsto (x, y) = (r \cos(\theta), r \sin(\theta))
 \end{aligned}$$

The cylindrical coordinates are related to the Cartesian coordinates in space by the transformation

$$\begin{aligned}
 \mathbb{R}^3 &\rightarrow \mathbb{R}^3 \\
 (r, \theta, z) &\mapsto (x, y, z) = (r \cos(\theta), r \sin(\theta), z)
 \end{aligned}$$

The cylindrical coordinates are the polar coordinates in the x, y plane combined with the normal z coordinate. For this reason, the results about partial differential equations in polar coordinates will apply to partial differential equations in cylindrical coordinates which are independent of z and vice-versa.

The mappings above are not bijections from $\mathbb{R}^2$ (or $\mathbb{R}^3$) to itself. They are periodic of period 2π in θ , $(0, \theta)$ is mapped to $(0, 0)$ whatever the value of θ , (r, θ) and $(-r, \theta + \pi)$ are mapped to the same point of $\mathbb{R}^2$, and so on.

The spherical coordinates are related to the Cartesian coordinates by the transformation

$$\mathbb{R}^3 \rightarrow \mathbb{R}^3$$

$$(r, \theta, \phi) \mapsto (x, y, z) = (r \cos(\theta) \sin(\phi), r \sin(\theta) \sin(\phi), r \cos(\phi))$$

As for the previous change of coordinates, this is not a bijection from $\mathbb{R}^3$ to itself. It is periodic of period 2π in θ and ϕ , $(0, \theta, \phi)$ is mapped to $(0, 0, 0)$ whatever the value of the θ and ϕ , $(r, \theta, 0)$ is mapped to $(r, 0, 0)$ whatever the value of θ , (r, θ, ϕ) and $(-r, \theta, \phi + \pi)$ are mapped to the same point of $\mathbb{R}^3$, and so on.

Nevertheless, locally and away from the origin, these transformations are bijections.

Using the chain rule, the reader can verify that the Laplacian in polar coordinates is

$$\Delta u = \frac{\partial^2 u}{\partial r^2} + \frac{1}{r} \frac{\partial u}{\partial r} + \frac{1}{r^2} \frac{\partial^2 u}{\partial \theta^2}.$$

The Laplacian in cylindrical coordinates is

$$\Delta u = \frac{\partial^2 u}{\partial r^2} + \frac{1}{r} \frac{\partial u}{\partial r} + \frac{1}{r^2} \frac{\partial^2 u}{\partial \theta^2} + \frac{\partial^2 u}{\partial z^2}.$$

Finally, the Laplacian in spherical coordinates is

$$\Delta u = \frac{\partial^2 u}{\partial r^2} + \frac{2}{r} \frac{\partial u}{\partial r} + \frac{1}{r^2} \frac{\partial^2 u}{\partial \phi^2} + \frac{\cot(\phi)}{r^2} \frac{\partial u}{\partial \phi} + \frac{1}{r^2 \sin^2(\phi)} \frac{\partial^2 u}{\partial \theta^2}.$$

Example 8.7.3

Solve the wave equation

$$\frac{\partial^2 u}{\partial t^2} = c^2 \Delta u \quad , \quad x^2 + y^2 < R^2 \text{ and } t > 0,$$

with the boundary condition $u(x, y, t) = 0$ for $x^2 + y^2 = R^2$ and $t > 0$, and the initial conditions $u(x, y, 0) = f(\sqrt{x^2 + y^2})$ and $\frac{\partial u}{\partial t}(x, y, 0) = g(\sqrt{x^2 + y^2})$ for $x^2 + y^2 \leq R^2$.

Since the boundary and initial conditions are functions of the distance $r = \sqrt{x^2 + y^2}$ between (x, y) and the origin, we may use polar coordinates and assume that u is a function of r and θ . As is traditional, we keep the same notation for u as a function of (x, y) and u as a function of (r, θ) . So $u(x, y) = u(r, \theta)$ where $x = r \cos(\theta)$ and $y = r \sin(\theta)$.

In polar coordinates, the boundary condition is $u(R, \theta, t) = 0$ for $0 \leq \theta < 2\pi$ and $t > 0$, and the initial conditions are $u(r, \theta, 0) = f(r)$ and $\frac{\partial u}{\partial t}(r, \theta, 0) = g(r)$ for $0 \leq r \leq R$ and $0 \leq \theta < 2\pi$.

The wave equation in polar coordinates is

$$\frac{\partial^2 u}{\partial t^2} = c^2 \Delta u = c^2 \left(\frac{\partial^2 u}{\partial r^2} + \frac{1}{r} \frac{\partial u}{\partial r} \right). \quad (8.7.1)$$

Separation of Variables: If we substitute $u(r, t) = F(r)G(t)$ in (8.7.1), we get

$$F(r) \frac{d^2 G}{dt^2}(t) = c^2 \left(\frac{d^2 F}{dr^2}(r) G(t) + \frac{1}{r} \frac{dF}{dr}(r) G(t) \right).$$

If we divide both side by $c^2 F(r)G(t)$, we get

$$\frac{1}{c^2 G(t)} \frac{d^2 G}{dt^2}(t) = \frac{1}{F(r)} \left(\frac{d^2 F}{dr^2}(r) + \frac{1}{r} \frac{dF}{dr}(r) \right) .$$

Since the right hand side is independent of t and the left hand side is independent of r , both sides must be constant. We have

$$\frac{1}{c^2 G(t)} \frac{d^2 G}{dt^2}(t) = \frac{1}{F(r)} \left(\frac{d^2 F}{dr^2}(r) + \frac{1}{r} \frac{dF}{dr}(r) \right) = q ,$$

where q is a constant. This gives two ordinary differential equations,

$$\frac{d^2 G}{dt^2}(t) - c^2 q G(t) = 0 \quad (8.7.2)$$

and

$$\frac{d^2 F}{dr^2}(r) + \frac{1}{r} \frac{dF}{dr}(r) - q F(r) = 0 . \quad (8.7.3)$$

The boundary condition $u(R, t) = F(R)G(t) = 0$ for all t yields $F(R) = 0$.

Simple Functions: To get a bounded solution of (8.7.2) for all t , we must have $q \leq 0$. Let $q = -k^2$.

We consider (8.7.3) with $q = -k^2$ and the condition $F(R) = 0$. Let $s = kr$. Since $\frac{dF}{dr} = \frac{dF}{ds} \frac{ds}{dr} = k \frac{dF}{ds}$ and $\frac{d^2 F}{dr^2} = \frac{d}{dr} \left(\frac{dF}{dr} \right) = \frac{d}{ds} \left(k \frac{dF}{ds} \right) \frac{ds}{dr} = k^2 \frac{d^2 F}{ds^2}$, The differential equation in (8.7.3) becomes

$$k^2 \frac{d^2 F}{ds^2}(s) + \frac{k^2}{s} \frac{dF}{ds}(s) + k^2 F(s) = 0 .$$

As usual, we keep the same name for F as a function of r or s . If we multiply both sides of the previous ordinary differential equation by s^2/k^2 , we get

$$s^2 \frac{d^2 F}{ds^2}(s) + s \frac{dF}{ds}(s) + s^2 F(s) = 0 .$$

This is the Bessel equation of order zero (i.e. $\nu = 0$ in the Bessel equation). The general solution of this ordinary differential equation is $F(s) = C J_0(s) + D Y_0(s)$, where $J_0(s)$ is a Bessel function of the first kind and Y_0 is a Bessel function of the second kind. Since we are seeking bounded solutions, we must have $D = 0$ because the Bessel functions of the second kind are not bounded at the origin.

The bounded solutions of (8.7.3) with $q = -k^2$ are therefore of the form $F(r) = C J_0(kr)$. To determine k , we use the boundary condition $F(R) = 0$. From $F(R) = C J_0(kR) = 0$, we get $k = k_n \equiv \alpha_{n,0}/R$ for $n > 0$, where $\alpha_{n,0}$ for $n > 0$ are the non-null zeros of J_0 . We have shown that the bounded solutions of (8.7.3) with $q = -k^2$ and satisfying $F(R) = 0$ exist only for $k = k_n = \alpha_{n,0}/R$ and are multiples of $F_n(r) \equiv J_0(\alpha_{n,0}r/R)$ for $n > 0$.

For $n > 0$ fix, the general solution of (8.7.2) with $q = -k_n^2 = -(\alpha_{n,0}/R)^2$ is

$$G(t) = G_n(t) \equiv A_n \cos\left(\frac{c\alpha_{n,0}t}{R}\right) + B_n \sin\left(\frac{c\alpha_{n,0}t}{R}\right).$$

We have found the simple functions

$$u_n(r, t) \equiv F_n(r)G_n(t) = \left(A_n \cos\left(\frac{c\alpha_{n,0}t}{R}\right) + B_n \sin\left(\frac{c\alpha_{n,0}t}{R}\right)\right) J_0\left(\frac{\alpha_{n,0}r}{R}\right)$$

for $n > 0$.

Initial Conditions: We seek a solution of the form

$$u(r, t) = \sum_{n=1}^{\infty} \left(A_n \cos\left(\frac{c\alpha_{n,0}t}{R}\right) + B_n \sin\left(\frac{c\alpha_{n,0}t}{R}\right)\right) J_0\left(\frac{\alpha_{n,0}r}{R}\right).$$

From $u(r, 0) = f(r)$, we get

$$f(r) = \sum_{n=1}^{\infty} A_n J_0\left(\frac{\alpha_{n,0}r}{R}\right).$$

This is the Fourier Bessel series of f on the interval $]0, R[$. Thus

$$A_n = \frac{2}{R^2 J_1^2(\alpha_{n,0})} \int_0^R r f(r) J_0\left(\frac{\alpha_{n,0}r}{R}\right) dr$$

for $n > 0$. From $\frac{\partial u}{\partial t}(r, 0) = g(r)$, we get

$$g(r) = \sum_{n=1}^{\infty} \frac{c\alpha_{n,0}B_n}{R} J_0\left(\frac{\alpha_{n,0}r}{R}\right).$$

Again, this is a Fourier Bessel series. This time, it is the Fourier Bessel series of g on the interval $]0, R[$. Thus

$$\frac{c\alpha_{n,0}B_n}{R} = \frac{2}{R^2 J_1^2(\alpha_{n,0})} \int_0^R r g(r) J_0\left(\frac{\alpha_{n,0}r}{R}\right) dr$$

for $n > 0$, which yields

$$B_n = \frac{2}{c\alpha_{n,0}R J_1^2(\alpha_{n,0})} \int_0^R r g(r) J_0\left(\frac{\alpha_{n,0}r}{R}\right) dr$$

for $n > 0$. ♣

Remark 8.7.4

The simple function $u_n(r, t)$ of the previous example is called the **nth-normal mode** or **nth-harmonic**. The first-harmonic is called the **fundamental harmonic**. Its frequency is $c\alpha_{n,0}/(2\pi R)$. The circles associated to the values of r for which $u_n(r, \theta) = 0$ are called the

modal lines of u_n . For n fixed, there are $n - 1$ modal lines given by $r_{n,i} = \alpha_{i,0}R/\alpha_{n,0}$ for $0 < i < n$. The modal lines are not evenly distributed.

If we imagine that the domain of u defined by $x^2 + y^2 \leq R^2$ is the membrane of a circular drum, then the modal lines $x^2 + y^2 \leq r_{n,i}^2$ are regions of the membrane that not vibrate under the frequency $c\alpha_{n,0}/(2\pi R)$. The frequencies obviously determine the sound of the drum. ♠

Remark 8.7.5

Using separation of variables, we find that the bounded solutions of the wave equation in (8.7.1) on a disk $B_R(\mathbf{0})$ are linear combinations of simple functions of the form

$$u(r, t) = (A \cos(ckt) + B \sin(ckt)) J_0(kr) .$$

♠

8.8 Exercises

Suggested exercises:

- In [22]: numbers 1 to 4 in Sections 3.2.
- In [24]: numbers 9.5, 9.6, 9.14 and 9.18 in Section 9.13.
- In [38]: numbers 3 to 7, 9, 11 to 14, 16 to 18 in Section 9.2; numbers 1 and 2 in Section 10.1; numbers 1 to 4 in Section 10.2; numbers 1 to 4 in Section 10.3.

Chapter 9

Laplace Equation

Let Ω be an open subset of $\mathbb{R}^n$. The **Laplace equation** is the partial differential equation

$$\Delta u = 0 \quad , \quad \mathbf{x} \in \Omega . \quad (9.0.1)$$

The problem is to find u that satisfies the Laplace equation and some boundary conditions. The boundary of Ω is denoted by $\partial\Omega$. The most frequent boundary conditions are listed below.

Dirichlet Boundary Conditions: Given $f : \partial\Omega \rightarrow \mathbb{R}$, we require that $u(\mathbf{x}) = f(\mathbf{x})$ for $\mathbf{x} \in \partial\Omega$. We write $u|_{\partial\Omega} = f$.

Neumann Boundary Conditions: Given $g : \partial\Omega \rightarrow \mathbb{R}$, we require that $\frac{\partial u}{\partial \nu}(\mathbf{x}) = g(\mathbf{x})$ for $\mathbf{x} \in \partial\Omega$, where $\nu(\mathbf{x})$ is the outward unit normal to the surface $\partial\Omega$ at $\mathbf{x}$. We write $\frac{\partial u}{\partial \nu}\Big|_{\partial\Omega} = g$.

Mixed Boundary Conditions: A mix of the previous two conditions. We assign u on some parts of $\partial\Omega$ and $\frac{\partial u}{\partial \nu}$ on the other parts of $\partial\Omega$.

Robin Boundary Conditions: Given $u_0 : \partial\Omega \rightarrow \mathbb{R}$, we require that

$$\frac{\partial u}{\partial \nu}(\mathbf{x}) + \alpha (u(\mathbf{x}) - u_0(\mathbf{x})) = 0 \quad , \quad \mathbf{x} \in \partial\Omega ,$$

where α is a constant. We write $\left(\frac{\partial u}{\partial \nu} + \alpha(u - u_0)\right)\Big|_{\partial\Omega} = 0$.

For the steady-state heat problem. Robin Boundary Conditions is used to express the fact that heat flow across $\partial\Omega$ is proportional to the difference between the temperature u inside Ω and the temperature u_0 outside Ω . This is the **Newton's Law** of heating (or cooling).

Equation (9.0.1) is also called **potential Equation** and the solutions of this equation are called **potential functions** or **harmonic functions**. We will give in Definition 9.1.3 a more convenient definition of harmonic function.

The non-homogeneous equation

$$\Delta u = f(\mathbf{x}) \quad , \quad \mathbf{x} \in \Omega ,$$

is called the **Poisson equation**. The types of boundary conditions mentioned above are also imposed on the Poisson equation. We will also study this equation in this and the next chapter.

9.1 Harmonic Functions

The following proposition is a consequence of the Divergence Theorem.

Proposition 9.1.1 (Green's Identities)

Let Ω be an open subset of $\mathbb{R}^n$ with a smooth boundary (at least a boundary of class C^2). If u and v are two functions of class $C^2(\Omega) \cap C^1(\overline{\Omega})$, then

$$\int_{\partial\Omega} v \frac{\partial u}{\partial \nu} dS = \int_{\Omega} (v \Delta u + \nabla u \cdot \nabla v) d\mathbf{x} \quad (9.1.1)$$

and

$$\int_{\partial\Omega} \left(v \frac{\partial u}{\partial \nu} - u \frac{\partial v}{\partial \nu} \right) dS = \int_{\Omega} (v \Delta u - u \Delta v) d\mathbf{x}, \quad (9.1.2)$$

where $\nu(\mathbf{x})$ is the outward unit normal to the surface $\partial\Omega$ at $\mathbf{x} \in \partial\Omega$.

Remark 9.1.2

Recall that a function $f : \Omega \rightarrow \mathbb{R}$ is of class $C^k(\overline{\Omega})$ if $D^\alpha f$ can be continuously extended to $\overline{\Omega}$ for all multi-index α such that $|\alpha| \leq k$. ♠

Proof.

The Divergence Theorem states that

$$\int_{\partial\Omega} F \cdot \nu dS = \int_{\Omega} \operatorname{div} F d\mathbf{x}$$

for $F : \Omega \rightarrow \mathbb{R}^n$ of class $C^1(\Omega) \cap C(\overline{\Omega})$ and ν a unite normal to $\partial\Omega$ pointing outside of Ω .

Moreover, $\frac{\partial f}{\partial \nu} = \nabla f \cdot \nu$ for $f \in C^1(\Omega)$. Thus (9.1.1) follows from the Divergence Theorem applied to $F = v \nabla u$. To get (9.1.2), interchange u and v in (9.1.1) and subtract the resulting equation from (9.1.1). ■

Recall that, for $\mathbf{x} \in \mathbb{R}^n$ and $r > 0$, the open ball of radius r centred at $\mathbf{x}$ is the set $B_r(\mathbf{x}) = \{\mathbf{y} \in \mathbb{R}^n : \|\mathbf{y} - \mathbf{x}\| < r\}$. The closure of this ball is thus $\overline{B_r(\mathbf{x})} = \{\mathbf{y} \in \mathbb{R}^n : \|\mathbf{y} - \mathbf{x}\| \leq r\}$.

Definition 9.1.3

A continuous function $u : \Omega \rightarrow \mathbb{R}$ on an open set $\Omega \subset \mathbb{R}^n$ is **harmonic** if

$$u(\mathbf{x}) = M_u(\mathbf{x}, r) \equiv \frac{1}{\omega_n} \int_{\|\mathbf{y}\|=1} u(\mathbf{x} + r\mathbf{y}) dS_{\mathbf{y}} = \frac{1}{r^{n-1}\omega_n} \int_{\|\mathbf{y}-\mathbf{x}\|=r} u(\mathbf{y}) dS_{\mathbf{y}}$$

for every $\mathbf{x} \in \Omega$ and $r > 0$ such that $\overline{B_r(\mathbf{x})} \subset \Omega$, where $\omega_n = 2\pi^{n/2}/\Gamma(n/2)$ is the area of the surface of the ball of radius 1 in $\mathbb{R}^n$.

Theorem 9.1.4 (Mean Value Theorem)

Suppose that Ω is an open subset of $\mathbb{R}^n$ and that $u \in C^2(\Omega)$ satisfies $\Delta u = 0$ on Ω . If $\mathbf{x} \in \Omega$ and $r > 0$ are such that $\overline{B_r(\mathbf{x})} \subset \Omega$, then $u(\mathbf{x}) = M_u(\mathbf{x}, r)$. In particular, this prove that u is harmonic on Ω .

Proof.

It follows from the Green's identity (9.1.1) on $\overline{B_s(\mathbf{x})}$ with $v \equiv 1$ that

$$\int_{\partial B_s(\mathbf{x})} \frac{\partial u}{\partial \nu} dS = \int_{B_s(\mathbf{x})} \Delta u d\mathbf{x} = 0 \quad (9.1.3)$$

for all $0 \leq s \leq r$. Since u and ∇u are continuous on the compact set $\overline{B_r(\mathbf{x})}$ and so uniformly continuous on the compact set $\overline{B_r(\mathbf{x})}$, we may write

$$\begin{aligned} \frac{\partial}{\partial s} \int_{\partial B_1(\mathbf{0})} u(\mathbf{x} + s\mathbf{y}) dS_{\mathbf{y}} &= \int_{\partial B_1(\mathbf{0})} \frac{\partial}{\partial s} u(\mathbf{x} + s\mathbf{y}) dS_{\mathbf{y}} = \int_{\partial B_1(\mathbf{0})} \nabla u(\mathbf{x} + s\mathbf{y}) \cdot \mathbf{y} dS_{\mathbf{y}} \\ &= s^{1-n} \int_{\partial B_s(\mathbf{x})} \nabla u(\mathbf{y}) \cdot (s^{-1}(\mathbf{y} - \mathbf{x})) dS_{\mathbf{y}} = s^{1-n} \int_{\partial B_s(\mathbf{x})} \frac{\partial u}{\partial \nu} dS = 0 \end{aligned}$$

for all $0 < s < r$, where we note that the unit normal vector to $\partial B_s(\mathbf{x})$ at $\mathbf{y} \in \partial B_s(\mathbf{x})$ is $\nu(\mathbf{y}) = s^{-1}(\mathbf{y} - \mathbf{x})$. Therefore, $s \mapsto \int_{\partial B_1(\mathbf{0})} u(\mathbf{x} + s\mathbf{y}) dS_{\mathbf{y}}$ is a constant function for $0 \leq s \leq r$.

Thus,

$$\int_{\partial B_1(\mathbf{0})} u(\mathbf{x} + r\mathbf{y}) dS_{\mathbf{y}} = \int_{\partial B_1(\mathbf{0})} u(\mathbf{x}) dS_{\mathbf{y}} = \omega_n u(\mathbf{x}) . \quad \blacksquare$$

Theorem 9.1.4 has a converse.

Theorem 9.1.5

Let Ω be an open subset of $\mathbb{R}^n$ and $u : \Omega \rightarrow \mathbb{R}$ be an harmonic function, then $u \in C^\infty(\Omega)$ and $\Delta u = 0$ on Ω .

Proof.

i) We first prove that u is of class $C^\infty(\Omega)$.

Choose $\psi \in \mathcal{D}([-1, 1])$ such that

$$\int_{\mathbb{R}^n} \phi(\mathbf{x}) d\mathbf{x} = 1 , \quad (9.1.4)$$

where $\phi(\mathbf{x}) = \psi(\|\mathbf{x}\|^2)$ for $\mathbf{x} \in \mathbb{R}^n$. We note that $\text{supp } \phi \subset B_1(\mathbf{0})$.

Given $\epsilon \in]0, 1[$, let $\phi_\epsilon(\mathbf{x}) = \epsilon^{-n} \phi(\epsilon^{-1} \mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^n$ and $\Omega_\epsilon = \{\mathbf{x} : \overline{B_\epsilon(\mathbf{x})} \subset \Omega\}$. For $\mathbf{x} \in \Omega_\epsilon$, the support of the function $\mathbf{y} \mapsto \phi_\epsilon(\mathbf{x} - \mathbf{y})$ is a subset of $B_\epsilon(\mathbf{x}) \subset \Omega$ and $\int_{\mathbb{R}^n} u(\mathbf{y}) \phi_\epsilon(\mathbf{x} - \mathbf{y}) d\mathbf{y}$ is well defined. The integral outside of Ω is irrelevant because $\phi_\epsilon(\mathbf{x} - \mathbf{y}) = 0$ for $\mathbf{y} \notin \Omega$ by construction. One can set $u(\mathbf{y}) = 0$ for $\mathbf{y} \notin \Omega$ to get an integral defined on all $\mathbb{R}^n$.

For $\mathbf{x} \in \Omega_\epsilon$, we have

$$\begin{aligned} \int_{\mathbb{R}^n} u(\mathbf{y}) \phi_\epsilon(\mathbf{x} - \mathbf{y}) d\mathbf{y} &= \int_{\mathbb{R}^n} u(\mathbf{x} - \mathbf{y}) \phi_\epsilon(\mathbf{y}) d\mathbf{y} = \int_{B_\epsilon(\mathbf{0})} u(\mathbf{x} - \mathbf{y}) \epsilon^{-n} \phi(\epsilon^{-1} \mathbf{y}) d\mathbf{y} \\ &= \int_{B_1(\mathbf{0})} u(\mathbf{x} - \epsilon \mathbf{y}) \phi(\mathbf{y}) d\mathbf{y} = \int_0^1 \left(\int_{\|\mathbf{y}\|=1} u(\mathbf{x} - \epsilon r \mathbf{y}) dS_{\mathbf{y}} \right) \psi(r^2) r^{n-1} dr, \end{aligned}$$

where we have used the fact that $\phi(r\mathbf{y}) = \psi(r^2)$ for $\|\mathbf{y}\| = 1$. However, by hypothesis,

$$\int_{\|\mathbf{y}\|=1} u(\mathbf{x} - \epsilon r \mathbf{y}) dS_{\mathbf{y}} = \omega_n u(\mathbf{x})$$

because $\epsilon r \leq \epsilon$. Thus,

$$\begin{aligned} \int_{\mathbb{R}^n} u(\mathbf{y}) \phi_\epsilon(\mathbf{x} - \mathbf{y}) d\mathbf{y} &= \omega_n u(\mathbf{x}) \int_0^1 \psi(r^2) r^{n-1} dr = u(\mathbf{x}) \int_0^1 \left(\int_{\|\mathbf{y}\|=1} dS_{\mathbf{y}} \right) \psi(r^2) r^{n-1} dr \\ &= u(\mathbf{x}) \int_0^1 \int_{\|\mathbf{y}\|=1} \phi(r\mathbf{y}) r^{n-1} dS(\mathbf{y}) dr = u(\mathbf{x}) \int_{B_1(\mathbf{0})} \phi(\mathbf{y}) d\mathbf{y} = u(\mathbf{x}), \end{aligned}$$

where the last equality comes from (9.1.4) with $\text{supp } \phi \subset B_1(\mathbf{0})$.

We have that

$$\begin{aligned} \Omega_\epsilon &\rightarrow \mathbb{R} \\ \mathbf{x} &\mapsto \int_{\mathbb{R}^n} u(\mathbf{y}) \phi_\epsilon(\mathbf{x} - \mathbf{y}) d\mathbf{y} \end{aligned}$$

is of class C^∞ on Ω_ϵ because $\phi_\epsilon \in C_c^\infty(\mathbb{R}^n)$ and, for any $\mathbf{x}_0 \in \Omega_\epsilon$ and $\delta > 0$ such that $\overline{B_\delta(\mathbf{x}_0)} \subset \Omega_\epsilon$, the function

$$\begin{aligned} h_\alpha &: \Omega_\epsilon \times \Omega \rightarrow \mathbb{R} \\ (\mathbf{x}, \mathbf{y}) &\mapsto u(\mathbf{y}) D_{\mathbf{x}}^\alpha \phi_\epsilon(\mathbf{x} - \mathbf{y}) \end{aligned}$$

is uniformly continuous on $\overline{B_\delta(\mathbf{x}_0)} \times \Omega$ for all multi-index α . This follows from the fact that $h_\alpha : \Omega_\epsilon \times \Omega$ is continuous, so it is uniformly continuous on the compact set $\overline{B_\delta(\mathbf{x}_0)} \times \overline{B_{\delta+\epsilon}(\mathbf{x}_0)}$, h_α . Since $h_\alpha(\mathbf{x}, \mathbf{y}) = 0$ for $\mathbf{x} \in \overline{B_\delta(\mathbf{x}_0)}$ and $\mathbf{y} \notin \overline{B_{\delta+\epsilon}(\mathbf{x}_0)}$, h_α is uniformly continuous on $\overline{B_\delta(\mathbf{x}_0)} \times \Omega$.

ii) Since

$$\int_{\|\mathbf{y}\|=1} u(\mathbf{x} + r\mathbf{y}) dS_{\mathbf{y}} = \omega_n u(\mathbf{x})$$

for all r such that $\overline{B_r(\mathbf{x})} \subset \Omega$ by hypothesis, then the function defined by

$r \mapsto \int_{\|\mathbf{y}\|=1} u(\mathbf{x} + r\mathbf{y}) dS_{\mathbf{y}}$ is constant on its domain. Hence, using the substitution $\mathbf{y} =$

$r^{-1}(\mathbf{z} - \mathbf{x})$, we get

$$\begin{aligned} 0 &= \frac{\partial}{\partial r} \int_{\|\mathbf{y}\|=1} u(\mathbf{x} + r\mathbf{y}) \, dS_{\mathbf{y}} = \int_{\|\mathbf{y}\|=1} \nabla u(\mathbf{x} + r\mathbf{y}) \cdot \mathbf{y} \, dS_{\mathbf{y}} \\ &= \int_{\partial B_r(\mathbf{x})} \nabla u(\mathbf{z}) \cdot (r^{-1}(\mathbf{z} - \mathbf{x})) \, r^{1-n} \, dS_{\mathbf{z}} = r^{1-n} \int_{\partial B_r(\mathbf{x})} \frac{\partial u}{\partial \nu} \, dS, \end{aligned}$$

where we note that the outward unit normal vector to $\partial B_r(\mathbf{x})$ at $\mathbf{z} \in \partial B_r(\mathbf{x})$ is $\nu(\mathbf{z}) = r^{-1}(\mathbf{z} - \mathbf{x})$.

It follows from (9.1.1) with $\Omega = B_r(\mathbf{x})$ and $v = 1$ on Ω that

$$0 = r^{1-n} \int_{\partial B_r(\mathbf{x})} \frac{\partial u}{\partial \nu} \, dS = r^{1-n} \int_{B_r(\mathbf{x})} \Delta u \, d\mathbf{y}.$$

Since this is true for all $\mathbf{x} \in \Omega$ and all r such that $\overline{B_r(\mathbf{x})} \subset \Omega$, and since Δu is continuous in Ω by the first part of the proof, then $\Delta u = 0$ on Ω . ■

Corollary 9.1.6

Let $\{u_j\}_{j=1}^{\infty}$ be a sequence of harmonic functions on an open set $\Omega \subset \mathbb{R}^n$. If the sequence $\{u_j\}_{j=1}^{\infty}$ converges uniformly on compact subsets of Ω to a function u , then u is harmonic in Ω .

Proof.

Since the u_j are harmonic in Ω , we have by definition that

$$u_j(\mathbf{x}) = \frac{1}{\omega_n} \int_{\|\mathbf{y}\|=1} u_j(\mathbf{x} + r\mathbf{y}) \, dS_{\mathbf{y}}$$

for $j > 0$, $\mathbf{x} \in \Omega$ and r small enough such that $\overline{B_r(\mathbf{x})} \subset \Omega$. The left side converges to $u(\mathbf{x})$ as $j \rightarrow \infty$ by hypothesis and the right side converges to $\frac{1}{\omega_n} \int_{\|\mathbf{y}\|=1} u(\mathbf{x} + r\mathbf{y}) \, dS_{\mathbf{y}}$ because $u_j(\mathbf{x} + r\mathbf{y}) \rightarrow u(\mathbf{x} + r\mathbf{y})$ uniformly on the compact set $\{\mathbf{y} : \|\mathbf{y}\| = 1\}$.

Being the uniform limit on compact sets of the continuous functions u_j , $u \in C(\Omega)$.

Thus u is a continuous function on Ω such that

$$u(\mathbf{x}) = \frac{1}{\omega_n} \int_{\|\mathbf{y}\|=1} u(\mathbf{x} + r\mathbf{y}) \, dS_{\mathbf{y}}$$

for any $\mathbf{x} \in \Omega$ and r such that $\overline{B_r(\mathbf{x})} \subset \Omega$. This means that u is harmonic on Ω by definition. ■

It is also interesting to note that Liouville's Theorem for complex entire (and so harmonic) functions generalizes to harmonic functions on $\mathbb{R}^n$ in general.

Theorem 9.1.7 (Liouville's Theorem)

If $u : \mathbb{R}^n \rightarrow \mathbb{R}$ is a bounded harmonic function, then u is constant on $\mathbb{R}^n$.

Proof.

Given $\mathbf{x} \in \mathbb{R}^n$ and $R > \|\mathbf{x}\|$, we multiply both sides of

$$u(\mathbf{x}) = \frac{1}{\omega_n} \int_{\|\mathbf{y}\|=1} u(\mathbf{x} + \rho R \mathbf{y}) dS_{\mathbf{y}}$$

by ρ^{n-1} and integrate with respect to ρ between 0 and 1 to get

$$\begin{aligned} u(\mathbf{x}) \int_0^1 \rho^{n-1} d\rho &= \frac{1}{\omega_n} \int_{\|\mathbf{y}\|=1} \int_0^1 u(\mathbf{x} + \rho R \mathbf{y}) \rho^{n-1} d\rho dS_{\mathbf{y}} \\ &= \frac{1}{R^n \omega_n} \int_{\|\mathbf{y}\|=1} \int_0^R u(\mathbf{x} + r \mathbf{y}) r^{n-1} dr dS_{\mathbf{y}} = \frac{1}{R^n \omega_n} \int_{B_R(\mathbf{x})} u(\mathbf{y}) d\mathbf{y}, \end{aligned} \quad (9.1.5)$$

where we have used the change of variable $r = \rho R$ to get the second equality. Since $\int_0^1 \rho^{n-1} = \frac{1}{n}$, we get from (9.1.5) that

$$u(\mathbf{x}) = \frac{n}{R^n \omega_n} \int_{B_R(\mathbf{x})} u(\mathbf{y}) d\mathbf{y}.$$

Thus,

$$|u(\mathbf{x}) - u(\mathbf{0})| = \frac{n}{R^n \omega_n} \left| \int_{B_R(\mathbf{x})} u(\mathbf{y}) d\mathbf{y} - \int_{B_R(\mathbf{0})} u(\mathbf{y}) d\mathbf{y} \right| = \frac{n}{R^n \omega_n} \left| \int_D u(\mathbf{y}) d\mathbf{y} \right|,$$

where $D = (B_R(\mathbf{x}) \setminus B_R(\mathbf{0})) \cup (B_R(\mathbf{0}) \setminus B_R(\mathbf{x}))$ is the symmetric difference of the sets $B_R(\mathbf{x})$ and $B_R(\mathbf{0})$ (Figure 9.1). Since $D \subset D_R = \{\mathbf{y} \in \mathbb{R}^n : R - \|\mathbf{x}\| \leq \|\mathbf{y}\| \leq R + \|\mathbf{x}\|\}$, we get

$$\begin{aligned} |u(\mathbf{x}) - u(\mathbf{0})| &\leq \frac{n}{R^n \omega_n} \|u\|_{\infty} \int_D d\mathbf{y} \leq \frac{n}{R^n \omega_n} \|u\|_{\infty} \int_{D_R} d\mathbf{y} \\ &= \frac{n}{R^n \omega_n} \|u\|_{\infty} \int_{\|\mathbf{y}\|=1} \int_{R-\|\mathbf{x}\|}^{R+\|\mathbf{x}\|} r^{n-1} dr dS_{\mathbf{y}} = \|u\|_{\infty} \frac{(R + \|\mathbf{x}\|)^n - (R - \|\mathbf{x}\|)^n}{R^n} \rightarrow 0 \end{aligned}$$

as $R \rightarrow \infty$. Thus $u(\mathbf{x}) = u(\mathbf{0})$. ■

9.2 Maximum Principle

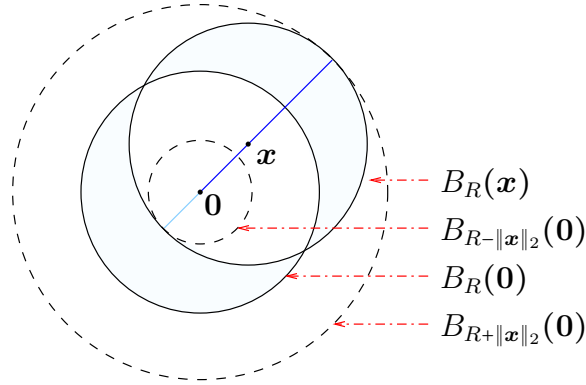


Figure 9.1: The region D described in the proof of Liouville's Theorem is in pale blue.

Definition 9.2.1

A continuous function $u : \Omega \rightarrow \mathbb{R}$ on an open set $\Omega \subset \mathbb{R}^n$ is **subharmonic** if, for every $\mathbf{x} \in \Omega$, there exists $R > 0$ (that may depend on $\mathbf{x}$) such that $B_R(\mathbf{x}) \subset \Omega$ and

$$u(\mathbf{x}) \leq M_u(\mathbf{x}, r) \equiv \frac{1}{\omega_n} \int_{\|\mathbf{y}\|=1} u(\mathbf{x} + r\mathbf{y}) \, dS_{\mathbf{y}} = \frac{1}{r^{n-1}\omega_n} \int_{\|\mathbf{y}-\mathbf{x}\|=1} u(\mathbf{y}) \, dS_{\mathbf{y}}$$

for all $0 < r < R$.

Any harmonic functions on an open set Ω is obviously subharmonic on Ω .

By slightly modifying the proof of Theorem 9.1.4, we obtain the following result.

Proposition 9.2.2

A function $u : \Omega \rightarrow \mathbb{R}$ of class C^2 on an open set $\Omega \subset \mathbb{R}^n$ is subharmonic if and only if $\Delta u \geq 0$ on Ω .

Proof.

$\Leftarrow$) This is basically the proof given for Theorem 9.1.4. The only difference is that (9.1.3) shows that $\int_{\partial B_s(\mathbf{x})} \frac{\partial u}{\partial \nu} \, dS_{\mathbf{y}} \geq 0$. It then follows that $s \mapsto \int_{\partial B_1(\mathbf{0})} u(\mathbf{x} + s\mathbf{y}) \, dS_{\mathbf{y}}$ is a non decreasing function for $0 \leq s \leq r$. Thus, $u(\mathbf{x}) = M_u(\mathbf{x}, 0) \leq M_u(\mathbf{x}, s)$ for $0 < s \leq r$.

$\Rightarrow$) Suppose that $\Delta u(\mathbf{z}) < 0$ for some $\mathbf{z} \in \Omega$. Since Δu is continuous, there exists $r > 0$ such that $\overline{B_r(\mathbf{z})} \subset \Omega$ and $\Delta u(\mathbf{x}) < 0$ for all $\mathbf{x} \in \overline{B_r(\mathbf{z})}$.

Since u and ∇u are uniformly continuous on the compact set $\overline{B_r(\mathbf{z})}$, we may write

$$\frac{\partial}{\partial s} \int_{\partial B_1(\mathbf{0})} u(\mathbf{z} + s\mathbf{y}) \, dS_{\mathbf{y}} = \int_{\partial B_1(\mathbf{0})} \frac{\partial}{\partial s} u(\mathbf{z} + s\mathbf{y}) \, dS_{\mathbf{y}} = s^{1-n} \int_{\partial B_s(\mathbf{x})} \frac{\partial u}{\partial \nu} \, dS$$

for $0 < s < r$. It then follows from Green's identity (9.1.1) on $B_x(\mathbf{x})$ with $v \equiv 1$ that

$$\frac{\partial}{\partial s} \int_{\partial B_1(\mathbf{0})} u(\mathbf{z} + s\mathbf{y}) dS_{\mathbf{y}} = s^{1-n} \int_{\partial B_s(\mathbf{x})} \frac{\partial u}{\partial \nu} dS = s^{1-n} \int_{B_s(\mathbf{x})} \Delta u d\mathbf{y} < 0$$

for $0 < s < r$. Hence, $s \mapsto M_u(\mathbf{z}, s)$ is a decreasing function on $[0, r]$. This implies that $M_u(\mathbf{z}, s) < M_u(\mathbf{z}, 0) = u(\mathbf{z})$ for $0 < s < r$. This is a contradiction that u is subharmonic on Ω . ■

Theorem 9.2.3 (Maximum Principle)

Let Ω be a connected open subset of $\mathbb{R}^n$ and $u : \overline{\Omega} \rightarrow \mathbb{R}$ be continuous function which is subharmonic on Ω . Moreover, let $A = \sup_{\mathbf{x} \in \Omega} u(\mathbf{x})$. If $A < \infty$, then either $u(\mathbf{x}) < A$ for all $\mathbf{x} \in \Omega$ or $u(\mathbf{x}) = A$ for all $\mathbf{x} \in \Omega$.

Remark 9.2.4

Recall that a set $\Omega \subset \mathbb{R}^n$ is **connected** if it cannot be expressed as the disjoint union of two non-empty open sets with respect to the induced topology on Ω . ♠

Proof.

The set $u^{-1}(\{A\}) = \{\mathbf{x} \in \Omega : u(\mathbf{x}) = A\}$ is closed in Ω because u is continuous (the inverse image of a closed set by a continuous function is a closed set).

We now show that $u^{-1}(\{A\})$ is also open in the induced topology on Ω . If $\mathbf{x} \in u^{-1}(\{A\})$, it follows that

$$A = u(\mathbf{x}) \leq \frac{1}{r^{n-1}\omega_n} \int_{\|\mathbf{y}-\mathbf{x}\|=r} u(\mathbf{y}) dS_{\mathbf{y}}$$

for r such that $\overline{B_r(\mathbf{x})} \subset \Omega$ because u is subharmonic on Ω . There exists such a r since Ω is open. Since A is the maximum of u in Ω and u is continuous in Ω , this implies that $u(\mathbf{y}) = A$ for all $\mathbf{y} \in B_r(\mathbf{x})$. Hence, $B_r(\mathbf{x})$ is an open neighbourhood of $\mathbf{x}$ inside $u^{-1}(\{A\})$. This shows that $u^{-1}(\{A\})$ is open since $\mathbf{x} \in u^{-1}(\{A\})$ is arbitrary.

Since Ω is connected, either $u^{-1}(\{A\}) = \emptyset$ or $u^{-1}(\{A\}) = \Omega$ otherwise $\Omega = u^{-1}(\{A\}) \cup (\Omega \setminus u^{-1}(\{A\}))$ is the disjoint union of two non-empty open sets in Ω . ■

Corollary 9.2.5

Let Ω be a bounded open subset of $\mathbb{R}^n$. If $u : \overline{\Omega} \rightarrow \mathbb{R}$ is continuous and u is harmonic in Ω , then u reaches its maximal (and minimal) value on $\partial\Omega$.

Proof.

Since u is continuous on the compact set $\overline{\Omega}$, there exists $\mathbf{x} \in \overline{\Omega}$ such that $u(\mathbf{x}) = \sup_{\mathbf{y} \in \overline{\Omega}} u(\mathbf{y})$. If

$\mathbf{x} \in \Omega$, it follows from Theorem 9.2.3 that $u(\mathbf{y}) = u(\mathbf{x})$ for all $\mathbf{y} \in \tilde{\Omega}$, a connected component of Ω . Thus $u(\mathbf{y}) = u(\mathbf{x})$ for $\mathbf{y} \in \partial\tilde{\Omega} \subset \partial\Omega$.

To prove the conclusion of the corollary for the minimal value, apply the previous reasoning to $-u$. ■

Theorem 9.2.6 (Uniqueness Theorem)

Let Ω be an open subset of $\mathbb{R}^n$ with compact closure. If $u_1 : \overline{\Omega} \rightarrow \mathbb{R}$ and $u_2 : \overline{\Omega} \rightarrow \mathbb{R}$ are two continuous functions which are harmonic on Ω and satisfy $u_1 = u_2$ on $\partial\Omega$, then $u_1 = u_2$ in Ω .

Proof.

Since $u_1 - u_2$ is continuous functions on $\overline{\Omega}$ and harmonic in Ω , we get from Corollary 9.2.5 that

$$\begin{aligned} 0 &= \inf_{\mathbf{x} \in \partial\Omega} (u_1(\mathbf{x}) - u_2(\mathbf{x})) = \inf_{\mathbf{x} \in \overline{\Omega}} (u_1(\mathbf{x}) - u_2(\mathbf{x})) \leq \sup_{\mathbf{x} \in \overline{\Omega}} (u_1(\mathbf{x}) - u_2(\mathbf{x})) \\ &= \sup_{\mathbf{x} \in \partial\Omega} (u_1(\mathbf{x}) - u_2(\mathbf{x})) = 0 . \end{aligned}$$

Thus $u_1(\mathbf{x}) - u_2(\mathbf{x}) = 0$ for all $\mathbf{x} \in \Omega$. ■

9.3 Fundamental Solutions

Theorem 9.3.1

Let

$$N(\mathbf{x}) = \begin{cases} \frac{1}{(2-n)\omega_n} \|\mathbf{x}\|^{2-n} & \text{if } n > 2 \\ \frac{1}{\omega_2} \ln(\|\mathbf{x}\|) & \text{if } n = 2 \end{cases}$$

for $\mathbf{x} \neq \mathbf{0}$. The function N is a fundamental solution of the Laplace operator $L(\mathbf{x}, D) = \Delta$.

Proof.

i) Though N is not defined at the origin, N is absolutely integrable on any ball $B_r(\mathbf{0})$ with $0 < r < \infty$.

For $n > 2$,

$$\begin{aligned} \int_{B_r(\mathbf{0})} |N(\mathbf{x})| d\mathbf{x} &= \frac{1}{(n-2)\omega_n} \int_{\|\mathbf{y}\|_2=1} \left(\int_0^r s^{2-n} s^{n-1} ds \right) dS_{\mathbf{y}} = \frac{1}{(n-2)} \int_0^r s ds \\ &= \frac{s^2}{2(n-2)} \Big|_0^r = \frac{r^2}{2(n-2)} < \infty . \end{aligned}$$

For $n = 2$ and $r > 1$,

$$\begin{aligned} \int_{B_r(\mathbf{0})} |N(\mathbf{x})| d\mathbf{x} &= \frac{1}{\omega_2} \int_{\|\mathbf{y}\|_2=1} \left(\int_0^r |\ln(s)| s ds \right) dS_{\mathbf{y}} = \int_0^r s |\ln(s)| ds \\ &= - \lim_{\epsilon \rightarrow 0^+} \int_{\epsilon}^1 s \ln(s) ds + \int_1^r s \ln(s) ds = - \lim_{\epsilon \rightarrow 0} \left(\frac{s^2}{2} \ln(s) - \frac{s^2}{4} \right) \Big|_{\epsilon}^1 + \left(\frac{s^2}{2} \ln(s) - \frac{s^2}{4} \right) \Big|_1^r \\ &= \frac{r^2}{2} \ln(r) - \frac{r^2}{2} + \frac{1}{2} < \infty. \end{aligned}$$

The case $0 < r < 1$ is left to the reader. Hence, N is in $L^1_{loc}(\mathbb{R}^n)$ and thus in $\mathcal{D}'(\mathbb{R}^n)$.

ii) For $\epsilon > 0$, let

$$N_{\epsilon}(\mathbf{x}) = \begin{cases} \frac{1}{(2-n)\omega_n} (\|\mathbf{x}\|^2 + \epsilon^2)^{(2-n)/2} & \text{if } n > 2 \\ \frac{1}{2\omega_2} \ln(\|\mathbf{x}\|^2 + \epsilon^2) & \text{if } n = 2 \end{cases}$$

Note that N_{ϵ} is of class C^{∞} on $\mathbb{R}^n$.

We first prove that $N_{\epsilon} \rightarrow N$ in $L^1(K)$ as $\epsilon \rightarrow 0$ for any compact set $K \subset \mathbb{R}^n$. It is obvious that $\lim_{\epsilon \rightarrow 0} N_{\epsilon}(\mathbf{x}) = N(\mathbf{x})$ for $\mathbf{x} \neq \mathbf{0}$. Moreover, for $\epsilon \leq 1$,

$$|N_{\epsilon}(\mathbf{x})| < \begin{cases} |N(\mathbf{x})| & \text{if } n > 2 \\ \frac{1}{\omega_2} (|\ln(\|\mathbf{x}\|)| + 1) & \text{if } n = 2 \end{cases}$$

for all $\mathbf{x} \neq \mathbf{0}$ ¹. In both cases, the dominating function is integrable on K (even if $\mathbf{0} \in K$). It follows from the Lebesgue Dominate Convergence Theorem that

$$\int_K |N_{\epsilon}(\mathbf{x}) - N(\mathbf{x})| d\mathbf{x} \rightarrow 0$$

as $\epsilon \rightarrow 0$.

We now prove that $N_{\epsilon} \rightarrow N$ in $\mathcal{D}'(\mathbb{R}^n)$ as $\epsilon \rightarrow 0$. If $\phi \in \mathcal{D}(\mathbb{R}^n)$, let $K = \text{supp } \phi$. K is a compact set by definition of $\mathcal{D}(\mathbb{R}^n)$. Hence

$$\begin{aligned} |\langle N_{\epsilon}, \phi \rangle - \langle N, \phi \rangle| &= |\langle N_{\epsilon} - N, \phi \rangle| = \left| \int_K (N_{\epsilon}(\mathbf{x}) - N(\mathbf{x})) \phi(\mathbf{x}) d\mathbf{x} \right| \\ &\leq \sup_{\mathbf{x} \in K} |\phi(\mathbf{x})| \int_K |N_{\epsilon}(\mathbf{x}) - N(\mathbf{x})| d\mathbf{x} \rightarrow 0 \end{aligned}$$

¹If $\|\mathbf{x}\|^2 + \epsilon^2 \leq 1$, then $(1/2)|\ln(\|\mathbf{x}\|^2 + \epsilon^2)| \leq (1/2)|\ln(\|\mathbf{x}\|^2)| = |\ln(\|\mathbf{x}\|)|$. If $\|\mathbf{x}\| + \epsilon > 1$ and $\epsilon \leq 1$, then

$$\begin{aligned} (1/2)|\ln(\|\mathbf{x}\|^2 + \epsilon^2)| &= (1/2) \ln(\|\mathbf{x}\|^2 + \epsilon^2) \leq (1/2) \ln(\|\mathbf{x}\|^2 + 1) \\ &\leq \begin{cases} (1/2) \ln(2) \leq 1 & \text{if } \|\mathbf{x}\| \leq 1 \\ (1/2)(\ln(\|\mathbf{x}\|^2) + 1) = \ln(\|\mathbf{x}\|) + 1/2 & \text{if } \|\mathbf{x}\| > 1 \end{cases} \end{aligned}$$

It is easier to look at the graph of $|\ln(x)|$ to convince yourself of these inequalities.

as $\epsilon \rightarrow 0$ because $N_\epsilon \rightarrow N$ in $L^1(K)$ as $\epsilon \rightarrow 0$ and $\sup_{\mathbf{x} \in K} |\phi(\mathbf{x})|$ is bounded.

iii) We have

$$\Delta N_\epsilon(\mathbf{x}) = \frac{\partial^2 N_\epsilon}{\partial r^2}(r) + \frac{n-1}{r} \frac{\partial N_\epsilon}{\partial r}(r) = \frac{n\epsilon^2}{\omega_n} (r^2 + \epsilon^2)^{-(n+2)/2} = \frac{n\epsilon^2}{\omega_n} (\|\mathbf{x}\|^2 + \epsilon^2)^{-(n+2)/2}$$

for $r = \|\mathbf{x}\|$. Thus $\Delta N_\epsilon(\mathbf{x}) = \epsilon^{-n} \psi(\epsilon^{-1} \mathbf{x})$, where

$$\psi(\mathbf{x}) = \Delta N_1(\mathbf{x}) = \frac{n}{\omega_n} (\|\mathbf{x}\|^2 + 1)^{-(n+2)/2}$$

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

Since $N_\epsilon(\mathbf{x}) = N_\epsilon(-\mathbf{x})$ for all $\mathbf{x}$, we get

$$\langle \Delta N_\epsilon, \phi \rangle = \int_{\mathbb{R}^n} \Delta N_\epsilon(\mathbf{x}) \phi(\mathbf{x}) d\mathbf{x} = \int_{\mathbb{R}^n} \Delta N_\epsilon(-\mathbf{x}) \phi(\mathbf{x}) d\mathbf{x} = (\phi * \Delta N_\epsilon)(\mathbf{0})$$

for $\phi \in \mathcal{D}(\mathbb{R}^n)$.

We have that $\psi \in L^1(\mathbb{R}^n)$. In fact, $\psi(\mathbf{x}) \geq 0$ for all $\mathbf{x}$ and

$$\begin{aligned} a &= \int_{\mathbb{R}^n} \psi(\mathbf{x}) d\mathbf{x} = \frac{n}{\omega_n} \int_{\|\mathbf{y}\|=1} \int_0^\infty (r^2 + 1)^{-(2+n)/2} r^{n-1} dr dS_{\mathbf{y}} \\ &= n \int_0^\infty (r^2 + 1)^{-(2+n)/2} r^{n-1} dr = n \int_0^\infty \left(\frac{r^2}{r^2 + 1} \right)^{(n-2)/2} \frac{r}{(r^2 + 1)^2} dr \\ &= \frac{n}{2} \int_0^1 s^{(n-2)/2} ds = s^{n/2} \Big|_0^1 = 1, \end{aligned}$$

where we have used the substitution $s = \frac{r^2}{r^2 + 1}$ and $ds = \frac{2r}{(r^2 + 1)^2} dr$.

We may therefore use Proposition 6.5.3 with $a = \int_{\mathbb{R}^n} \psi(\mathbf{x}) d\mathbf{x} = 1$ to conclude that

$$\langle \Delta N_\epsilon, \phi \rangle = (\phi * \Delta N_\epsilon)(\mathbf{0}) \rightarrow \phi(\mathbf{0}) \quad (9.3.1)$$

as $\epsilon \rightarrow 0$ for any $\phi \in \mathcal{D}(\mathbb{R}^n)$.

iv) Finally, from (ii) and (9.3.1), we have

$$\langle \Delta N, \phi \rangle = \lim_{\epsilon \rightarrow 0} \langle \Delta N_\epsilon, \phi \rangle = \phi(\mathbf{0}) = \delta(\phi)$$

for all $\phi \in \mathcal{D}(\mathbb{R}^n)$. ■

Corollary 9.3.2

If $u \in \mathcal{D}'(\mathbb{R}^n)$ satisfies $\Delta u = f \in C^\infty(\mathbb{R}^n)$, then $u \in C^\infty(\mathbb{R}^n)$. In particular, we recover the result that harmonic functions on $\mathbb{R}^n$ are of class $C^\infty(\mathbb{R}^n)$.

Proof.

It follows from Theorem 9.3.1 that the first sentence of the equivalence list of Theorem 6.7.11 is satisfied. $\blacksquare$

The distribution solution of $\Delta u = f \in \mathcal{E}'(\mathbb{R}^n)$ is given by $u = N * f$. The same conclusion may sometime be drawn when $f \notin \mathcal{E}'(\mathbb{R}^n)$.

Theorem 9.3.3

If $f \in L^1(\mathbb{R}^n)$ for $n > 2$ or $\int_{\mathbb{R}^n} |f(\mathbf{y}) \ln(\|\mathbf{y}\|)| d\mathbf{y} < \infty$ for $n = 2$, then $u = N * f$ is a locally integrable function satisfying $\Delta u = f$ in the sense of distributions.

Proof.

We provide only the proof for $n > 0$. Let

$$\chi_r(\mathbf{x}) = \begin{cases} 1 & \text{if } \|\mathbf{x}\| < r \\ 0 & \text{otherwise} \end{cases}$$

We have that $\chi_r N \in L^1(\mathbb{R}^n)$ because

$$\begin{aligned} \int_{\mathbb{R}^n} |\chi_r(\mathbf{y}) N(\mathbf{y})| d\mathbf{y} &= \frac{1}{(n-2)\omega_n} \int_{\|\mathbf{y}\| < r} \|\mathbf{y}\|^{2-n} d\mathbf{y} = \frac{1}{(n-2)\omega_n} \int_{\|\mathbf{y}\|=1} \int_0^r \rho d\rho dS_{\mathbf{y}} \\ &= \frac{r^2}{2(n-2)} < \infty. \end{aligned}$$

Moreover, $(1 - \chi_r)N \in L^\infty(\mathbb{R}^n)$ because $\mathbf{x} \mapsto N(\mathbf{x})$ is a decreasing function as $\|\mathbf{x}\| \rightarrow \infty$ with support in $\{\mathbf{x} : \|\mathbf{x}\| > r\}$. Thus, it follows from Young's Inequality, Theorem 6.5.2, that $(\chi_r N) * f \in L^1(\mathbb{R}^n)$ and $((1 - \chi_r)N) * f \in L^\infty(\mathbb{R}^n)$.

We first show that $\chi_r f \rightarrow f$ and $N * (\chi_r f) \rightarrow N * f$ in $\mathcal{D}'(\mathbb{R}^n)$ as $r \rightarrow \infty$.

It follows from Lebesgue Dominated Convergence Theorem that $\chi_r f$ converges to f in $L^1(\mathbb{R}^n)$ as $r \rightarrow \infty$ because $|\chi_r f| \leq |f| \in L^1(\mathbb{R}^n)$ and $\lim_{r \rightarrow \infty} \chi_r(\mathbf{x}) f(\mathbf{x}) = f(\mathbf{x})$ for all $\mathbf{x}$. Therefore, $\langle \chi_r f, \phi \rangle \rightarrow \langle f, \phi \rangle$ as $r \rightarrow \infty$ for all test function $\phi \in \mathcal{D}(\mathbb{R}^n)$; namely, $\chi_r f \rightarrow f$ in $\mathcal{D}'(\mathbb{R}^n)$ as $r \rightarrow \infty$.

Since

$$\|(\chi_1 N) * (\chi_r f - f)\|_1 \leq \|\chi_r f - f\|_1 \|\chi_1 N\|_1 \rightarrow 0$$

as $r \rightarrow \infty$ and

$$\|((1 - \chi_1)N) * (\chi_r f - f)\|_\infty \leq \|\chi_r f - f\|_1 \|(1 - \chi_1)N\|_\infty \rightarrow 0$$

as $r \rightarrow \infty$, we have for all $\phi \in \mathcal{D}(\mathbb{R}^n)$ that

$$\begin{aligned} |\langle N * (\chi_r f), \phi \rangle - \langle N * f, \phi \rangle| &= |\langle N * (\chi_r f - f), \phi \rangle| \\ &= |\langle (\chi_1 N) * (\chi_r f - f), \phi \rangle + \langle ((1 - \chi_1)N) * (\chi_r f - f), \phi \rangle| \end{aligned}$$

$$\begin{aligned}
&\leq \int_{\mathbb{R}^n} |(\chi_1 N) * (\chi_r f - f)| |\phi| d\mathbf{x} + \int_{\mathbb{R}^n} |((1 - \chi_1)N) * (\chi_r f - f)| |\phi| d\mathbf{x} \\
&\leq \|\phi\|_\infty \|(\chi_1 N) * (\chi_r f - f)\|_1 + \|((1 - \chi_1)N) * (\chi_r f - f)\|_\infty \int_{\mathbb{R}^n} |\phi| d\mathbf{x} \rightarrow 0
\end{aligned}$$

as $r \rightarrow \infty$. Thus, $N * (\chi_r f) \rightarrow N * f$ in $\mathcal{D}'(\mathbb{R}^n)$ as $r \rightarrow \infty$.

Finally, we have

$$\Delta(N * f) = \lim_{r \rightarrow \infty} \Delta(N * (\chi_r f)) = \lim_{r \rightarrow \infty} \chi_r f = f$$

in the sense of distributions. The first equality comes from the fact that Δ is continuous on $\mathcal{D}'(\mathbb{R}^n)$ and $N * (\chi_r f) \rightarrow N * f$ in $\mathcal{D}'(\mathbb{R}^n)$ as $r \rightarrow \infty$. The second equality comes from the fact that $\chi_r f \in \mathcal{E}'(\mathbb{R}^n)$ for all r . So $N * (\chi_r f)$ is a solution in the sense of distribution of $\Delta u = \chi_r f$. The last equality is just $\chi_r f \rightarrow f$ in $\mathcal{D}'(\mathbb{R}^n)$ as $r \rightarrow \infty$ that we have proved previously. ■

9.4 Dirichlet and Neumann Problems

Let

$$N(\mathbf{x}, \mathbf{y}) = N(\mathbf{x} - \mathbf{y}) \quad (9.4.1)$$

for $\mathbf{x} \neq \mathbf{y}$ in $\mathbb{R}^n$ and

$$N_\epsilon(\mathbf{x}, \mathbf{y}) = N_\epsilon(\mathbf{x} - \mathbf{y})$$

for $\mathbf{x}, \mathbf{y}$ in $\mathbb{R}^n$.

Theorem 9.4.1

Let $\Omega \subset \mathbb{R}^n$ be an open bounded set. If $u \in C^1(\overline{\Omega})$ is harmonic in Ω , then

$$u(\mathbf{x}) = \int_{\partial\Omega} \left(u(\mathbf{y}) \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) - N(\mathbf{x}, \mathbf{y}) \frac{\partial u}{\partial \nu_{\mathbf{y}}}(\mathbf{y}) \right) dS_{\mathbf{y}} \quad , \quad \mathbf{x} \in \Omega \quad , \quad (9.4.2)$$

where $\frac{\partial N}{\partial \nu_{\mathbf{y}}}$ denotes the directional derivative of $\mathbf{y} \mapsto N(\mathbf{x}, \mathbf{y})$ in the direction of the outward unit normal $\nu(\mathbf{y})$ to $\partial\Omega$ at $\mathbf{y}$

Proof.

Consider $\mathbf{x} \in \Omega$. Since $\Delta u = 0$ on Ω , we get from Green's identity (9.1.2) that

$$\int_{\Omega} u(\mathbf{y}) \Delta_{\mathbf{y}} N_\epsilon(\mathbf{x}, \mathbf{y}) d\mathbf{y} = \int_{\partial\Omega} \left(u(\mathbf{y}) \frac{\partial N_\epsilon}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) - N_\epsilon(\mathbf{x}, \mathbf{y}) \frac{\partial u}{\partial \nu_{\mathbf{y}}}(\mathbf{y}) \right) dS_{\mathbf{y}} \quad . \quad (9.4.3)$$

i) If we set $u(\mathbf{y}) = 0$ for $\mathbf{y} \in \mathbb{R}^n \setminus \overline{\Omega}$, the left hand side of (9.4.3) becomes $(u * \Delta N_\epsilon)(\mathbf{x})$ since $\Delta_{\mathbf{y}} N_\epsilon(\mathbf{x}, \mathbf{y}) = \Delta_{\mathbf{y}} N_\epsilon(\mathbf{x} - \mathbf{y}) = (\Delta N_\epsilon)(\mathbf{x} - \mathbf{y})$ for $\mathbf{y} \in \mathbb{R}^n$.

Since $u \in L^\infty(\mathbb{R}^n)$, u is uniformly continuous on $\overline{\Omega}$, $\Delta N_\epsilon(\mathbf{y}) = \epsilon^{-n} \Delta N_1(\epsilon^{-1}\mathbf{y})$ and $\int_{\mathbb{R}^n} \Delta N_1(\mathbf{y}) d\mathbf{y} = 1$ (proved in (iii) of the proof of Theorem 9.3.1), we can use Proposition 6.5.3 to conclude that $u * \Delta N_\epsilon \rightarrow u$ uniformly on any compact subset of Ω as $\epsilon \rightarrow 0$. In particular, $(u * \Delta N_\epsilon)(\mathbf{x}) \rightarrow u(\mathbf{x})$ as $\epsilon \rightarrow 0$.

ii) Since $\mathbf{x} \notin \partial\Omega$ and $\partial\Omega$ is a compact subset of $\mathbb{R}^n$, we have that

$$\left| u(\mathbf{y}) \frac{\partial N_\epsilon}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) \right| \leq \left| u(\mathbf{y}) \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) \right|$$

for all $\mathbf{y} \in \partial\Omega$ with $\mathbf{y} \mapsto u(\mathbf{y}) \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y})$ in $L^1(\partial\Omega)$. Moreover,

$$\lim_{\epsilon \rightarrow 0} u(\mathbf{y}) \frac{\partial N_\epsilon}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) = u(\mathbf{y}) \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y})$$

for all $\mathbf{y} \in \partial\Omega$. We have similar results for $N_\epsilon(\mathbf{x}, \mathbf{y}) \frac{\partial u}{\partial \nu_{\mathbf{y}}}(\mathbf{y})$; namely,

$$\left| N_\epsilon(\mathbf{x}, \mathbf{y}) \frac{\partial u}{\partial \nu_{\mathbf{y}}}(\mathbf{y}) \right| \leq h(\mathbf{x}, \mathbf{y}) \equiv \begin{cases} N(\mathbf{x}, \mathbf{y}) \frac{\partial u}{\partial \nu_{\mathbf{y}}}(\mathbf{y}) & \text{if } n > 2 \\ \frac{1}{\omega_2} (|\ln(\|\mathbf{x} - \mathbf{y}\|)| + 1) \frac{\partial u}{\partial \nu_{\mathbf{y}}}(\mathbf{y}) & \text{if } n = 2 \end{cases}$$

for all $\mathbf{y} \in \partial\Omega$ if $\epsilon \leq 1$ with $\mathbf{y} \mapsto h(\mathbf{x}, \mathbf{y})$ in $L^1(\partial\Omega)$. Moreover,

$$\lim_{\epsilon \rightarrow 0} N_\epsilon(\mathbf{x}, \mathbf{y}) \frac{\partial u}{\partial \nu_{\mathbf{y}}}(\mathbf{y}) = N(\mathbf{x}, \mathbf{y}) \frac{\partial u}{\partial \nu_{\mathbf{y}}}(\mathbf{y})$$

for all $\mathbf{y} \in \partial\Omega$.

It follows from the Lebesgue dominate convergence theorem that the right hand side of (9.4.3) converges to

$$\int_{\partial\Omega} \left(u(\mathbf{y}) \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) - N(\mathbf{x}, \mathbf{y}) \frac{\partial u}{\partial \nu_{\mathbf{y}}}(\mathbf{y}) \right) dS_{\mathbf{y}} .$$

■

Remark 9.4.2

Replacing $u|_{\partial\Omega}$ by f and $\frac{\partial u}{\partial \nu}|_{\partial\Omega}$ by g in (9.4.2) does not generally give a solution of $\Delta u = 0$

in Ω with the boundary conditions $u|_{\partial\Omega} = f$ and $\frac{\partial u}{\partial \nu}|_{\partial\Omega} = g$. There may not be a solution of $\Delta u = 0$ in Ω with such boundary conditions. This is a consequence of the uniqueness theorem, Theorem 9.2.6, which states that the solution is completely determined by the values of u on $\partial\Omega$. ♠

Corollary 9.4.3

Let Ω be an open bounded subset of $\mathbb{R}^n$. If $u \in \mathcal{D}'(\Omega)$ satisfies $\Delta u = f$, where f is real analytic on Ω , then u is real analytic on Ω .

Proof.

We give only a sketch of the proof. A complete proof can be found in [10, 39].

From Corollary 9.3.2, we have $u \in C^\infty(\Omega)$. Given $\mathbf{x}_0 \in \Omega$, we prove that u is analytic in a neighbourhood of $\mathbf{x}_0$.

According to Cauchy-Kowalevski Theorem 3.1.1, there exists a analytic solution v to $\Delta v = f$ in a neighbourhood $B_r(\mathbf{x}_0) \subset \Omega$ of $\mathbf{x}_0$.

Let $w = u - v$. We have that $\Delta w = 0$ in $B = B_{r/2}(\mathbf{x}_0)$ and $w \in C^\infty(\overline{B})$. Hence, from Theorem 9.4.1, we have

$$w(\mathbf{x}) = \int_{\partial B} \left(w(\mathbf{y}) \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) - N(\mathbf{x}, \mathbf{y}) \frac{\partial w}{\partial \nu_{\mathbf{y}}}(\mathbf{y}) \right) dS_{\mathbf{y}} \quad (9.4.4)$$

for $\mathbf{x} \in B$.

Given $\mathbf{y} \in \partial B$, the functions $\mathbf{x} \mapsto N(\mathbf{x}, \mathbf{y})$ and

$$\mathbf{x} \mapsto \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) = \frac{-1}{\omega_n \|\mathbf{x} - \mathbf{y}\|_2^n \|\mathbf{y} - \mathbf{x}_0\|_2} (\mathbf{x} - \mathbf{y}) \cdot (\mathbf{y} - \mathbf{x}_0)$$

are real analytic functions on $B_{r/2}(\mathbf{x}_0)$. We may extended these functions to complex analytic (and so holomorphic) functions for

$$\mathbf{x} \in W \equiv \left\{ \mathbf{x} \in \mathbb{C}^n : \|\operatorname{Re}(\mathbf{x}) - \mathbf{x}_0\| < \frac{r}{4} \text{ and } \|\operatorname{Im}(\mathbf{x})\| < \frac{r}{4} \right\}$$

More precisely, the real series expansion for $\mathbf{x} \mapsto N(\mathbf{x}, \mathbf{y})$ and $\mathbf{x} \mapsto \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y})$ can be used to defined the complex extension since

$$\|\mathbf{x} - \mathbf{x}_0\|_2 = \|\operatorname{Re}(\mathbf{x}) - \mathbf{x}_0\|_2 + \|\operatorname{Im}(\mathbf{x})\|_2 < \frac{r}{2}$$

for $\mathbf{x} \in W$. We can use the fact that the integrant in (9.4.4) is holomorphic for $\mathbf{x} \in W$ to extent w to an holomorphic function on W . Thus w is real analytic on $B_{r/4}(\mathbf{x}_0) = \mathbb{R}^n \cap W$. It follows that $u = v + w$ is real analytic on $B_{r/4}(\mathbf{x}_0)$. ■

Let $\Omega \subset \mathbb{R}^n$ be an open subset with a smooth boundary. Our goal is to find a solution (in the space of distributions) for the **Dirichlet Problem**

$$\Delta u = f \quad \text{in } \Omega \quad \text{with} \quad u|_{\partial\Omega} = g, \quad (9.4.5)$$

and a solution for the **Neumann Problem**

$$\Delta u = f \quad \text{in } \Omega \quad \text{with} \quad \left. \frac{\partial u}{\partial \nu} \right|_{\partial \Omega} = g . \quad (9.4.6)$$

There are a few issues to raise before attempting to solve the Dirichlet and Neumann problems above.

1. We assume for now that the function g used for the boundary conditions is at least continuous on $\partial \Omega$. We will relax the hypothesis later, when working in Sobolev spaces.
2. For the Neumann problem, the solution is not unique. If u is a solution, then $u + c$ is also a solution for any constant c .
3. Moreover, for the Neumann problem, f and g cannot be arbitrary. If $\tilde{\Omega}$ is a connected component of Ω , then it follows from Green's identity (9.1.1) with $v = 1$ on $\tilde{\Omega}$ that

$$\int_{\tilde{\Omega}} f(\mathbf{x}) \, d\mathbf{x} = \int_{\tilde{\Omega}} \Delta u(\mathbf{x}) \, d\mathbf{x} = \int_{\partial \tilde{\Omega}} \frac{\partial u}{\partial \nu}(\mathbf{x}) \, dS_{\mathbf{x}} = \int_{\partial \tilde{\Omega}} g(\mathbf{x}) \, dS_{\mathbf{x}} .$$

4. For the Dirichlet, we may assume without loss of generality that $f = 0$ in Ω or $g = 0$ on $\partial \Omega$. If v satisfies

$$\Delta v = f \quad \text{in } \Omega \quad \text{with} \quad v|_{\partial \Omega} = 0 , \quad (9.4.7)$$

and w satisfies

$$\Delta w = 0 \quad \text{in } \Omega \quad \text{with} \quad w|_{\partial \Omega} = g , \quad (9.4.8)$$

then $u = v + w$ satisfies (9.4.5). Moreover, if we can solve problems of the form (9.4.7) then we can solve problems of the form (9.4.8) and vice-versa.

Suppose that we can solve problems of the form (9.4.7). Choose $\tilde{g} \in C^2(\Omega) \cap C(\overline{\Omega})$ such that $\tilde{g} = g$ on $\partial \Omega$. Let v be the solution of

$$\Delta v = \Delta \tilde{g} \quad \text{in } \Omega \quad \text{with} \quad v|_{\partial \Omega} = 0 .$$

Then $w = \tilde{g} - v$ satisfies (9.4.8).

Suppose that we can solve problems of the form (9.4.8). Let

$$\tilde{f} = \begin{cases} f & \text{in } \Omega \\ 0 & \text{in } \mathbb{R}^n \setminus \Omega \end{cases}$$

and $\tilde{v} = N * \tilde{f}$. Then $\Delta \tilde{v} = \Delta N * \tilde{f} = \delta * \tilde{f} = \tilde{f}$. Let w be the solution of

$$\Delta w = 0 \quad \text{in } \Omega \quad \text{with} \quad w|_{\partial \Omega} = \tilde{v}|_{\partial \Omega} .$$

Then $v = \tilde{v} - w$ satisfies (9.4.7).

5. For the Neumann, we may assume without loss of generality that $f = 0$ in Ω or $g = 0$ on $\partial\Omega$. If v satisfies

$$\Delta v = f \quad \text{in } \Omega \quad \text{with} \quad \left. \frac{\partial v}{\partial \nu} \right|_{\partial\Omega} = 0, \quad (9.4.9)$$

and w satisfies

$$\Delta w = 0 \quad \text{in } \Omega \quad \text{with} \quad \left. \frac{\partial w}{\partial \nu} \right|_{\partial\Omega} = g, \quad (9.4.10)$$

then $u = v + w$ satisfies (9.4.6). Moreover, if we can solve problems of the form (9.4.9) then we can solve problems of the form (9.4.10) and vice-versa.

Suppose that we can solve problems of the form (9.4.9). Choose $\tilde{g} \in C^2(\Omega) \cap C^1(\overline{\Omega})$ such that $\frac{\partial \tilde{g}}{\partial \nu} = g$ on $\partial\Omega$. Let v be the solution of

$$\Delta v = \Delta \tilde{g} \quad \text{in } \Omega \quad \text{with} \quad \left. \frac{\partial v}{\partial \nu} \right|_{\partial\Omega} = 0.$$

Then $w = \tilde{g} - v$ satisfies (9.4.10).

Suppose that we can solve problems of the form (9.4.10). Let

$$\tilde{f} = \begin{cases} f & \text{in } \Omega \\ 0 & \text{in } \mathbb{R}^n \setminus \Omega \end{cases}$$

and $\tilde{v} = N * \tilde{f}$. Then $\Delta \tilde{v} = \tilde{f}$ as we have seen before. Let w be the solution of

$$\Delta w = 0 \quad \text{in } \Omega \quad \text{with} \quad \left. \frac{\partial w}{\partial \nu} \right|_{\partial\Omega} = \left. \frac{\partial \tilde{v}}{\partial \nu} \right|_{\partial\Omega}.$$

Then $v = \tilde{v} - w$ satisfies (9.4.9).

9.5 Green's Functions

Definition 9.5.1

Let Ω be an open subset of $\mathbb{R}^n$ with a boundary of class C^2 . Suppose that there exists a function $U : \Omega \times \overline{\Omega} \rightarrow \mathbb{R}$ such that

1. the function $\mathbf{y} \mapsto U(\mathbf{x}, \mathbf{y})$ is harmonic in Ω and continuous on $\overline{\Omega}$ for all $\mathbf{x} \in \Omega$, and
2. $U(\mathbf{x}, \mathbf{y}) = N(\mathbf{x}, \mathbf{y})$ for $\mathbf{x} \in \Omega$ and $\mathbf{y} \in \partial\Omega$,

where N is the fundamental solution defined in (9.4.1). Then, the function G defined by $G(\mathbf{x}, \mathbf{y}) = N(\mathbf{x}, \mathbf{y}) - U(\mathbf{x}, \mathbf{y})$ for $(\mathbf{x}, \mathbf{y}) \in \Omega \times \overline{\Omega} \setminus \{(\mathbf{x}, \mathbf{x}) : \mathbf{x} \in \Omega\}$ is the **Green's**

function for Ω .

In the definition above, the Green's function is obviously not specified on the diagonal $\{(\mathbf{x}, \mathbf{x}) : \mathbf{x} \in \Omega\}$.

Example 9.5.2

The first example of Green's function that students generally encounter is for ordinary differential equations over an interval. This is the one-dimensional Green's function that we have not mentioned but it must be mentioned for completeness.

We have that $N(x) = |x|/2$ for $x \in \mathbb{R}$. It is the fundamental solution of $\frac{d^2 u}{dx^2}$ because

$$\begin{aligned} \frac{d^2 N}{dx^2}(\phi) &= N\left(\frac{d^2 \phi}{dx^2}\right) = \int_{-\infty}^{\infty} N(x) \frac{d^2 \phi}{dx^2} dx = - \int_{-\infty}^0 \frac{x}{2} \frac{d^2 \phi}{dx^2} dx + \int_0^{\infty} \frac{x}{2} \frac{d^2 \phi}{dx^2} dx \\ &= \int_{-\infty}^0 \frac{1}{2} \frac{d\phi}{dx} dx - \int_0^{\infty} \frac{1}{2} \frac{d\phi}{dx} dx = \frac{1}{2} \phi(x) \Big|_{-\infty}^0 - \frac{1}{2} \phi(x) \Big|_0^{\infty} = \phi(0) = \delta(\phi) \end{aligned}$$

for all $\phi \in \mathcal{D}(\mathbb{R})$.

Let $\Omega =]0, L[$. We seek $U : \Omega \times \overline{\Omega} \rightarrow \mathbb{R}$ such that $U_x : \overline{\Omega} \rightarrow \mathbb{R}$ defined by $U(y) \equiv U(x, y)$ for all $y \in \overline{\Omega}$ satisfies $\frac{d^2 U_x}{dy^2} = 0$ on Ω , U_x is continuous on $\overline{\Omega}$, $U_x(0) = N(x, 0) = N(x) = x/2$ and $U_x(L) = N(x, L) = N(x - L) = (L - x)/2$.

Since $\frac{d^2 U_x}{dy^2} = 0$, we get $U_x(y) = Ay + B$ for some constants A and B that may depend on x . From $U_x(0) = x/2$, we get $B = x/2$. From $U_x(L) = (L - x)/2$, we get $AL + B = (L - x)/2$. Hence, $U(x, y) = U_x(y) = (1/2 - x/L)y + x/2$ for $(x, y) \in \Omega \times \overline{\Omega}$.

The Green's function is

$$\begin{aligned} G(x, y) &= N(x, y) - U(x, y) = \frac{|x - y|}{2} - \left(\frac{1}{2} - \frac{x}{L}\right)y - \frac{x}{2} \\ &= \begin{cases} \frac{x - y}{2} - \left(\frac{1}{2} - \frac{x}{L}\right)y - \frac{x}{2} = \frac{y(x - L)}{L} & \text{if } x > y \\ \frac{y - x}{2} - \left(\frac{1}{2} - \frac{x}{L}\right)y - \frac{x}{2} = \frac{x(y - L)}{L} & \text{if } x < y \end{cases} \end{aligned}$$

♣

The Green's function is unique because, for each $\mathbf{x} \in \Omega$, the function $u_{\mathbf{x}} : \overline{\Omega} \rightarrow \mathbb{R}$ defined by $u_{\mathbf{x}}(\mathbf{y}) = U(\mathbf{x}, \mathbf{y})$ is the unique solution of the Dirichlet problem

$$\Delta u = 0 \quad \text{in } \Omega \quad \text{with} \quad u(\mathbf{y}) = N(\mathbf{x}, \mathbf{y}) \quad \text{for } \mathbf{y} \in \partial\Omega. \quad (9.5.1)$$

Theorem 9.5.3

Let Ω be any bounded open subset of $\mathbb{R}^n$ with a boundary of class C^2 . The Green's function G for Ω exists ². Moreover, $\mathbf{y} \mapsto G(\mathbf{x}, \mathbf{y})$ is of class C^∞ on $\bar{\Omega} \setminus \{\mathbf{x}\}$.

Proof.

The proof of this theorem is based on Theorem 9.8.12 that we will present later.

Let N be the fundamental solution given in (9.4.1). For each $\mathbf{x} \in \Omega$, let $u_{\mathbf{x}}$ be the solution of the Dirichlet problem $\Delta u = 0$ on Ω with $u_{\mathbf{x}}(\mathbf{y}) = N(\mathbf{x}, \mathbf{y})$ for $\mathbf{y} \in \partial\Omega$ given by Theorem 9.8.12. We also have that $u_{\mathbf{x}} \in C(\bar{\Omega})$ according to Theorem 9.8.12.

Let $U(\mathbf{x}, \mathbf{y}) = u_{\mathbf{x}}(\mathbf{y})$ for $(\mathbf{x}, \mathbf{y}) \in \Omega \times \bar{\Omega}$. Then U satisfies the conditions of Definition 9.5.1 and so $G(\mathbf{x}, \mathbf{y}) = N(\mathbf{x}, \mathbf{y}) - U(\mathbf{x}, \mathbf{y})$ for $(\mathbf{x}, \mathbf{y}) \in \Omega \times \bar{\Omega} \setminus \{(\mathbf{x}, \mathbf{x}) : \mathbf{x} \in \Omega\}$ is the Green's function in the statement of the theorem.

Since $\mathbf{y} \mapsto N(\mathbf{x}, \mathbf{y})$ is harmonic on $\Omega \setminus \{\mathbf{x}\}$ and $\mathbf{y} \mapsto U(\mathbf{x}, \mathbf{y})$ is harmonic on Ω , we have that $\mathbf{y} \mapsto G(\mathbf{x}, \mathbf{y})$ is harmonic on $\Omega \setminus \{\mathbf{x}\}$. It follows from Theorem 9.1.5 that $\mathbf{y} \mapsto G(\mathbf{x}, \mathbf{y})$ is of class C^∞ on $\Omega \setminus \{\mathbf{x}\}$.

Another proof of this theorem is given in Section 12.7. It will then be proved that $\mathbf{y} \mapsto G(\mathbf{x}, \mathbf{y})$ is of class C^∞ on $\bar{\Omega} \setminus \{\mathbf{x}\}$. ■

Proposition 9.5.4

If G is the Green's function for a bounded open set $\Omega \subset \mathbb{R}^n$ with a boundary of class C^2 , then $G(\mathbf{x}, \mathbf{y}) = G(\mathbf{y}, \mathbf{x})$ for all $\mathbf{x}, \mathbf{y}$ in Ω with $\mathbf{x} \neq \mathbf{y}$.

Proof.

Given $\mathbf{z} \in \Omega$, let $G_{\mathbf{x}} : \Omega \setminus \{\mathbf{z}\} \rightarrow \mathbb{R}$ be the function defined by $G_{\mathbf{z}}(\mathbf{w}) = G(\mathbf{z}, \mathbf{w})$ for all $\mathbf{w} \in \Omega \setminus \{\mathbf{z}\}$.

Given $\mathbf{x}$ and $\mathbf{y}$, two distinct points in Ω , choose $\epsilon > 0$ small enough such that $\overline{B_\epsilon(\mathbf{x})} \subset \Omega$, $\overline{B_\epsilon(\mathbf{y})} \subset \Omega$ and $\overline{B_\epsilon(\mathbf{x})} \cap \overline{B_\epsilon(\mathbf{y})} = \emptyset$. Let $\Omega_\epsilon = \Omega \setminus (\overline{B_\epsilon(\mathbf{x})} \cup \overline{B_\epsilon(\mathbf{y})})$.

We have from Theorem 9.5.3 that $G_{\mathbf{x}} \in C^\infty(\bar{\Omega} \setminus B_\epsilon(\mathbf{x}))$ and $G_{\mathbf{y}} \in C^\infty(\bar{\Omega} \setminus B_\epsilon(\mathbf{y}))$. Thus $G_{\mathbf{x}}$ and $G_{\mathbf{y}}$ are in $C^\infty(\bar{\Omega}_\epsilon)$. We may therefore use Green's second identity (9.1.2) on Ω_ϵ . Thus

$$\begin{aligned} \int_{\Omega_\epsilon} (G_{\mathbf{x}} \Delta G_{\mathbf{y}} - G_{\mathbf{y}} \Delta G_{\mathbf{x}}) \, d\mathbf{z} &= \int_{\partial\Omega_\epsilon} \left(G_{\mathbf{x}} \frac{\partial G_{\mathbf{y}}}{\partial \nu} - G_{\mathbf{y}} \frac{\partial G_{\mathbf{x}}}{\partial \nu} \right) \, dS_{\mathbf{z}} \\ &= \int_{\partial\Omega} \left(G_{\mathbf{x}} \frac{\partial G_{\mathbf{y}}}{\partial \nu} - G_{\mathbf{y}} \frac{\partial G_{\mathbf{x}}}{\partial \nu} \right) \, dS_{\mathbf{z}} - \int_{\partial B_\epsilon(\mathbf{x})} \left(G_{\mathbf{x}} \frac{\partial G_{\mathbf{y}}}{\partial \nu} - G_{\mathbf{y}} \frac{\partial G_{\mathbf{x}}}{\partial \nu} \right) \, dS_{\mathbf{z}} \\ &\quad - \int_{\partial B_\epsilon(\mathbf{y})} \left(G_{\mathbf{x}} \frac{\partial G_{\mathbf{y}}}{\partial \nu} - G_{\mathbf{y}} \frac{\partial G_{\mathbf{x}}}{\partial \nu} \right) \, dS_{\mathbf{z}}. \end{aligned} \quad (9.5.2)$$

²We should probably say that the function U in Definition 9.5.1 exists and that we may therefore define $G = N - U$. However, we will stick to the tradition of formulating the statement in terms of G instead of U .

By definition of the Green's functions, $\mathbf{z} \mapsto U(\mathbf{x}, \mathbf{z}) = N(\mathbf{x}, \mathbf{z}) - G_{\mathbf{x}}(\mathbf{z})$ is harmonic on Ω . Since, $\mathbf{z} \mapsto N(\mathbf{x}, \mathbf{z})$ is harmonic on $\Omega \setminus B_\epsilon(\mathbf{x})$, we have that $G_{\mathbf{x}}$ is harmonic on $\Omega \setminus B_\epsilon(\mathbf{x})$. Similarly, $G_{\mathbf{y}}$ is harmonic on $\Omega \setminus B_\epsilon(\mathbf{y})$.

Since $G_{\mathbf{x}}$ and $G_{\mathbf{y}}$ are harmonic in Ω_ϵ , the left hand side in (9.5.2) is null. Moreover, $G_{\mathbf{x}} = 0$ and $G_{\mathbf{y}} = 0$ on $\partial\Omega$ by definition of the Green's functions. Thus, the first integral on the right hand side of (9.5.2) is also null. We get

$$\int_{\partial B_\epsilon(\mathbf{x})} \left(G_{\mathbf{x}} \frac{\partial G_{\mathbf{y}}}{\partial \nu} - G_{\mathbf{y}} \frac{\partial G_{\mathbf{x}}}{\partial \nu} \right) dS_{\mathbf{z}} + \int_{\partial B_\epsilon(\mathbf{y})} \left(G_{\mathbf{x}} \frac{\partial G_{\mathbf{y}}}{\partial \nu} - G_{\mathbf{y}} \frac{\partial G_{\mathbf{x}}}{\partial \nu} \right) dS_{\mathbf{z}} = 0 \quad (9.5.3)$$

for all $\epsilon > 0$.

We consider the first integral in (9.5.3).

$$\begin{aligned} & \int_{\partial B_\epsilon(\mathbf{x})} \left(G_{\mathbf{x}} \frac{\partial G_{\mathbf{y}}}{\partial \nu} - G_{\mathbf{y}} \frac{\partial G_{\mathbf{x}}}{\partial \nu} \right) dS_{\mathbf{z}} \\ &= \int_{\|\mathbf{z}\|=1} \left(G_{\mathbf{x}}(\mathbf{x} + \epsilon \mathbf{z}) \frac{\partial G_{\mathbf{y}}}{\partial \nu}(\mathbf{x} + \epsilon \mathbf{z}) - G_{\mathbf{y}}(\mathbf{x} + \epsilon \mathbf{z}) \frac{\partial G_{\mathbf{x}}}{\partial \nu}(\mathbf{x} + \epsilon \mathbf{z}) \right) \epsilon^{n-1} dS_{\mathbf{z}}. \end{aligned} \quad (9.5.4)$$

i) Let $u_{\mathbf{x}}(\mathbf{y}) = U(\mathbf{x}, \mathbf{y})$ for $(\mathbf{x}, \mathbf{y}) \in \Omega \times \overline{\Omega}$, where U is given in the Definition of the Green's function.

Since $u_{\mathbf{x}}$ is continuous on $\overline{\Omega}$ by definition of Green's functions, and $G_{\mathbf{y}}$ is continuously differentiable on $\Omega \setminus B_\epsilon(\mathbf{y})$, then $u_{\mathbf{x}}$ and $\frac{\partial G_{\mathbf{y}}}{\partial \nu}$ are bounded on a bounded open neighbourhood $V \subset \Omega$ of $\mathbf{x}$. Suppose that $|u_{\mathbf{x}}| < M_1$ and $\left| \frac{\partial G_{\mathbf{y}}}{\partial \nu} \right| < M_2$ on V .

Suppose that ϵ is small enough to have $B_\epsilon(\mathbf{x}) \subset V$. For the first term in (9.5.4), we get for $n > 2$ that

$$\begin{aligned} & \left| \int_{\|\mathbf{z}\|=1} G_{\mathbf{x}}(\mathbf{x} + \epsilon \mathbf{z}) \frac{\partial G_{\mathbf{y}}}{\partial \nu}(\mathbf{x} + \epsilon \mathbf{z}) \epsilon^{n-1} dS_{\mathbf{z}} \right| \\ &= \left| \int_{\|\mathbf{z}\|=1} (N(-\epsilon \mathbf{z}) - u_{\mathbf{x}}(\mathbf{x} + \epsilon \mathbf{z})) \frac{\partial G_{\mathbf{y}}}{\partial \nu}(\mathbf{x} + \epsilon \mathbf{z}) \epsilon^{n-1} dS_{\mathbf{z}} \right| \\ &= \left| \int_{\|\mathbf{z}\|=1} \left(\frac{1}{(2-n)\omega_n} \epsilon^{2-n} - u_{\mathbf{x}}(\mathbf{x} + \epsilon \mathbf{z}) \right) \frac{\partial G_{\mathbf{y}}}{\partial \nu}(\mathbf{x} + \epsilon \mathbf{z}) \epsilon^{n-1} dS_{\mathbf{z}} \right| \\ &\leq \int_{\|\mathbf{z}\|=1} \left(\frac{1}{(n-2)\omega_n} \epsilon^{2-n} + |u_{\mathbf{x}}(\mathbf{x} + \epsilon \mathbf{z})| \right) \left| \frac{\partial G_{\mathbf{y}}}{\partial \nu}(\mathbf{x} + \epsilon \mathbf{z}) \right| \epsilon^{n-1} dS_{\mathbf{z}} \leq \left(\frac{\epsilon}{n-2} + \epsilon^{n-1} \omega_n M_1 \right) M_2. \end{aligned}$$

Similarly, we get for $n = 2$ that

$$\begin{aligned} & \left| \int_{\|\mathbf{z}\|=1} G_{\mathbf{x}}(\mathbf{x} + \epsilon \mathbf{z}) \frac{\partial G_{\mathbf{y}}}{\partial \nu}(\mathbf{x} + \epsilon \mathbf{z}) \epsilon dS_{\mathbf{z}} \right| = \left| \int_{\|\mathbf{z}\|=1} (N(-\epsilon \mathbf{z}) - u_{\mathbf{x}}(\mathbf{x} + \epsilon \mathbf{z})) \frac{\partial G_{\mathbf{y}}}{\partial \nu}(\mathbf{x} + \epsilon \mathbf{z}) \epsilon dS_{\mathbf{z}} \right| \\ &= \left| \int_{\|\mathbf{z}\|=1} \left(\frac{1}{\omega_2} \ln(\epsilon) - u_{\mathbf{x}}(\mathbf{x} + \epsilon \mathbf{z}) \right) \frac{\partial G_{\mathbf{y}}}{\partial \nu}(\mathbf{x} + \epsilon \mathbf{z}) \epsilon dS_{\mathbf{z}} \right| \\ &\leq \int_{\|\mathbf{z}\|=1} \left(\frac{1}{\omega_2} |\ln(\epsilon)| + |u_{\mathbf{x}}(\mathbf{x} + \epsilon \mathbf{z})| \right) \left| \frac{\partial G_{\mathbf{y}}}{\partial \nu}(\mathbf{x} + \epsilon \mathbf{z}) \right| \epsilon dS_{\mathbf{z}} \leq (\epsilon |\ln(\epsilon)| + \epsilon \omega_2 M_1) M_2. \end{aligned}$$

We therefore have that $\int_{\|z\|=1} G_{\mathbf{x}}(\mathbf{x} + \epsilon \mathbf{z}) \frac{\partial G_{\mathbf{y}}}{\partial \nu}(\mathbf{x} + \epsilon \mathbf{z}) \epsilon \, dS_{\mathbf{z}} \rightarrow 0$ as $\epsilon \rightarrow 0$.

ii) We now consider the second term of the integral in (9.5.4).

$$\begin{aligned} & \int_{\|z\|=1} G_{\mathbf{y}}(\mathbf{x} + \epsilon \mathbf{z}) \frac{\partial G_{\mathbf{x}}}{\partial \nu}(\mathbf{x} + \epsilon \mathbf{z}) \epsilon^{n-1} \, dS_{\mathbf{z}} \\ &= \int_{\|z\|=1} G_{\mathbf{y}}(\mathbf{x} + \epsilon \mathbf{z}) \left(\frac{\partial N}{\partial \nu}(-\epsilon \mathbf{z}) - \frac{\partial u_{\mathbf{x}}}{\partial \nu}(\mathbf{x} + \epsilon \mathbf{z}) \right) \epsilon^{n-1} \, dS_{\mathbf{z}}. \end{aligned}$$

Since $u_{\mathbf{x}}$ is continuously differentiable on Ω by definition of the Green's functions, and $G_{\mathbf{y}}$ is continuous on $\Omega \setminus B_{\epsilon}(\mathbf{y})$, then $G_{\mathbf{y}}$ and $\frac{\partial u_{\mathbf{x}}}{\partial \nu}(\mathbf{z})$ are bounded on a bounded open neighbourhood $V \subset \Omega$ of $\mathbf{x}$. Suppose that $|G_{\mathbf{y}}| < M_3$ and $\left| \frac{\partial u_{\mathbf{x}}}{\partial \nu} \right| < M_4$ on V . For ϵ small enough to have $B_{\epsilon}(\mathbf{x}) \subset V$, we get that

$$\begin{aligned} \left| \int_{\|z\|=1} G_{\mathbf{y}}(\mathbf{x} + \epsilon \mathbf{z}) \frac{\partial u_{\mathbf{x}}}{\partial \nu}(\mathbf{x} + \epsilon \mathbf{z}) \epsilon^{n-1} \, dS_{\mathbf{z}} \right| &\leq \int_{\|z\|=1} |G_{\mathbf{y}}(\mathbf{x} + \epsilon \mathbf{z})| \left| \frac{\partial u_{\mathbf{x}}}{\partial \nu}(\mathbf{x} + \epsilon \mathbf{z}) \right| \epsilon^{n-1} \, dS_{\mathbf{z}} \\ &\leq M_3 M_4 \omega_n \epsilon^{n-1} \rightarrow 0 \end{aligned}$$

as $\epsilon \rightarrow 0$.

Moreover, for $n > 2$, we have

$$\begin{aligned} & \int_{\|z\|=1} G_{\mathbf{y}}(\mathbf{x} + \epsilon \mathbf{z}) \frac{\partial N}{\partial \nu}(-\epsilon \mathbf{z}) \epsilon^{n-1} \, dS_{\mathbf{z}} = \int_{\|z\|=1} G_{\mathbf{y}}(\mathbf{x} + \epsilon \mathbf{z}) \frac{d}{d\epsilon} \left(\frac{1}{(2-n)\omega_n} \epsilon^{2-n} \right) \epsilon^{n-1} \, dS_{\mathbf{z}} \\ &= \frac{1}{\omega_n} \int_{\|z\|=1} G_{\mathbf{y}}(\mathbf{x} + \epsilon \mathbf{z}) \, dS_{\mathbf{z}} \rightarrow \frac{1}{\omega_n} \int_{\|z\|=1} G_{\mathbf{y}}(\mathbf{x}) \, dS_{\mathbf{z}} = G_{\mathbf{y}}(\mathbf{x}) \end{aligned}$$

as $\epsilon \rightarrow 0$ and, for $n = 2$, we have

$$\begin{aligned} & \int_{\|z\|=1} G_{\mathbf{y}}(\mathbf{x} + \epsilon \mathbf{z}) \frac{\partial N}{\partial \nu}(-\epsilon \mathbf{z}) \epsilon \, dS_{\mathbf{z}} = \int_{\|z\|=1} G_{\mathbf{y}}(\mathbf{x} + \epsilon \mathbf{z}) \frac{d}{d\epsilon} \left(\frac{1}{\omega_2} \ln(\epsilon) \right) \epsilon \, dS_{\mathbf{z}} \\ &= \frac{1}{\omega_2} \int_{\|z\|=1} G_{\mathbf{y}}(\mathbf{x} + \epsilon \mathbf{z}) \, dS_{\mathbf{z}} \rightarrow \frac{1}{\omega_2} \int_{\|z\|=1} G_{\mathbf{y}}(\mathbf{x}) \, dS_{\mathbf{z}} = G_{\mathbf{y}}(\mathbf{x}) \end{aligned}$$

as $\epsilon \rightarrow 0$. Note that $\mathbf{z} \rightarrow G_{\mathbf{y}}(\mathbf{x} + \epsilon \mathbf{z})$ converges uniformly to $\mathbf{z} \rightarrow G_{\mathbf{y}}(\mathbf{x})$ on $\{\mathbf{z} : \|\mathbf{z}\| = 1\}$ as $\epsilon \rightarrow 0$ because $G_{\mathbf{y}}$ is continuous on $\bar{V}$ and so uniformly continuous on $\bar{V}$.

Note that is $f(\mathbf{x}) = g(\|\mathbf{x}\|)$ and $\nu(\mathbf{x})$ is the outward unit normal to $\partial B_{\epsilon}(\mathbf{0})$ at the point $\mathbf{x} \in \partial B_{\epsilon}(\mathbf{x})$, then

$$\frac{\partial f}{\partial \nu}(\mathbf{x}) = \lim_{s \rightarrow 0} \frac{f(\mathbf{x} + s(\mathbf{x}/\|\mathbf{x}\|)) - f(\mathbf{x})}{s} = \lim_{s \rightarrow 0} \frac{g(\epsilon + s) - g(\epsilon)}{s} = \frac{dg}{d\epsilon}(\epsilon).$$

Therefore, for $\mathbf{x} \in \partial B_{\epsilon}(\mathbf{0})$, we have that

$$\frac{\partial N}{\partial \nu}(\mathbf{x}) = \begin{cases} \frac{d}{d\epsilon} \left(\frac{1}{(2-n)\omega_n} \epsilon^{2-n} \right) & \text{if } n > 2 \\ \frac{d}{d\epsilon} \left(\frac{1}{\omega_2} \ln(\epsilon) \right) & \text{if } n = 2 \end{cases}$$

iii) (i) and (ii) imply that

$$\int_{\partial B_\epsilon(\mathbf{x})} \left(G_{\mathbf{x}} \frac{\partial G_{\mathbf{y}}}{\partial \nu} - G_{\mathbf{y}} \frac{\partial G_{\mathbf{x}}}{\partial \nu} \right) dS_{\mathbf{z}} \rightarrow -G_{\mathbf{y}}(\mathbf{x}) \quad (9.5.5)$$

as $\epsilon \rightarrow 0$.

iv) As we have done in (i) and (ii), we can prove that

$$\int_{\partial B_\epsilon(\mathbf{y})} \left(G_{\mathbf{x}} \frac{\partial G_{\mathbf{y}}}{\partial \nu} - G_{\mathbf{y}} \frac{\partial G_{\mathbf{x}}}{\partial \nu} \right) dS_{\mathbf{z}} \rightarrow G_{\mathbf{x}}(\mathbf{y}) \quad (9.5.6)$$

as $\epsilon \rightarrow 0$.

v) It follows from (9.5.5) and (9.5.6) that the limit of (9.5.3) as $\epsilon \rightarrow 0$ is $-G_{\mathbf{y}}(\mathbf{x}) + G_{\mathbf{x}}(\mathbf{y}) = 0$. ■

Let U be the function given in Definition 9.5.1 and assume that $\mathbf{y} \in \Omega$. A consequence of the previous proposition is that $\mathbf{x} \mapsto U(\mathbf{x}, \mathbf{y})$ is harmonic in Ω and can be extended continuously to $\bar{\Omega}$ by $U(\mathbf{x}, \mathbf{y}) = N(\mathbf{x}, \mathbf{y})$ for $\mathbf{x} \in \partial\Omega$.

Theorem 9.5.5

Let Ω be a bounded open subset of $\mathbb{R}^n$ with a boundary of class C^2 and let $f : \Omega \rightarrow \mathbb{R}$ be a function in $L^1(\Omega)$. The solution of the Dirichlet problem $\Delta u = f$ in Ω with $u = 0$ on $\partial\Omega$ is given by

$$u(\mathbf{x}) = \int_{\Omega} G(\mathbf{x}, \mathbf{y}) f(\mathbf{y}) d\mathbf{y} \quad , \quad \mathbf{x} \in \bar{\Omega} \quad ,$$

where G is the Green's function for Ω .

Proof.

If we set $f(\mathbf{x}) = 0$ for $\mathbf{x} \in \mathbb{R}^n \setminus \Omega$, we can write

$$u(\mathbf{x}) = \int_{\Omega} f(\mathbf{y}) N(\mathbf{x}, \mathbf{y}) d\mathbf{y} - \int_{\Omega} f(\mathbf{y}) U(\mathbf{x}, \mathbf{y}) d\mathbf{y} = (N * f)(\mathbf{x}) - \int_{\Omega} f(\mathbf{y}) U(\mathbf{x}, \mathbf{y}) d\mathbf{y} \quad ,$$

where U is the function given in Definition 9.5.1.

Since $f \in \mathcal{E}'(\Omega)$,

$$(\Delta(N * f))(\phi) = ((\Delta N) * f)(\phi) = (\delta * f)(\phi) = f(\phi)$$

for all $\phi \in \mathcal{D}(\Omega)$.

We first show that $h : \Omega \rightarrow \mathbb{R}$ defined by $h(\mathbf{x}) = \int_{\Omega} f(\mathbf{y}) U(\mathbf{x}, \mathbf{y}) d\mathbf{y}$ is locally integrable on Ω and so belongs to $\mathcal{D}'(\Omega)$. Let K be a compact subset of Ω and set $d_m = \min \{ \|\mathbf{z} - \mathbf{y}\| : \mathbf{z} \in K \text{ and } \mathbf{y} \in \partial\Omega \}$ and $d_M = \max \{ \|\mathbf{z} - \mathbf{y}\| : \mathbf{z} \in K \text{ and } \mathbf{y} \in \partial\Omega \}$. Since K and $\partial\Omega$ are compact, we have that $0 < d_m \leq d_M < \infty$. Since $\mathbf{y} \mapsto U(\mathbf{x}, \mathbf{y})$ is harmonic on Ω and continuous on $\bar{\Omega}$, we have that

$$\max_{\mathbf{y} \in \Omega} |U(\mathbf{x}, \mathbf{y})| \leq \max_{\mathbf{y} \in \partial\Omega} |U(\mathbf{x}, \mathbf{y})| = \max_{\mathbf{y} \in \partial\Omega} |N(\mathbf{x}, \mathbf{y})| = \begin{cases} \max_{\mathbf{y} \in \partial\Omega} \frac{1}{(2-n)\omega_n} \|\mathbf{x} - \mathbf{y}\|^{2-n} & \text{if } n > 2 \\ \max_{\mathbf{y} \in \partial\Omega} \frac{1}{\omega_2} |\ln(\|\mathbf{x} - \mathbf{y}\|)| & \text{if } n = 2 \end{cases}$$

$$\leq C \equiv \begin{cases} \frac{1}{(2-n)\omega_n d_m^{n-2}} & \text{if } n > 2 \\ \frac{1}{\omega_2} \max\{|\ln(d_m)|, |\ln(d_M)|\} & \text{if } n = 2 \end{cases}$$

for all $\mathbf{x} \in K$. Hence,

$$\begin{aligned} \int_K |h(\mathbf{x})| d\mathbf{x} &\leq \int_K \int_\Omega |f(\mathbf{y})| |U(\mathbf{x}, \mathbf{y})| d\mathbf{y} d\mathbf{x} \leq C \int_K \int_\Omega |f(\mathbf{y})| d\mathbf{y} d\mathbf{x} \\ &= C \underbrace{\int_K d\mathbf{x}}_{<\infty \text{ because } K \text{ is compact}} \underbrace{\int_\Omega |f(\mathbf{y})| d\mathbf{y}}_{<\infty \text{ because } f \in L^1(\Omega)} < \infty. \end{aligned}$$

We find that

$$\begin{aligned} \Delta h(\phi) &= h(\Delta\phi) = \int_\Omega h(\mathbf{x}) \Delta\phi(\mathbf{x}) d\mathbf{x} = \int_\Omega \left(\int_\Omega f(\mathbf{y}) U(\mathbf{x}, \mathbf{y}) d\mathbf{y} \right) \Delta\phi(\mathbf{x}) d\mathbf{x} \\ &= \int_\Omega f(\mathbf{y}) \left(\int_\Omega U(\mathbf{x}, \mathbf{y}) \Delta\phi(\mathbf{x}) d\mathbf{x} \right) d\mathbf{y} = \int_\Omega f(\mathbf{y}) \left(\int_\Omega \Delta_{\mathbf{x}} U(\mathbf{x}, \mathbf{y}) \phi(\mathbf{x}) d\mathbf{x} \right) d\mathbf{y} = 0 \end{aligned}$$

for all $\phi \in \mathcal{D}(\Omega)$ because $\mathbf{x} \mapsto U(\mathbf{x}, \mathbf{y})$ is harmonic on Ω .

We have found that

$$\Delta u(\phi) = \Delta(N * f)(\phi) + \Delta h(\phi) = f(\phi)$$

for all $\phi \in \mathcal{D}(\Omega)$. Hence, $\Delta u = f$ almost everywhere on Ω and in the sense of distributions.

Finally, since $G(\mathbf{x}, \mathbf{y}) = 0$ for $\mathbf{x} \in \partial\Omega$ and $\mathbf{y} \in \Omega$, we have $u(\mathbf{x}) = \int_\Omega G(\mathbf{x}, \mathbf{y}) f(\mathbf{y}) d\mathbf{y} = 0$ for $\mathbf{x} \in \partial\Omega$. ■

Remark 9.5.6

In a sense, it follows from the previous theorem that the Green's function for Ω plays a role similar to the fundamental solution for the operator Δ on $\mathbb{R}^n$.

Choose $\mathbf{x} \in \Omega$. By definition of the Green's function on Ω , the function $\tilde{u}_{\mathbf{x}} : \mathbb{R}^n \rightarrow \mathbb{R}$ defined by

$$\tilde{u}_{\mathbf{x}}(\mathbf{y}) = \begin{cases} U(\mathbf{x}, \mathbf{y}) & \text{if } \mathbf{y} \in \overline{\Omega} \\ 0 & \text{otherwise} \end{cases}$$

is harmonic in Ω . Consider the function $\tilde{G}_{\mathbf{x}} : \mathbb{R}^n \rightarrow \mathbb{R}$ defined by

$$\tilde{G}_{\mathbf{x}}(\mathbf{y}) = \begin{cases} G_{\mathbf{x}}(\mathbf{y}) = G(\mathbf{x}, \mathbf{y}) & \text{if } \mathbf{y} \in \overline{\Omega} \setminus \{\mathbf{x}\} \\ 0 & \text{otherwise} \end{cases}$$

and let $N_{\mathbf{x}}(\mathbf{y}) = N(\mathbf{x}, \mathbf{y})$ for $\mathbf{y} \in \mathbb{R}^n \setminus \{\mathbf{x}\}$. Then $\tilde{G}_{\mathbf{x}} = N_{\mathbf{x}} - \tilde{u}_{\mathbf{x}}$ defines a distributions on $\mathbb{R}^n$ and

$$\langle \Delta \tilde{G}_{\mathbf{x}}, \phi \rangle = \langle \Delta N_{\mathbf{x}}, \phi \rangle - \langle \Delta \tilde{u}_{\mathbf{x}}, \phi \rangle = \delta_{\mathbf{x}}(\phi) - 0 = \phi(\mathbf{x})$$

for all $\phi \in \mathcal{D}(\overline{\Omega})$. Thus, $\Delta G_{\mathbf{x}} = \delta_{\mathbf{x}}$ on Ω . This property is often used as the definition of a Green function for Ω . ♠

Theorem 9.5.7

Let Ω be a bounded open subset of $\mathbb{R}^n$ with a boundary of class C^2 , and let $g : \partial\Omega \rightarrow \mathbb{R}$ be a continuous function. The solution of the Dirichlet problem $\Delta u = 0$ in Ω with $u = g$ on $\partial\Omega$ is given by

$$u(\mathbf{x}) = \int_{\partial\Omega} g(\mathbf{y}) \frac{\partial G_{\mathbf{x}}}{\partial \nu}(\mathbf{y}) dS_{\mathbf{y}} \quad , \quad \mathbf{x} \in \Omega \quad ,$$

where $G_{\mathbf{x}}(\mathbf{y}) = G(\mathbf{x}, \mathbf{y})$ for $\mathbf{y} \in \overline{\Omega} \setminus \{\mathbf{x}\}$ and G is the Green's function for Ω . In particular, u can be continuously extended to $\overline{\Omega}$ by $u(\mathbf{x}) = g(\mathbf{x})$ for $\mathbf{x} \in \partial\Omega$.

Proof.

The proof is given in Section 12.7. ■

We proceed as in the proof of Proposition 9.5.4 to justify the formula for u in the previous theorem.

Suppose that $u \in C^1(\overline{\Omega}) \cap C^2(\Omega)$ is a solution of $\Delta u = 0$ in Ω with $u = g$ on $\partial\Omega$. Fix $\mathbf{x} \in \Omega$ and choose $\epsilon > 0$ small enough such that $\overline{B_{\epsilon}(\mathbf{x})} \subset \Omega$. Let $\Omega_{\epsilon} = \Omega \setminus \overline{B_{\epsilon}(\mathbf{x})}$.

We have from Theorem 9.5.3 that $G_{\mathbf{x}} \in C^{\infty}(\overline{\Omega_{\epsilon}})$. We may therefore use Green's second identity (9.1.2) on Ω_{ϵ} to get

$$\begin{aligned} \int_{\Omega_{\epsilon}} (G_{\mathbf{x}} \Delta u - u \Delta G_{\mathbf{x}}) d\mathbf{z} &= \int_{\partial\Omega_{\epsilon}} \left(G_{\mathbf{x}} \frac{\partial u}{\partial \nu} - u \frac{\partial G_{\mathbf{x}}}{\partial \nu} \right) dS \\ &= \int_{\partial\Omega} \left(G_{\mathbf{x}} \frac{\partial u}{\partial \nu} - u \frac{\partial G_{\mathbf{x}}}{\partial \nu} \right) dS - \int_{\partial B_{\epsilon}(\mathbf{x})} \left(G_{\mathbf{x}} \frac{\partial u}{\partial \nu} - u \frac{\partial G_{\mathbf{x}}}{\partial \nu} \right) dS \end{aligned}$$

Since u and $G_{\mathbf{x}}$ are harmonic in Ω , $G_{\mathbf{x}}(\mathbf{y}) = 0$ for all $\mathbf{y} \in \partial\Omega$ and $u(\mathbf{y}) = g(\mathbf{y})$ for all $\mathbf{y} \in \partial\Omega$, we get

$$0 = - \int_{\partial\Omega} g \frac{\partial G_{\mathbf{x}}}{\partial \nu} dS - \int_{\partial B_{\epsilon}(\mathbf{x})} \left(G_{\mathbf{x}} \frac{\partial u}{\partial \nu} - u \frac{\partial G_{\mathbf{x}}}{\partial \nu} \right) dS \quad (9.5.7)$$

for all $\epsilon > 0$.

We first consider the integral $\int_{\partial B_{\epsilon}(\mathbf{x})} G_{\mathbf{x}} \frac{\partial u}{\partial \nu} dS$.

Let $u_{\mathbf{x}}(\mathbf{y}) = U(\mathbf{x}, \mathbf{y})$ for $(\mathbf{x}, \mathbf{y}) \in \Omega \times \overline{\Omega}$, where U is given in the Definition of the Green's function. Since $u_{\mathbf{x}}$ is continuous on $\overline{\Omega}$ by definition of Green's functions, and u is continuously differentiable on Ω , then $u_{\mathbf{x}}$ and $\frac{\partial u}{\partial \nu}$ are bounded on a bounded open neighbourhood $V \subset \Omega$ of $\mathbf{x}$. Suppose that $|u_{\mathbf{x}}| < M_1$ and $\left| \frac{\partial u}{\partial \nu} \right| < M_2$ on V . If ϵ is small enough to have $B_{\epsilon}(\mathbf{x}) \subset V$, we get for $n > 2$ that

$$\begin{aligned} \left| \int_{\partial B_{\epsilon}(\mathbf{x})} G_{\mathbf{x}} \frac{\partial u}{\partial \nu} dS \right| &= \left| \int_{\|\mathbf{z}\|=1} G_{\mathbf{x}}(\mathbf{x} + \epsilon \mathbf{z}) \frac{\partial u}{\partial \nu}(\mathbf{x} + \epsilon \mathbf{z}) \epsilon^{n-1} dS_{\mathbf{z}} \right| \\ &= \left| \int_{\|\mathbf{z}\|=1} (N(-\epsilon \mathbf{z}) - u_{\mathbf{x}}(\mathbf{x} + \epsilon \mathbf{z})) \frac{\partial u}{\partial \nu}(\mathbf{x} + \epsilon \mathbf{z}) \epsilon^{n-1} dS_{\mathbf{z}} \right| \end{aligned}$$

$$\begin{aligned}
&= \left| \int_{\|z\|=1} \left(\frac{1}{(2-n)\omega_n} \epsilon^{2-n} - u_{\mathbf{x}}(\mathbf{x} + \epsilon \mathbf{z}) \right) \frac{\partial u}{\partial \nu}(\mathbf{x} + \epsilon \mathbf{z}) \epsilon^{n-1} dS_{\mathbf{z}} \right| \\
&\leq \int_{\|z\|=1} \left(\frac{1}{(n-2)\omega_n} \epsilon^{2-n} + |u_{\mathbf{x}}(\mathbf{x} + \epsilon \mathbf{z})| \right) \left| \frac{\partial u}{\partial \nu}(\mathbf{x} + \epsilon \mathbf{z}) \right| \epsilon^{n-1} dS_{\mathbf{z}} \leq \left(\frac{\epsilon}{n-2} + \epsilon^{n-1} \omega_n M_1 \right) M_2 .
\end{aligned}$$

Similarly, we get for $n = 2$ that

$$\begin{aligned}
&\left| \int_{\partial B_{\epsilon}(\mathbf{x})} G_{\mathbf{x}} \frac{\partial u}{\partial \nu} dS \right| = \left| \int_{\|z\|=1} G_{\mathbf{x}}(\mathbf{x} + \epsilon \mathbf{z}) \frac{\partial u}{\partial \nu}(\mathbf{x} + \epsilon \mathbf{z}) \epsilon dS_{\mathbf{z}} \right| \\
&= \left| \int_{\|z\|=1} (N(-\epsilon \mathbf{z}) - u_{\mathbf{x}}(\mathbf{x} + \epsilon \mathbf{z})) \frac{\partial u}{\partial \nu}(\mathbf{x} + \epsilon \mathbf{z}) \epsilon dS_{\mathbf{z}} \right| \\
&= \left| \int_{\|z\|=1} \left(\frac{1}{\omega_2} \ln(\epsilon) - u_{\mathbf{x}}(\mathbf{x} + \epsilon \mathbf{z}) \right) \frac{\partial u}{\partial \nu}(\mathbf{x} + \epsilon \mathbf{z}) \epsilon dS_{\mathbf{z}} \right| \\
&\leq \int_{\|z\|=1} \left(\frac{1}{\omega_2} |\ln(\epsilon)| + |u_{\mathbf{x}}(\mathbf{x} + \epsilon \mathbf{z})| \right) \left| \frac{\partial u}{\partial \nu}(\mathbf{x} + \epsilon \mathbf{z}) \right| \epsilon dS_{\mathbf{z}} \leq (\epsilon |\ln(\epsilon)| + \epsilon \omega_2 M_1) M_2 .
\end{aligned}$$

We therefore have that

$$\int_{\partial B_{\epsilon}(\mathbf{x})} G_{\mathbf{x}} \frac{\partial u}{\partial \nu} dS \rightarrow 0$$

as $\epsilon \rightarrow 0$.

We now consider the integral

$$\begin{aligned}
\int_{\partial B_{\epsilon}(\mathbf{x})} u \frac{\partial G_{\mathbf{x}}}{\partial \nu} dS &= \int_{\|z\|=1} u(\mathbf{x} + \epsilon \mathbf{z}) \frac{\partial G_{\mathbf{x}}}{\partial \nu}(\mathbf{x} + \epsilon \mathbf{z}) \epsilon^{n-1} dS_{\mathbf{z}} \\
&= \int_{\|z\|=1} u(\mathbf{x} + \epsilon \mathbf{z}) \left(\frac{\partial N}{\partial \nu}(-\epsilon \mathbf{z}) - \frac{\partial u_{\mathbf{x}}}{\partial \nu}(\mathbf{x} + \epsilon \mathbf{z}) \right) \epsilon^{n-1} dS_{\mathbf{z}} .
\end{aligned}$$

Since $u_{\mathbf{x}}$ is continuously differentiable on Ω by definition of the Green's functions, and u is continuous on Ω , then u and $\frac{\partial u_{\mathbf{x}}}{\partial \nu}$ are bounded on a bounded open neighbourhood $V \subset \Omega$ of $\mathbf{x}$. Suppose that $|G_{\mathbf{y}}| < M_3$ and $\left| \frac{\partial u_{\mathbf{x}}}{\partial \nu} \right| < M_4$ on V . We get for ϵ small enough to have $B_{\epsilon}(\mathbf{x}) \subset V$ that

$$\begin{aligned}
\left| \int_{\|z\|=1} u(\mathbf{x} + \epsilon \mathbf{z}) \frac{\partial u_{\mathbf{x}}}{\partial \nu}(\mathbf{x} + \epsilon \mathbf{z}) \epsilon^{n-1} dS_{\mathbf{z}} \right| &\leq \int_{\|z\|=1} |G_{\mathbf{y}}(\mathbf{x} + \epsilon \mathbf{z})| \left| \frac{\partial u_{\mathbf{x}}}{\partial \nu}(\mathbf{x} + \epsilon \mathbf{z}) \right| \epsilon^{n-1} dS_{\mathbf{z}} \\
&\leq M_3 M_4 \omega_n \epsilon^{n-1} \rightarrow 0
\end{aligned}$$

as $\epsilon \rightarrow 0$. Moreover, we get for $n > 2$ that

$$\begin{aligned}
\int_{\|z\|=1} u(\mathbf{x} + \epsilon \mathbf{z}) \frac{\partial N}{\partial \nu}(-\epsilon \mathbf{z}) \epsilon^{n-1} dS_{\mathbf{z}} &= \int_{\|z\|=1} u(\mathbf{x} + \epsilon \mathbf{z}) \frac{d}{d\epsilon} \left(\frac{1}{(2-n)\omega_n} \epsilon^{2-n} \right) \epsilon^{n-1} dS_{\mathbf{z}} \\
&= \frac{1}{\omega_n} \int_{\|z\|=1} u(\mathbf{x} + \epsilon \mathbf{z}) dS_{\mathbf{z}} \rightarrow \frac{1}{\omega_n} \int_{\|z\|=1} u(\mathbf{x}) dS_{\mathbf{z}} = u(\mathbf{x})
\end{aligned}$$

as $\epsilon \rightarrow 0$ and, we get for $n = 2$ that

$$\int_{\|z\|=1} u(\mathbf{x} + \epsilon \mathbf{z}) \frac{\partial N}{\partial \nu}(-\epsilon \mathbf{z}) \epsilon dS_{\mathbf{z}} = \int_{\|z\|=1} u(\mathbf{x} + \epsilon \mathbf{z}) \frac{d}{d\epsilon} \left(\frac{1}{\omega_2} \ln(\epsilon) \right) \epsilon dS_{\mathbf{z}}$$

$$= \frac{1}{\omega_2} \int_{\|z\|=1} u(\mathbf{x} + \epsilon z) dS_z \rightarrow \frac{1}{\omega_2} \int_{\|z\|=1} u(\mathbf{x}) dS_z = u(\mathbf{x})$$

as $\epsilon \rightarrow 0$. Note that $z \rightarrow u(\mathbf{x} + \epsilon z)$ converges uniformly to $z \rightarrow u(\mathbf{x})$ on $\{z : \|z\| = 1\}$ as $\epsilon \rightarrow 0$ because u is continuous on $\bar{V}$ and so uniformly continuous on $\bar{V}$.

It follows from (9.5.7) that $u(\mathbf{x}) = \int_{\partial\Omega} g \frac{\partial G_{\mathbf{x}}}{\partial \nu} dS$.

Remark 9.5.8

There is also a formal justification of the previous theorem.

As we have seen in Remark 9.5.6, $\langle \Delta \tilde{G}_{\mathbf{x}}, \phi \rangle = \phi(\mathbf{x})$ for all $\phi \in \mathcal{D}(\bar{\Omega})$. Since $\mathcal{D}(\bar{\Omega})$ is uniformly dense in $C(\bar{\Omega})$, we get

$$u(\mathbf{x}) = \int_{\Omega} u(\mathbf{y}) \Delta G_{\mathbf{x}}(\mathbf{y}) d\mathbf{y} \quad (9.5.8)$$

for $\mathbf{x} \in \Omega$ and $u \in C(\bar{\Omega})$ because $\tilde{G}_{\mathbf{x}} = G_{\mathbf{x}}$ on $\Omega \setminus \{\mathbf{x}\}$. Suppose that $u \in C^1(\bar{\Omega}) \cap C^2(\Omega)$ is the solution of $\Delta u = 0$ in Ω and $u = g$ on $\partial\Omega$. It is here that we are very formal and assume that the Green's identity (9.1.2) can be used with $G_{\mathbf{x}}$ on Ω though we do not have that $G_{\mathbf{x}} \in C^1(\bar{\Omega}) \cap C^2(\Omega)$. We get

$$\int_{\partial\Omega} \left(G_{\mathbf{x}} \frac{\partial u}{\partial \nu} - u \frac{\partial G_{\mathbf{x}}}{\partial \nu} \right) dS = \int_{\Omega} (G_{\mathbf{x}} \Delta u - u \Delta G_{\mathbf{x}}) d\mathbf{y}.$$

Since $\Delta u = 0$ in Ω , $u = g$ on $\partial\Omega$ and $G_{\mathbf{x}}(\mathbf{y}) = G(\mathbf{x}, \mathbf{y}) = 0$ for $\mathbf{y} \in \partial\Omega$ and $\mathbf{x} \in \Omega$, we get from (9.5.8) that

$$\int_{\partial\Omega} u \frac{\partial G_{\mathbf{x}}}{\partial \nu} dS = \int_{\Omega} u \Delta G_{\mathbf{x}} d\mathbf{y} = u(\mathbf{x}).$$

♠

Definition 9.5.9

For $\mathbf{x} \in \Omega$, consider $G_{\mathbf{x}}(\mathbf{y}) = G(\mathbf{x}, \mathbf{y})$ for $\mathbf{y} \in \bar{\Omega} \setminus \{\mathbf{x}\}$. The function $\mathbf{y} \mapsto \frac{\partial G_{\mathbf{x}}}{\partial \nu}(\mathbf{y})$ for $\mathbf{y} \in \partial\Omega$ is called the **Poisson Kernel** for Ω .

9.5.1 Dirichlet Problem on the Half Space

In this section, we consider

$$\mathbb{R}_+^n = \{\mathbf{x} \in \mathbb{R}^n : x_n > 0\}.$$

We will also use the following notation. If $\mathbf{x} = (x_1, x_2, \dots, x_n)$, then $\tilde{\mathbf{x}} = (x_1, x_2, \dots, x_{n-1}, -x_n)$.

The **Green's function** on $\mathbb{R}_+^n$ is defined by $G(\mathbf{x}, \mathbf{y}) = N(\mathbf{x}, \mathbf{y}) - U(\mathbf{x}, \mathbf{y})$ for $\mathbf{x} \in \mathbb{R}_+^n$ and $\mathbf{y} \in \bar{\mathbb{R}}_+^n$ with $\mathbf{x} \neq \mathbf{y}$, where $U(\mathbf{x}, \mathbf{y}) = N(\tilde{\mathbf{x}}, \mathbf{y})$ for $\mathbf{x} \in \mathbb{R}_+^n$ and $\mathbf{y} \in \bar{\mathbb{R}}_+^n$.

i) The function $\mathbf{y} \mapsto U(\mathbf{x}, \mathbf{y}) = N(\tilde{\mathbf{x}}, \mathbf{y})$ is harmonic in $\mathbb{R}_+^n$ and continuous on $\bar{\mathbb{R}}_+^n$ because $\tilde{\mathbf{x}} \notin \bar{\mathbb{R}}_+^n$. In fact, it is harmonic in $\mathbb{R}^n \setminus \{\tilde{\mathbf{x}}\}$.

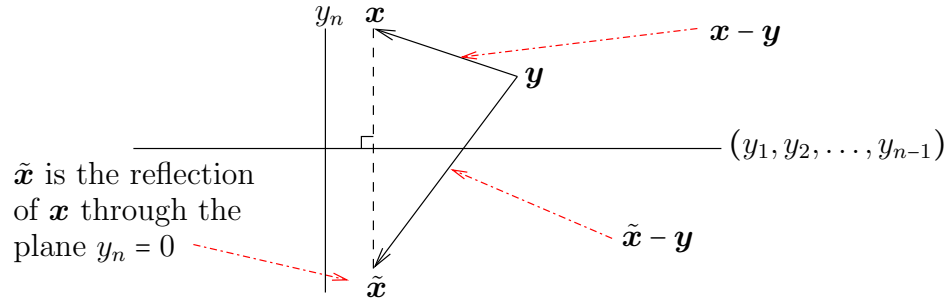


Figure 9.2: Domain of the Green's function for the half space.

ii) We have $U(\mathbf{x}, \mathbf{y}) = N(\tilde{\mathbf{x}}, \mathbf{y}) = N(\mathbf{x}, \mathbf{y})$ for $\mathbf{y} \in \partial\mathbb{R}_+^n$ because $\|\mathbf{x} - \mathbf{y}\| = \|\tilde{\mathbf{x}} - \mathbf{y}\|$ when $y_n = 0$ since $\tilde{\mathbf{x}}$ is the reflection of $\mathbf{x}$ through the plane $y_n = 0$ (Figure 9.2).

Let $\tilde{G}_{\mathbf{x}}(\mathbf{y}) = N(\mathbf{x}, \mathbf{y}) - N(\tilde{\mathbf{x}}, \mathbf{y})$ for $\mathbf{y} \in \mathbb{R}^n$ and $\mathbf{y} \neq \mathbf{x}, \tilde{\mathbf{x}}$. The function $\tilde{G}_{\mathbf{x}}$ defines a distribution on $\mathbb{R}^n$ such that

$$\Delta \tilde{G}_{\mathbf{x}} = \delta_{\mathbf{x}} - \delta_{\tilde{\mathbf{x}}}$$

on $\mathcal{D}(\mathbb{R}^n)$ according to Theorem 9.3.1. Hence,

$$\Delta \tilde{G}_{\mathbf{x}}(\phi) = \delta_{\mathbf{x}}(\phi) - \delta_{\tilde{\mathbf{x}}}(\phi) = \phi(\mathbf{x}) - \phi(\tilde{\mathbf{x}}) = \phi(\mathbf{x})$$

for $\phi \in \mathcal{D}(\mathbb{R}_+^n)$ assuming that a function $\phi \in \mathcal{D}(\mathbb{R}_+^n)$ is expanded to $\mathbb{R}^n$ by $\phi(\mathbf{y}) = 0$ for $\mathbf{y} \in \mathbb{R}^n \setminus \mathbb{R}_+^n$. Since $\tilde{G}_{\mathbf{x}}(\mathbf{y}) = G_{\mathbf{x}}(\mathbf{y}) = G(\mathbf{x}, \mathbf{y})$ for all $\mathbf{y} \in \overline{\mathbb{R}_+^n}$ and $\mathbf{x} \in \mathbb{R}_+^n$, we get $\Delta G_{\mathbf{x}} = \delta_{\mathbf{x}}$ in the sense of distribution on $\mathbb{R}_+^n$ for $\mathbf{x} \in \mathbb{R}_+^n$. As mentioned in Remark 9.5.6, this is expected from the Green's function for $\mathbb{R}_+^n$.

Since our domain $\mathbb{R}_+^n$ is not bounded, we may not apply Theorem 9.5.5 and Theorem 9.5.7. The reader may verify that

$$u(\mathbf{x}) = \int_{\mathbb{R}_+^n} G(\mathbf{x}, \mathbf{y}) f(\mathbf{y}) d\mathbf{y}$$

is the solution of $\Delta u = f$ in $\mathbb{R}_+^n$ with $u = 0$ on $\partial\mathbb{R}_+^n$ if $f \in L^1(\mathbb{R}_+^n)$ has compact support. We will not consider this problem but instead we will focus on the Dirichlet Problem $\Delta u = 0$ in $\mathbb{R}_+^n$ with $u = g$ on $\partial\mathbb{R}_+^n$ if $g \in C(\partial\mathbb{R}_+^n)$. More precisely, we will prove a version of Theorem 9.5.7 for $\mathbb{R}_+^n$; namely, Theorem 9.5.11 below.

Since

$$N(\mathbf{x}) = \begin{cases} \frac{1}{(2-n)\omega_n} \|\mathbf{x}\|^{2-n} & \text{if } n > 2 \\ \frac{1}{\omega_2} \ln(\|\mathbf{x}\|) & \text{if } n = 2 \end{cases}$$

for $\mathbf{x} \neq \mathbf{0}$, The Green's function is

$$G(\mathbf{x}, \mathbf{y}) = N(\mathbf{x}, \mathbf{y}) - N(\tilde{\mathbf{x}}, \mathbf{y})$$

$$= \begin{cases} \frac{1}{(2-n)\omega_n} \|\mathbf{x} - \mathbf{y}\|^{2-n} - \frac{1}{(2-n)\omega_n} \|\tilde{\mathbf{x}} - \mathbf{y}\|^{2-n} & \text{if } n > 2 \\ \frac{1}{\omega_2} \ln(\|\mathbf{x} - \mathbf{y}\|) - \frac{1}{\omega_2} \ln(\|\tilde{\mathbf{x}} - \mathbf{y}\|) & \text{if } n = 2 \end{cases}$$

for $\mathbf{x} \in \mathbb{R}_+^n$ and $\mathbf{y} \in \overline{\mathbb{R}_+^n}$ with $\mathbf{x} \neq \mathbf{y}$. The **Poisson kernel** is

$$\frac{\partial G_{\mathbf{x}}}{\partial \nu}(\mathbf{y}) = -\frac{\partial G}{\partial y_n}(\mathbf{x}, \mathbf{y}) = \frac{x_n - y_n}{\omega_n} \|\mathbf{x} - \mathbf{y}\|^{-n} + \frac{x_n + y_n}{\omega_n} \|\tilde{\mathbf{x}} - \mathbf{y}\|^{-n} = \frac{2x_n}{\omega_n} \|\tilde{\mathbf{x}} - \mathbf{y}\|^{-n}$$

for $\mathbf{x} \in \mathbb{R}_+^n$ and $\mathbf{y} \in \partial\mathbb{R}_+^n$. Recall that $\|\mathbf{x} - \mathbf{y}\| = \|\tilde{\mathbf{x}} - \mathbf{y}\|$ for $\mathbf{y} \in \partial\mathbb{R}_+^n$.

For the remainder of this section, we use the following convention. Every $\mathbf{x} \in \mathbb{R}_+^n$ is written $\mathbf{x} = (\check{\mathbf{x}}, x_n)$ for $\check{\mathbf{x}} \in \mathbb{R}^{n-1}$ and $x_n > 0$,

Let

$$P_{x_n}(\check{\mathbf{x}}) \equiv \frac{2x_n}{\omega_n} (\|\check{\mathbf{x}}\|^2 + x_n^2)^{-n/2}, \quad \check{\mathbf{x}} \in \mathbb{R}^{n-1}.$$

Then

$$u(\mathbf{x}) = \int_{\partial\mathbb{R}_+^n} g(\mathbf{y}) \frac{\partial G_{\mathbf{x}}}{\partial \nu}(\mathbf{y}) dS_{\mathbf{y}} = \int_{\mathbb{R}^{n-1}} g(\check{\mathbf{y}}, 0) P_{x_n}(\check{\mathbf{x}} - \check{\mathbf{y}}) d\check{\mathbf{y}} = (g * P_{x_n})(\check{\mathbf{x}}),$$

where the convolution is on $\mathbb{R}^{n-1}$.

Lemma 9.5.10

$$\int_{\mathbb{R}^{n-1}} P_1(\check{\mathbf{y}}) d\check{\mathbf{y}} = 1 \text{ and } P_{x_n}(\check{\mathbf{x}}) = x_n^{-(n-1)} P_1(x_n^{-1}\check{\mathbf{x}}).$$

Proof.

We have that

$$\begin{aligned} \int_{\mathbb{R}^{n-1}} P_1(\check{\mathbf{y}}) d\check{\mathbf{y}} &= \int_{\|\check{\mathbf{y}}\|=1} \int_0^\infty \frac{2}{\omega_n} (r^2 + 1)^{-n/2} r^{n-2} dr dS_{\check{\mathbf{y}}} = \frac{2\omega_{n-1}}{\omega_n} \int_0^\infty (r^2 + 1)^{-n/2} r^{n-2} dr \\ &= \frac{\omega_{n-1}}{\omega_n} \int_0^1 s^{(n-3)/2} (1-s)^{-1/2} ds = \frac{\Gamma(n/2)}{\Gamma(1/2)\Gamma((n-1)/2)} \int_0^1 s^{(n-3)/2} (1-s)^{-1/2} ds = 1, \end{aligned}$$

where we have used the substitution $s = \frac{r^2}{1+r^2}$ and $ds = \frac{2r}{(1+r^2)^2} dr$, and the formulae $\omega_n = 2\pi^{n/2}/\Gamma(n/2)$ and $\Gamma(1/2) = \sqrt{\pi}$. The last integral is one of the identities satisfied by the Beta function (see [9]). ■

Theorem 9.5.11

Suppose that $g \in L^p(\partial\mathbb{R}_+^n)$,³ for some integer $p \geq 1$. Then $u(\mathbf{x}) = (g * P_{x_n})(\check{\mathbf{x}})$ for $\mathbf{x} \in \mathbb{R}_+^n$ is well defined and harmonic in $\mathbb{R}_+^n$. Moreover, $u_{x_n}(\check{\mathbf{x}}) \equiv u(\check{\mathbf{x}}, x_n)$ for $\check{\mathbf{x}} \in \mathbb{R}^{n-1}$ satisfies $u_{x_n} \rightarrow g$ in $L^p(\mathbb{R}^{n-1})$ as $x_n \rightarrow 0$.

If $g \in C(\partial\mathbb{R}_+^n)$ is bounded, then u can be continuously extended to $R = \{\mathbf{x} \in \mathbb{R}^n : x_n \geq 0\}$

by $u(\mathbf{x}) = g(\mathbf{x})$ for $\mathbf{x} \in \partial\mathbb{R}_+^n$.

Proof.

i) We first show that u is well defined. We have that $P_{x_n} \in L^q(\mathbb{R}^{n-1})$ for all integer $q \geq 1$. This is a consequence of the following two facts. P_{x_n} is bounded at the origin; namely, $|P_{x_n}(\check{\mathbf{x}})| \leq \frac{2}{\omega_n x_n^{n-1}}$ for all $\check{\mathbf{x}} \in \mathbb{R}^{n-1}$. P_{x_n} is L^q -integrable away of the origin because $|P_{x_n}(\check{\mathbf{x}})| \leq \frac{2x_n}{\omega_n \|\check{\mathbf{x}}\|^n}$ for all $\check{\mathbf{x}} \in \mathbb{R}^{n-1} \setminus B_1(\mathbf{0})$. Hence, for all $\mathbf{x} \in \mathbb{R}^n$,

$$\begin{aligned} \mathbb{R}^{n-1} &\rightarrow \mathbb{R} \\ \check{\mathbf{y}} &\mapsto g(\check{\mathbf{y}}) P_{x_n}(\check{\mathbf{x}} - \check{\mathbf{y}}) \end{aligned}$$

is in $L^1(\mathbb{R}^{n-1})$. For instance, if $1 < p < \infty$ we have

$$\int_{\mathbb{R}^{n-1}} |g(\check{\mathbf{y}}) P_{x_n}(\check{\mathbf{x}} - \check{\mathbf{y}})| \, d\check{\mathbf{y}} \leq \left(\int_{\mathbb{R}^{n-1}} |g(\check{\mathbf{y}})|^p \, d\check{\mathbf{y}} \right)^{1/p} \left(\int_{\mathbb{R}^{n-1}} |P_{x_n}(\check{\mathbf{x}} - \check{\mathbf{y}})|^q \, d\check{\mathbf{y}} \right)^{1/q} < \infty$$

for $\frac{1}{p} + \frac{1}{q} = 1$. Similarly, for $p = 1$,

$$\int_{\mathbb{R}^{n-1}} |g(\check{\mathbf{y}}) P_{x_n}(\check{\mathbf{x}} - \check{\mathbf{y}})| \, d\check{\mathbf{y}} \leq \sup_{\check{\mathbf{x}} \in \mathbb{R}^{n-1}} |P_{x_n}(\check{\mathbf{x}})| \int_{\mathbb{R}^{n-1}} |g(\check{\mathbf{y}})| \, d\check{\mathbf{y}} \leq \frac{2}{\omega_n x_n^{n-1}} \int_{\mathbb{R}^{n-1}} |g(\check{\mathbf{y}})| \, d\check{\mathbf{y}} < \infty.$$

Thus $u : \mathbb{R}_+^n \rightarrow \mathbb{R}$ is well defined.

ii) We show that u is harmonic in $\mathbb{R}_+^n$. The function $\mathbf{x} \mapsto P_{x_n}(\check{\mathbf{x}} - \check{\mathbf{y}})$ is of class C^∞ on $\mathbb{R}_+^n$ for all $\check{\mathbf{y}} \in \partial\mathbb{R}^{n-1}$. Choose $\delta > 0$ and $b > a > 0$. Given $k \geq 0$ and a multi-index $\alpha \in \mathbb{N}^{n-1}$, there exists $C > 0$ such that

$$\left| \left(D_{\check{\mathbf{x}}}^\alpha \frac{\partial^k}{\partial x_n^k} P_{x_n}(\check{\mathbf{x}} - \check{\mathbf{y}}) \right) \right| \leq h(\check{\mathbf{y}}) = \begin{cases} C & \text{if } \|\check{\mathbf{y}}\| \leq 2\delta \\ C \|\check{\mathbf{y}}\|^{-n} & \text{if } \|\check{\mathbf{y}}\| > 2\delta \end{cases}$$

for $\|\check{\mathbf{x}}\| \leq \delta$ and $a \leq x_n \leq b$. The constant C depends only on a, b, δ, k and α . Thus

$$\left| g(\check{\mathbf{y}}) \left(D_{\check{\mathbf{x}}}^\alpha \frac{\partial^k}{\partial x_n^k} P_{x_n}(\check{\mathbf{x}} - \check{\mathbf{y}}) \right) \right| \leq |g(\check{\mathbf{y}})| h(\check{\mathbf{y}})$$

for all $\check{\mathbf{y}} \in \mathbb{R}^{n-1}$ and $\mathbf{x} \in \mathbb{R}_+^n$ such that $\|\check{\mathbf{x}}\| \leq \delta$ and $a \leq x_n \leq b$. Proceeding as in (i), we can show that $gh \in L^1(\mathbb{R}^{n-1})$. Hence, since a, b and δ are arbitrary, we may interchange the integral with respect to $\check{\mathbf{y}}$ and the partial derivatives of any order with respect to the x_i (see [7]).

Since $\mathbf{x} \mapsto G(\mathbf{x}, \mathbf{y})$ is harmonic in $\mathbb{R}_+^n$ for $\mathbf{y} \in \partial\mathbb{R}_+^n$, we have that

$$\mathbf{x} \mapsto -\frac{\partial G}{\partial y_n}(\mathbf{x}, \mathbf{y}) = P_{x_n}(\check{\mathbf{x}} - \check{\mathbf{y}})$$

³In other words, $g \in L^p(\mathbb{R}^{n-1})$ since $\partial\mathbb{R}_+^n \cong \mathbb{R}^{n-1}$.

is harmonic in $\mathbb{R}_+^n$ for $\mathbf{y} \in \partial\mathbb{R}_+^n$. Hence,

$$\Delta u(\mathbf{x}) = \int_{\mathbb{R}^{n-1}} g(\check{\mathbf{y}}) \Delta_{\mathbf{x}} P_{x_n}(\check{\mathbf{x}} - \check{\mathbf{y}}) d\check{\mathbf{y}} = 0$$

for all $\mathbf{x} \in \mathbb{R}_+^n$.

iii) We consider the boundary conditions. If $g \in L^p(\mathbb{R}^{n-1})$ for some integer $p \geq 1$, it follows from Proposition 6.5.3 and Lemma 9.5.10 that $u_{x_n} = g * P_{x_n} \rightarrow g$ in $L^p(\mathbb{R}^{n-1})$ as $x_n \rightarrow 0$.

iv) If g is continuous and bounded on $\mathbb{R}^{n-1}$, in particular g is uniformly continuous on any compact subset of $\mathbb{R}^{n-1}$, it follows from Proposition 6.5.3 and Lemma 9.5.10 that $g * P_{x_n} \rightarrow g$ uniformly on compact subsets of $\mathbb{R}^{n-1}$ as $x_n \rightarrow 0$. In particular, $u(\mathbf{x}) = g(\mathbf{x})$ for $\mathbf{x} \in \partial\mathbb{R}_+^n$ defines a continuous extension of u on R . ■

Remark 9.5.12

The solution of the Dirichlet problem $\Delta u = 0$ on $\mathbb{R}_+^n$ with the condition $u = g$ on $\partial\mathbb{R}_+^n$ is not unique. If u is a solution, then $\mathbf{x} \rightarrow u(\mathbf{x}) + cx_n$ is also a solution. However, $\mathbf{x} \rightarrow u(\mathbf{x}) + cx_n$ is not bounded on $\mathbb{R}_+^n$. There is a unique bounded solution and it is given by $u(\mathbf{x}) = (g * P_{x_n})(\check{\mathbf{x}})$ for $\mathbf{x} \in \mathbb{R}_+^n$. ♠

Theorem 9.5.13

Suppose that $g \in C(\partial\mathbb{R}_+^n)$ and vanishes at infinity. Then $u(\mathbf{x}) = (g * P_{x_n})(\check{\mathbf{x}})$ for $\mathbf{x} \in \mathbb{R}_+^n$ vanishes at infinity on $R = \{\mathbf{x} \in \mathbb{R}^n : x_n \geq 0\}$. Moreover, it is the unique solution that vanishes at infinity for the Dirichlet problem $\Delta u = 0$ on $\mathbb{R}_+^n$ with the condition $u = g$ on $\partial\mathbb{R}_+^n$.

Proof.

i) Suppose first that g has a compact support inside $B_a(\mathbf{0}) \subset \partial(\mathbb{R}_+^n) \cong \mathbb{R}^{n-1}$. Since $g \in L^1(\mathbb{R}^{n-1})$ and $|P_{x_n}(\check{\mathbf{x}})| \leq \frac{2}{\omega_n x_n^{n-1}}$ for $\check{\mathbf{x}} \in \mathbb{R}^{n-1}$ and $x_n > 0$, we have

$$\begin{aligned} |u((\check{\mathbf{x}}, x_n))| &= \left| \int_{\mathbb{R}^{n-1}} g(\check{\mathbf{y}}) P_{x_n}(\check{\mathbf{x}} - \check{\mathbf{y}}) d\check{\mathbf{y}} \right| \leq \int_{\mathbb{R}^{n-1}} |g(\check{\mathbf{y}})| |P_{x_n}(\check{\mathbf{x}} - \check{\mathbf{y}})| d\check{\mathbf{y}} \\ &\leq \left(\int_{\mathbb{R}^{n-1}} |g(\check{\mathbf{y}})| d\check{\mathbf{y}} \right) \frac{2}{\omega_n x_n^{n-1}} \rightarrow 0 \quad \text{as } x_n \rightarrow \infty \end{aligned}$$

uniformly for $\check{\mathbf{x}} \in \mathbb{R}^{n-1}$. Moreover, $|P_{x_n}(\check{\mathbf{x}})| \leq \frac{2x_n}{\omega_n \|\check{\mathbf{x}}\|^n}$ for $\check{\mathbf{x}} \neq \mathbf{0}$. Hence, for $\|\check{\mathbf{x}}\| > 2a$ and $0 < x_n \leq X$, where X is arbitrary but fixed, we have

$$\begin{aligned} |u((\check{\mathbf{x}}, x_n))| &= \left| \int_{\mathbb{R}^{n-1}} g(\check{\mathbf{y}}) P_{x_n}(\check{\mathbf{x}} - \check{\mathbf{y}}) d\check{\mathbf{y}} \right| = \left| \int_{B_a(\mathbf{0})} g(\check{\mathbf{y}}) P_{x_n}(\check{\mathbf{x}} - \check{\mathbf{y}}) d\check{\mathbf{y}} \right| \\ &\leq \int_{B_a(\mathbf{0})} |g(\check{\mathbf{y}})| |P_{x_n}(\check{\mathbf{x}} - \check{\mathbf{y}})| d\check{\mathbf{y}} \leq \left(\int_{B_a(\mathbf{0})} |g(\check{\mathbf{y}})| d\check{\mathbf{y}} \right) \sup_{\check{\mathbf{y}} \in B_a(\mathbf{0})} |P_{x_n}(\check{\mathbf{x}} - \check{\mathbf{y}})| \\ &\leq \left(\int_{B_a(\mathbf{0})} |g(\check{\mathbf{y}})| d\check{\mathbf{y}} \right) \frac{2X}{\omega_n (\|\check{\mathbf{x}}\|_2 - a)^n} \rightarrow 0 \quad \text{as } \|\check{\mathbf{x}}\| \rightarrow \infty \end{aligned}$$

uniformly for $0 < x_n \leq X$ because $a \leq \|\check{\mathbf{x}}\|_2 - a = \|\check{\mathbf{x}}\|_2 - \|\check{\mathbf{y}}\|_2 \leq \|\check{\mathbf{x}} - \check{\mathbf{y}}\|_2$ for all $\check{\mathbf{y}} \in B_a(\mathbf{0})$ and $\|\check{\mathbf{x}}\| > 2a$.

Combining these two results proved that u vanishes at infinity on $\mathbb{R}_+^n$ when g has a compact support.

ii) If g vanished at infinity but does not have a compact support, we choose functions $g_k : \mathbb{R}^{n-1} \rightarrow \mathbb{R}$ with compact supports such that $g_k \rightarrow g$ uniformly as $k \rightarrow \infty$. For instance, just multiply g by smooth functions $\psi_k : \mathbb{R}^{n-1} \rightarrow \mathbb{R}$ with compact support such that $\psi(\mathbf{y}) \leq 1$ for all $\mathbf{y} \in \mathbb{R}^{n-1}$ and $\psi(\mathbf{y}) = 1$ for $\mathbf{y} \in B_k(\mathbf{0}) \cap R$ with $k > 0$.

Let $u_k(\mathbf{x}) = (g_k * P_{x_n})(\check{\mathbf{x}})$ for $\mathbf{x} \in \mathbb{R}_+^n$. We have shown in (i) that u_k vanished at infinity on $\mathbb{R}_+^n$. Moreover, u_k converges uniformly on $\mathbb{R}_+^n$ to u because

$$\begin{aligned} \|u - u_k\|_\infty &= \sup_{\mathbf{x} \in \mathbb{R}_+^n} \left| \int_{\check{\mathbf{y}} \in \mathbb{R}^{n-1}} (g(\check{\mathbf{y}}) - g_k(\check{\mathbf{y}})) P_{x_n}(\check{\mathbf{y}} - \check{\mathbf{x}}) d\check{\mathbf{y}} \right| \\ &\leq \sup_{\check{\mathbf{y}} \in \mathbb{R}^{n-1}} |g(\check{\mathbf{y}}) - g_k(\check{\mathbf{y}})| \sup_{\mathbf{x} \in \mathbb{R}_+^n} \left| \int_{\check{\mathbf{y}} \in \mathbb{R}^{n-1}} P_{x_n}(\check{\mathbf{y}} - \check{\mathbf{x}}) d\check{\mathbf{y}} \right| \\ &= \sup_{\check{\mathbf{y}} \in \mathbb{R}^{n-1}} |g(\check{\mathbf{y}}) - g_k(\check{\mathbf{y}})| \sup_{\mathbf{x} \in \mathbb{R}_+^n} \left| \int_{\check{\mathbf{y}} \in \mathbb{R}^{n-1}} P_{x_n}(\check{\mathbf{y}}) d\check{\mathbf{y}} \right| = \sup_{\check{\mathbf{y}} \in \mathbb{R}^{n-1}} |g(\check{\mathbf{y}}) - g_k(\check{\mathbf{y}})| \rightarrow 0 \quad \text{as } k \rightarrow \infty, \end{aligned}$$

where we have used Lemma 9.5.10. One can now easily show that u also vanishes at infinity on $\mathbb{R}_+^n$.

Since u can be continuously expanded to R by $u = g$ on $\partial\mathbb{R}_+^n$ according to Theorem 9.5.11, we have that u also vanishes at infinity on R .

iii) Suppose that v_1 and v_2 are two solutions of the Dirichlet problem $\Delta u = 0$ on $\mathbb{R}_+^n$ with the condition $u = g$ on $\partial\mathbb{R}_+^n$. Suppose moreover that they are continuous on R and vanishing at infinity on R . Then $w = v_1 - v_2$ is a solution of $\Delta u = 0$ on $\mathbb{R}_+^n$ that vanishes at infinity on R and such that $w = 0$ on $\partial\mathbb{R}_+^n$.

Choose $\epsilon > 0$. Since w vanished at infinity, there exists $X_0 > 0$ such that $|w| < \epsilon$ on $R \setminus B_{X_0}(\mathbf{0}) \times [0, X]$, where $B_{X_0}(\mathbf{0}) \subset \mathbb{R}^{n-1}$. In particular, $|w| < \epsilon$ on $\partial(B_X(\mathbf{0}) \times [0, X])$ for $X > X_0$. It follows from the Corollary 9.2.5 to the Maximum Principle, that $|w| \leq \epsilon$ on $B_X(\mathbf{0}) \times [0, X]$. Thus $|w| \leq \epsilon$ on R . Since $\epsilon > 0$ is arbitrary, we get that $w = 0$ on R . ■

9.5.2 Dirichlet Problem in a Ball

Lemma 9.5.14

If $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $\mathbf{x} \neq \mathbf{0}$ and $\|\mathbf{y}\| = 1$, then $\|\mathbf{x} - \mathbf{y}\| = \|\|\mathbf{x}\|^{-1}\mathbf{x} - \|\mathbf{x}\|\mathbf{y}\|$.

Proof.

We have that

$$\begin{aligned} \|\mathbf{x} - \mathbf{y}\|^2 &= \langle \mathbf{x} - \mathbf{y}, \mathbf{x} - \mathbf{y} \rangle = \|\mathbf{x}\|^2 - 2\langle \mathbf{x}, \mathbf{y} \rangle + 1 = \|\|\mathbf{x}\|\mathbf{y}\|^2 - 2\langle \|\mathbf{x}\|^{-1}\mathbf{x}, \|\mathbf{x}\|\mathbf{y} \rangle + \|\|\mathbf{x}\|^{-1}\mathbf{x}\|^2 \\ &= \langle \|\mathbf{x}\|^{-1}\mathbf{x} - \|\mathbf{x}\|\mathbf{y}, \|\mathbf{x}\|^{-1}\mathbf{x} - \|\mathbf{x}\|\mathbf{y} \rangle = \|\|\mathbf{x}\|^{-1}\mathbf{x} - \|\mathbf{x}\|\mathbf{y}\|^2. \end{aligned}$$

■

Without loss of generality, we may limit our work to finding the Green function in the ball $\Omega = B_1(\mathbf{0})$.

$n > 2$) The **Green's function** on $B_1(\mathbf{0})$ is defined by $G(\mathbf{x}, \mathbf{y}) = N(\mathbf{x}, \mathbf{y}) - U(\mathbf{x}, \mathbf{y})$ for $\mathbf{x} \in B_1(\mathbf{0})$ and $\mathbf{y} \in \overline{B_1(\mathbf{0})}$ with $\mathbf{x} \neq \mathbf{y}$, where

$$U(\mathbf{x}, \mathbf{y}) = \begin{cases} N(\|\mathbf{x}\|^{-1}\mathbf{x}, \|\mathbf{x}\|\mathbf{y}) & \text{if } \mathbf{x} \neq \mathbf{0} \\ \frac{1}{(2-n)\omega_n} & \text{if } \mathbf{x} = \mathbf{0} \end{cases}$$

$$= \begin{cases} \frac{1}{(2-n)\omega_n} \|\|\mathbf{x}\|^{-1}\mathbf{x} - \|\mathbf{x}\|\mathbf{y}\|^{2-n} & \text{if } \mathbf{x} \neq \mathbf{0} \\ \frac{1}{(2-n)\omega_n} & \text{if } \mathbf{x} = \mathbf{0} \end{cases}$$

for $\mathbf{x} \in B_1(\mathbf{0})$ and $\mathbf{y} \in \overline{B_1(\mathbf{x})}$. We note that

$$\frac{1}{(2-n)\omega_n} \|\|\mathbf{x}\|^{-1}\mathbf{x} - \|\mathbf{x}\|\mathbf{y}\|^{2-n} \rightarrow \frac{1}{(2-n)\omega_n}$$

uniformly for $\mathbf{y} \in \overline{B_1(\mathbf{0})}$ as $\mathbf{x} \rightarrow \mathbf{0}$ because

$$\begin{aligned} 1 - \|\mathbf{x}\| &\leq 1 - \|\mathbf{x}\|\|\mathbf{y}\| = \|\|\mathbf{x}\|^{-1}\mathbf{x}\| - \|\mathbf{x}\|\mathbf{y}\| \leq \|\|\mathbf{x}\|^{-1}\mathbf{x} - \|\mathbf{x}\|\mathbf{y}\| \\ &\leq \|\|\mathbf{x}\|^{-1}\mathbf{x}\| + \|\mathbf{x}\|\mathbf{y}\| = 1 + \|\mathbf{x}\|\mathbf{y}\| \leq 1 + \|\mathbf{x}\| \end{aligned}$$

for $\mathbf{y} \in \overline{B_1(\mathbf{0})}$ and $\mathbf{x} \in B_1(\mathbf{0})$.

The function $\mathbf{y} \mapsto U(\mathbf{x}, \mathbf{y})$ with $\mathbf{x} \neq \mathbf{0}$ is harmonic in $B_1(\mathbf{0})$ and continuous on $\overline{B_1(\mathbf{0})}$ because $\|\mathbf{x}\|^{-2}\mathbf{x} \notin \overline{B_1(\mathbf{0})}$ for all $\mathbf{x} \in B_1(\mathbf{0}) \setminus \{\mathbf{0}\}$ (Figure 9.3) and so $\|\mathbf{x}\|^{-1}\mathbf{x} - \|\mathbf{x}\|\mathbf{y} \neq \mathbf{0}$ for all $\mathbf{y} \in B_1(\mathbf{0})$. The previous statement about $\mathbf{y} \mapsto U(\mathbf{x}, \mathbf{y})$ is obviously true if $\mathbf{x} = \mathbf{0}$.

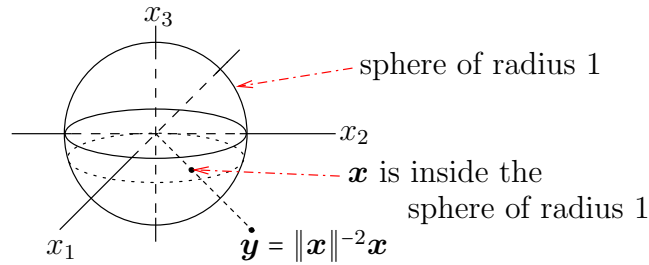


Figure 9.3: $\mathbf{y} = \|\mathbf{x}\|^{-2}\mathbf{x}$ is the reflection through the sphere of radius 1 of $\mathbf{x}$ because $\langle \mathbf{x}, \mathbf{y} \rangle = 1$, the radius of the sphere. Since $\|\mathbf{x}\| < 1$, we have that $\|\mathbf{y}\| = \|\mathbf{x}\|^{-1} > 1$.

We get from Lemma 9.5.14 that

$$U(\mathbf{x}, \mathbf{y}) = \frac{1}{(2-n)\omega_n} \|\|\mathbf{x}\|^{-1}\mathbf{x} - \|\mathbf{x}\|\mathbf{y}\|^{2-n} = \frac{1}{(2-n)\omega_n} \|\mathbf{x} - \mathbf{y}\|^{2-n} = N(\mathbf{x}, \mathbf{y})$$

for $\mathbf{y} \in \partial B_1(\mathbf{0})$ and $\mathbf{x} \in B_1(\mathbf{0}) \setminus \{\mathbf{0}\}$. For $\mathbf{y} \in \partial B_1(\mathbf{0})$ and $\mathbf{x} = \mathbf{0}$, we have

$$U(\mathbf{x}, \mathbf{y}) = \frac{1}{(2-n)\omega_n} = \frac{1}{(2-n)\omega_n} \|\mathbf{y}\|^{2-n} = N(\mathbf{x}, \mathbf{y}).$$

Explicitly, the Green's function for $B_1(\mathbf{0})$ is

$$\begin{aligned} G(\mathbf{x}, \mathbf{y}) &= \begin{cases} N(\mathbf{x}, \mathbf{y}) - N(\|\mathbf{x}\|^{-1}\mathbf{x}, \|\mathbf{x}\|\mathbf{y}) & \text{if } \mathbf{x} \neq \mathbf{0} \\ N(\mathbf{y}) - \frac{1}{(2-n)\omega_n} & \text{if } \mathbf{x} = \mathbf{0} \end{cases} \\ &= \begin{cases} \frac{1}{(2-n)\omega_n} \left(\|\mathbf{x} - \mathbf{y}\|^{2-n} - \|\|\mathbf{x}\|^{-1}\mathbf{x} - \|\mathbf{x}\|\mathbf{y}\|^{2-n} \right) & \text{if } \mathbf{x} \neq \mathbf{0} \\ \frac{1}{(2-n)\omega_n} \|\mathbf{y}\|^{2-n} - \frac{1}{(2-n)\omega_n} & \text{if } \mathbf{x} = \mathbf{0} \end{cases} \end{aligned}$$

for $\mathbf{x} \in B_1(\mathbf{0})$ and $\mathbf{y} \in \overline{B_1(\mathbf{0})}$ with $\mathbf{x} \neq \mathbf{y}$.

n = 2 The **Green's function** on $B_1(\mathbf{0})$ is defined by $G(\mathbf{x}, \mathbf{y}) = N(\mathbf{x}, \mathbf{y}) - U(\mathbf{x}, \mathbf{y})$ for $\mathbf{x} \in B_1(\mathbf{0})$ and $\mathbf{y} \in \overline{B_1(\mathbf{0})}$ with $\mathbf{x} \neq \mathbf{y}$, where

$$\begin{aligned} U(\mathbf{x}, \mathbf{y}) &= \begin{cases} N(\|\mathbf{x}\|^{-1}\mathbf{x}, \|\mathbf{x}\|\mathbf{y}) & \text{if } \mathbf{x} \neq \mathbf{0} \\ 0 & \text{if } \mathbf{x} = \mathbf{0} \end{cases} \\ &= \begin{cases} \frac{1}{(2-n)\omega_n} \ln(\|\|\mathbf{x}\|^{-1}\mathbf{x} - \|\mathbf{x}\|\mathbf{y}\|) & \text{if } \mathbf{x} \neq \mathbf{0} \\ 0 & \text{if } \mathbf{x} = \mathbf{0} \end{cases} \end{aligned}$$

for $\mathbf{x} \in B_1(\mathbf{0})$ and $\mathbf{y} \in \overline{B_1(\mathbf{x})}$. For a reason similar to the one given in the case $n > 2$, $\|\|\mathbf{x}\|^{-1}\mathbf{x} - \|\mathbf{x}\|\mathbf{y}\| \rightarrow 1$ uniformly for $\mathbf{y} \in \overline{B_1(\mathbf{x})}$ as $\mathbf{x} \rightarrow \mathbf{0}$ and so $\ln(\|\|\mathbf{x}\|^{-1}\mathbf{x} - \|\mathbf{x}\|\mathbf{y}\|) \rightarrow 0$ uniformly for $\mathbf{y} \in \overline{B_1(\mathbf{x})}$ as $\mathbf{x} \rightarrow \mathbf{0}$.

The function $\mathbf{y} \mapsto U(\mathbf{x}, \mathbf{y})$ is harmonic in $B_1(\mathbf{0})$ and continuous on $\overline{B_1(\mathbf{0})}$ because $\|\mathbf{x}\|^{-1}\mathbf{x} \notin B_1(\mathbf{0})$ for all $\mathbf{x} \in B_1(\mathbf{0}) \setminus \{\mathbf{0}\}$ (Figure 9.3) and so $\|\mathbf{x}\|^{-1}\mathbf{x} - \|\mathbf{x}\|\mathbf{y} \neq \mathbf{0}$ for all $\mathbf{y} \in \overline{B_1(\mathbf{0})}$. The previous statement about $\mathbf{y} \mapsto U(\mathbf{x}, \mathbf{y})$ is obviously true if $\mathbf{x} = \mathbf{0}$.

We get from Lemma 9.5.14 that

$$U(\mathbf{x}, \mathbf{y}) = \frac{1}{\omega_2} \ln(\|\mathbf{x}\|^{-1}\mathbf{x} - \|\mathbf{x}\|\mathbf{y}) = \frac{1}{\omega_2} \ln(\|\mathbf{x} - \mathbf{y}\|) = N(\mathbf{x}, \mathbf{y})$$

for $\mathbf{y} \in \partial B_1(\mathbf{0})$ and $\mathbf{x} \in B_1(\mathbf{0}) \setminus \{\mathbf{0}\}$. For $\mathbf{y} \in \partial B_1(\mathbf{0})$ and $\mathbf{x} = \mathbf{0}$, we have

$$N(\mathbf{x}, \mathbf{y}) = \frac{1}{\omega_2} \ln(\|\mathbf{y}\|) = 0 = U(\mathbf{x}, \mathbf{y})$$

because $\|\mathbf{y}\| = 1$.

Explicitly, the Green's function on $B_1(\mathbf{0})$ is

$$G(\mathbf{x}, \mathbf{y}) = \begin{cases} N(\mathbf{x}, \mathbf{y}) - N(\|\mathbf{x}\|^{-1}\mathbf{x}, \|\mathbf{x}\|\mathbf{y}) & \text{if } \mathbf{x} \neq \mathbf{0} \\ N(\mathbf{y}) & \text{if } \mathbf{x} = \mathbf{0} \end{cases}$$

$$= \begin{cases} \frac{1}{\omega_2} (\ln(\|\mathbf{x} - \mathbf{y}\|) - \ln(\|\|\mathbf{x}\|^{-1}\mathbf{x} - \|\mathbf{x}\|\mathbf{y}\|)) & \text{if } \mathbf{x} \neq \mathbf{0} \\ \frac{1}{\omega_2} \ln(\|\mathbf{y}\|) & \text{if } \mathbf{x} = \mathbf{0} \end{cases}$$

for $\mathbf{x} \in B_1(\mathbf{0})$ and $\mathbf{y} \in \overline{B_1(\mathbf{0})}$ with $\mathbf{x} \neq \mathbf{y}$.

Remark 9.5.15

It can be directly proved that $G(\mathbf{x}, \mathbf{y}) = G(\mathbf{y}, \mathbf{x})$ for $\mathbf{x}$ and $\mathbf{y}$ in $B_1(\mathbf{0})$ using an identity similar to the one in Lemma 9.5.14. $\spadesuit$

According to Theorem 9.5.5, the solution of $\Delta u = f$ on $B_1(\mathbf{0})$ with $u = 0$ on $\partial B_1(\mathbf{0})$ is given by

$$u(\mathbf{x}) = \int_{B_1(\mathbf{0})} G(\mathbf{x}, \mathbf{y}) f(\mathbf{y}) d\mathbf{y}$$

provide that $f \in L^1(B_1(\mathbf{0}))$.

According to Theorem 9.5.7 and provide that $g \in C(\partial B_1(\mathbf{0}))$, the solution of $\Delta u = 0$ on $B_1(\mathbf{0})$ with $u = g$ on $\partial B_1(\mathbf{0})$ is given by

$$u(\mathbf{x}) = \begin{cases} \int_{\partial B_1(\mathbf{0})} g(\mathbf{y}) P(\mathbf{x}, \mathbf{y}) dS_{\mathbf{y}} & \text{if } \mathbf{x} \in B_1(\mathbf{0}) \\ g(\mathbf{x}) & \text{if } \mathbf{x} \in \partial B_1(\mathbf{0}) \end{cases}$$

where $P(\mathbf{x}, \mathbf{y}) = \frac{\partial G_{\mathbf{x}}}{\partial \nu}(\mathbf{y})$ with $G_{\mathbf{x}}(\mathbf{y}) = G(\mathbf{x}, \mathbf{y})$ for all $\mathbf{y} \in \overline{B_1(\mathbf{0})} \setminus \{\mathbf{x}\}$. Moreover, $u \in C(\overline{B_1(\mathbf{0})})$.

We now show that the **Poisson kernel** for $B_1(\mathbf{0})$ is

$$P(\mathbf{x}, \mathbf{y}) = \frac{1 - \|\mathbf{x}\|^2}{\omega_n \|\mathbf{x} - \mathbf{y}\|^n} \quad (9.5.9)$$

for $\mathbf{x} \in B_1(\mathbf{0})$ and $\mathbf{y} \in \partial B_1(\mathbf{0})$.

For $n > 2$, we find using Lemma 9.5.14 that the Poisson kernel is

$$P(\mathbf{x}, \mathbf{y}) = \frac{\partial G_{\mathbf{x}}}{\partial \nu}(\mathbf{y}) = \nabla_{\mathbf{y}} G_{\mathbf{x}}(\mathbf{y}) \cdot \mathbf{y}$$

$$= -\frac{1}{\omega_n} \left(\frac{1}{\|\mathbf{x} - \mathbf{y}\|^n} (\mathbf{x} - \mathbf{y}) \cdot \mathbf{y} - \frac{\|\mathbf{x}\|}{\|\|\mathbf{x}\|^{-1}\mathbf{x} - \|\mathbf{x}\|\mathbf{y}\|^n} (\|\mathbf{x}\|^{-1}\mathbf{x} - \|\mathbf{x}\|\mathbf{y}) \cdot \mathbf{y} \right) = \frac{1 - \|\mathbf{x}\|^2}{\omega_n \|\mathbf{x} - \mathbf{y}\|^n}$$

for $\mathbf{x} \in B_1(\mathbf{0}) \setminus \{\mathbf{0}\}$ and $\mathbf{y} \in \partial B_1(\mathbf{0})$. Similarly, for $\mathbf{y} \in \partial B_1(\mathbf{0})$, we have

$$P(\mathbf{0}, \mathbf{y}) = \nabla_{\mathbf{y}} G_{\mathbf{0}}(\mathbf{y}) \cdot \mathbf{y} = \frac{1}{\omega_n} \frac{1}{\|\mathbf{y}\|^n} \mathbf{y} \cdot \mathbf{y} = \frac{1}{\omega_n}$$

because $\|\mathbf{y}\| = 1$. Thus (9.5.9) is true for $n > 2$ and all $\mathbf{x} \in B_1(\mathbf{0})$. Proceeding as we have just done for $n > 2$, we can show that (9.5.9) is also true for $n = 2$ and all $\mathbf{x} \in B_1(\mathbf{0})$.

As we have done in the case of the half plane, since Theorem 9.5.7 has not been proved yet, we now prove that this theorem is true in the present case. More precisely, we prove the following theorem.

Theorem 9.5.16

Suppose that $g \in L^1(\partial B_1(\mathbf{0}))$. Then

$$u(\mathbf{x}) = \int_{\partial B_1(\mathbf{0})} g(\mathbf{y}) P(\mathbf{x}, \mathbf{y}) dS_{\mathbf{y}} = \int_{\partial B_1(\mathbf{0})} g(\mathbf{y}) \frac{1 - \|\mathbf{x}\|^2}{\omega_n \|\mathbf{x} - \mathbf{y}\|^n} dS_{\mathbf{y}}$$

is well defined for $\mathbf{x} \in B_1(\mathbf{0})$. Moreover, u is harmonic in $B_1(\mathbf{0})$.

If $g \in C(\partial B_1(\mathbf{0}))$, then u can be continuously extended to $\overline{B_1(\mathbf{0})}$ by $u(\mathbf{x}) = g(\mathbf{x})$ for $\mathbf{x} \in \partial B_1(\mathbf{0})$.

If $g \in L^p(\partial B_1(\mathbf{0}))$ for some integer $p \geq 1$, then $u_r(\mathbf{x}) \equiv u(r\mathbf{x})$ for $0 \leq r < 1$ and $\mathbf{x} \in \partial B_1(\mathbf{0})$ satisfies $u_r \rightarrow g$ in $L^p(\partial B_1(\mathbf{0}))$ as $r \rightarrow 1^-$.

Remark 9.5.17

To solve the Dirichlet problem $\Delta u = 0$ on $B_\rho(\mathbf{x}_0)$ with $u = g$ on $\partial B_\rho(\mathbf{x}_0)$, we need to consider $\tilde{u}(\tilde{\mathbf{x}}) = u(\rho\tilde{\mathbf{x}} + \mathbf{x}_0)$ for $\tilde{\mathbf{x}} \in B_1(\mathbf{0})$ and $\tilde{g}(\tilde{\mathbf{x}}) = g(\rho\tilde{\mathbf{x}} + \mathbf{x}_0)$ for $\tilde{\mathbf{x}} \in \partial B_1(\mathbf{0})$ to transform the previous problem into $\Delta \tilde{u} = 0$ on $B_1(\mathbf{0})$ with $\tilde{u} = \tilde{g}$ on $\partial B_1(\mathbf{0})$. The solution of this problem is

$$\tilde{u}(\tilde{\mathbf{x}}) = \int_{\partial B_1(\mathbf{0})} \tilde{g}(\tilde{\mathbf{y}}) P(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) dS_{\tilde{\mathbf{y}}} = \int_{\partial B_1(\mathbf{0})} \tilde{g}(\tilde{\mathbf{y}}) \frac{1 - \|\tilde{\mathbf{x}}\|^2}{\omega_n \|\tilde{\mathbf{x}} - \tilde{\mathbf{y}}\|^n} dS_{\tilde{\mathbf{y}}}$$

for $\tilde{\mathbf{x}} \in B_1(\mathbf{0})$. If we restate this formula in terms of u and g , we get

$$u(\rho\tilde{\mathbf{x}} + \mathbf{x}_0) = \int_{\partial B_1(\mathbf{0})} g(\rho\tilde{\mathbf{y}} + \mathbf{x}_0) P(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) dS_{\tilde{\mathbf{y}}} = \int_{\partial B_1(\mathbf{0})} g(\rho\tilde{\mathbf{y}} + \mathbf{x}_0) \frac{1 - \|\tilde{\mathbf{x}}\|^2}{\omega_n \|\tilde{\mathbf{x}} - \tilde{\mathbf{y}}\|^n} dS_{\tilde{\mathbf{y}}}$$

for $\tilde{\mathbf{x}} \in B_1(\mathbf{0})$. Finally, the substitution $\tilde{\mathbf{x}} = (\mathbf{x} - \mathbf{x}_0)/\rho$ and $\tilde{\mathbf{y}} = (\mathbf{y} - \mathbf{x}_0)/\rho$ yields the solution

$$\begin{aligned} u(\mathbf{x}) &= \int_{\partial B_\rho(\mathbf{x}_0)} g(\mathbf{y}) P\left(\frac{\mathbf{x} - \mathbf{x}_0}{\rho}, \frac{\mathbf{y} - \mathbf{x}_0}{\rho}\right) \rho^{1-n} dS_{\mathbf{y}} \\ &= \int_{\partial B_\rho(\mathbf{x}_0)} g(\mathbf{y}) \frac{1 - \|(\mathbf{x} - \mathbf{x}_0)/\rho\|^2}{\omega_n \|(\mathbf{x} - \mathbf{x}_0)/\rho - (\mathbf{y} - \mathbf{x}_0)/\rho\|^n} \rho^{1-n} dS_{\mathbf{y}} \\ &= \int_{\partial B_\rho(\mathbf{x}_0)} g(\mathbf{y}) \frac{\rho^2 - \|\mathbf{x} - \mathbf{x}_0\|^2}{\rho \omega_n \|\mathbf{x} - \mathbf{y}\|^n} dS_{\mathbf{y}} \end{aligned} \quad (9.5.10)$$

for $\mathbf{x} \in B_\rho(\mathbf{x}_0)$ for the original Dirichlet problem on $B_\rho(\mathbf{x}_0)$. ♠

Before giving the proof of the previous theorem, we need a lemma.

Lemma 9.5.18

We have

$$\int_{\partial B_1(\mathbf{0})} P(\mathbf{x}, \mathbf{y}) dS_{\mathbf{y}} = 1 \quad , \quad \mathbf{x} \in B_1(\mathbf{0}) . \quad (9.5.11)$$

Moreover, given $\epsilon > 0$,

$$\int_{\partial B_1(\mathbf{0}) \setminus B_\epsilon(\mathbf{x}_0)} P(r\mathbf{x}_0, \mathbf{y}) dS_{\mathbf{y}} \rightarrow 0 \quad \text{as } r \rightarrow 0$$

uniformly for $\mathbf{x}_0 \in \partial B_1(\mathbf{0})$.

Proof.

Since $\mathbf{x} \rightarrow P(\mathbf{x}, \mathbf{y})$ is harmonic in $B_1(\mathbf{0})$ for all $\mathbf{y} \in \partial B_1(\mathbf{0})$, we get from the Main Value Theorem, Theorem 9.1.4, that

$$P(\mathbf{0}, \mathbf{y}) = \frac{1}{\omega_n} \int_{\partial B_1(\mathbf{0})} P(r\mathbf{x}, \mathbf{y}) dS_{\mathbf{x}}$$

for $\mathbf{y} \in \partial B_1(\mathbf{0})$ and $0 < r < 1$. Since $P(\mathbf{0}, \mathbf{y}) = 1/\omega_n$ for all $\mathbf{y} \in \partial B_1(\mathbf{0})$, we get

$$1 = \int_{\partial B_1(\mathbf{0})} P(r\mathbf{x}, \mathbf{y}) dS_{\mathbf{x}}$$

for $\mathbf{y} \in \partial B_1(\mathbf{0})$ and $0 < r < 1$. From Lemma 9.5.14, we have that $\|r\mathbf{x} - \mathbf{y}\| = \|r^{-1}(r\mathbf{x}) - r\mathbf{y}\| = \|\mathbf{x} - r\mathbf{y}\|$ for $\mathbf{y} \in \partial B_1(\mathbf{0})$. Therefore, $P(r\mathbf{x}, \mathbf{y}) = P(r\mathbf{y}, \mathbf{x})$ for $\mathbf{x}$ and $\mathbf{y}$ in $\partial B_1(\mathbf{0})$. It follows that

$$1 = \int_{\partial B_1(\mathbf{0})} P(r\mathbf{y}, \mathbf{x}) dS_{\mathbf{x}}$$

for $\mathbf{y} \in \partial B_1(\mathbf{0})$ and $0 < r < 1$. Since all elements of $B_1(\mathbf{0})$ are of the form $r\mathbf{y}$ for $0 \leq r < 1$ and $\mathbf{y} \in \partial B_1(\mathbf{0})$, we get (9.5.11).

To prove the second part of the lemma, we note that

$$P(r\mathbf{x}_0, \mathbf{y}) = \frac{1 - \|r\mathbf{x}_0\|^2}{\omega_n \|r\mathbf{x}_0 - \mathbf{y}\|^n} \leq \frac{1 - |r|^2}{\omega_n q^n}$$

for all $\mathbf{x}_0 \in \partial B_1(\mathbf{0})$ and $\mathbf{y} \in \partial B_1(\mathbf{0}) \setminus B_\epsilon(\mathbf{x}_0)$, where

$$q = \min \{ \|r\mathbf{x}_0 - \mathbf{y}\| : r > 0, \mathbf{y} \in \partial B_1(\mathbf{0}) \setminus B_\epsilon(\mathbf{x}_0) \} > 0$$

is the distance between the region $\partial B_1(\mathbf{0}) \setminus B_\epsilon(\mathbf{x}_0)$ and the segment $\{r\mathbf{x}_0 : r > 0\}$ as it can be seen in Figure 9.4. Let $M = \int_{\partial B_1(\mathbf{0}) \setminus B_\epsilon(\mathbf{x}_0)} dS$. M is the surface area of $\partial B_1(\mathbf{0}) \setminus B_\epsilon(\mathbf{x}_0)$. By symmetry, it is independent of the $\mathbf{x}_0$ considered.

Given $\delta > 0$, choose r_0 closed to 1 such that $0 \leq P(r\mathbf{x}_0, \mathbf{y}) < \delta/M$ for all $\mathbf{x}_0 \in \partial B_1(\mathbf{0})$, $\mathbf{y} \in \partial B_1(\mathbf{0}) \setminus B_\epsilon(\mathbf{x}_0)$, and $r_0 < r < 1$. Then

$$0 \leq \int_{\partial B_1(\mathbf{0}) \setminus B_\epsilon(\mathbf{x}_0)} P(r\mathbf{x}_0, \mathbf{y}) dS_{\mathbf{y}} < \delta$$

for all $\mathbf{x}_0 \in \partial B_1(\mathbf{0})$ and $r_0 < r < 1$. ■

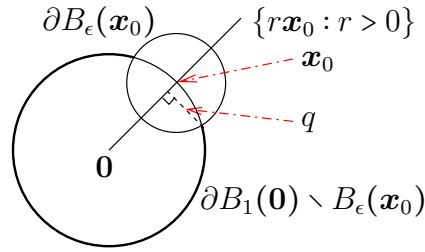


Figure 9.4: q is the distance between the region $\partial B_1(\mathbf{0}) \setminus B_\epsilon(\mathbf{x}_0)$ and the segment $\{r\mathbf{x}_0 : r > 0\}$ used in the proof of Lemma 9.5.18.

Proof of Theorem 9.5.16.

i) Since $\mathbf{y} \rightarrow P(\mathbf{x}, \mathbf{y})$ is continuous on $\partial B_1(\mathbf{0})$ for all $\mathbf{x} \in B_1(\mathbf{0})$ and $\partial B_1(\mathbf{0})$ is compact, then $\mathbf{y} \rightarrow P(\mathbf{x}, \mathbf{y})$ is bounded on $\partial B_1(\mathbf{0})$ and

$$|u(\mathbf{x})| \leq \sup_{\mathbf{y} \in \partial B_1(\mathbf{0})} |P(\mathbf{x}, \mathbf{y})| \int_{\partial B_1(\mathbf{0})} |g(\mathbf{y})| \, dS_{\mathbf{y}} < \infty$$

for all $\mathbf{x} \in B_1(\mathbf{0})$. Thus u is well defined on $B_1(\mathbf{0})$.

ii) Proceeding as we did in (i), we can show that, for any $0 \leq a < b < 1$, there exists a constant C such that $|g(\mathbf{y}) D_{\mathbf{x}}^{\alpha} P(\mathbf{x}, \mathbf{y})| \leq |g(\mathbf{y})|$ for $a \leq \|\mathbf{x}\| \leq b$ and $\mathbf{y} \in \partial B_1(\mathbf{0})$, and any multi-index $\alpha \in \mathbb{N}^n$ with $|\alpha| \leq 2$. Thus, $\mathbf{y} \mapsto g(\mathbf{y}) D_{\mathbf{x}}^{\alpha} P(\mathbf{x}, \mathbf{y})$ is bounded in absolute value by a function in $L^1(\partial B_1(\mathbf{0}))$ for $a \leq \|\mathbf{x}\| \leq b$ and $\mathbf{y} \in \partial B_1(\mathbf{0})$, and any multi-index $\alpha \in \mathbb{N}^n$ with $|\alpha| < 2$. We may therefore interchange the integral with respect to $\mathbf{y}$ that defined u and the differential operator $\Delta_{\mathbf{x}}$ (see [7]). We get

$$\Delta_{\mathbf{x}} u(\mathbf{x}) = \int_{\partial B_1(\mathbf{0})} \Delta_{\mathbf{x}} P(\mathbf{x}, \mathbf{y}) g(\mathbf{y}) \, dS_{\mathbf{y}} = 0$$

for $\mathbf{x} \in B_1(\mathbf{0})$ because $\mathbf{x} \mapsto P(\mathbf{x}, \mathbf{y})$ is harmonic in $B_1(\mathbf{0})$ for all $\mathbf{y} \in \partial B_1(\mathbf{0})$.

iii) Suppose that g is continuous on the compact set $\partial B_1(\mathbf{0})$, then g is uniformly continuous on $\partial B_1(\mathbf{0})$. Given $\epsilon > 0$, choose $\delta > 0$ such that $|g(\mathbf{y}) - g(\mathbf{x})| < \epsilon/2$ for $\|\mathbf{x} - \mathbf{y}\| < \delta$ and $\mathbf{x}, \mathbf{y} \in \partial B_1(\mathbf{0})$.

Let

$$V_{\mathbf{x}} = \partial B_1(\mathbf{0}) \cap B_{\delta}(\mathbf{x}) = \{\mathbf{y} \in \partial B_1(\mathbf{0}) : \|\mathbf{x} - \mathbf{y}\| < \delta\}$$

for all $\mathbf{x} \in \partial B_1(\mathbf{0})$. We have $|g(\mathbf{y}) - g(\mathbf{x})| < \epsilon/2$ for all $\mathbf{y} \in V_{\mathbf{x}}$. According to Lemma 9.5.18, there exists $r_0 < 1$ closed to 1 such that

$$\int_{\partial B_1(\mathbf{0}) \setminus V_{\mathbf{x}}} P(r\mathbf{x}, \mathbf{y}) \, dS_{\mathbf{y}} < \frac{\epsilon}{4\|g\|_{\infty}}$$

for all $\mathbf{x} \in \partial B_1(\mathbf{0})$ and $r_0 < r < 1$, where $\|g\|_{\infty} = \max_{\mathbf{y} \in \partial B_1(\mathbf{0})} |g(\mathbf{y})|$. We assume that $g \neq 0$ on $\partial B_1(\mathbf{0})$ otherwise there is nothing to prove.

Using (9.5.11), we have

$$g(\mathbf{x}) - u(r\mathbf{x}) = \int_{\partial B_1(\mathbf{0})} (g(\mathbf{x}) - g(\mathbf{y})) P(r\mathbf{x}, \mathbf{y}) dS_{\mathbf{y}}$$

for all $\mathbf{x} \in \partial B_1(\mathbf{0})$. Thus,

$$\begin{aligned} |g(\mathbf{x}) - u(r\mathbf{x})| &= \left| \int_{\partial B_1(\mathbf{0})} (g(\mathbf{x}) - g(\mathbf{y})) P(r\mathbf{x}, \mathbf{y}) dS_{\mathbf{y}} \right| \\ &\leq \int_{V_{\mathbf{x}}} |g(\mathbf{x}) - g(\mathbf{y})| P(r\mathbf{x}, \mathbf{y}) dS_{\mathbf{y}} + \int_{\partial B_1(\mathbf{0}) \setminus V_{\mathbf{x}}} |g(\mathbf{x}) - g(\mathbf{y})| P(r\mathbf{x}, \mathbf{y}) dS_{\mathbf{y}} \\ &< \frac{\epsilon}{2} \int_{V_{\mathbf{x}}} P(r\mathbf{x}, \mathbf{y}) dS_{\mathbf{y}} + 2\|g\|_{\infty} \int_{\partial B_1(\mathbf{0}) \setminus V_{\mathbf{x}}} P(r\mathbf{x}, \mathbf{y}) dS_{\mathbf{y}} \leq \frac{\epsilon}{2} + 2\|g\|_{\infty} \frac{\epsilon}{4\|g\|_{\infty}} < \epsilon \end{aligned}$$

for all $\mathbf{x} \in \partial B_1(\mathbf{0})$ and $r_0 < r < 1$, where again we have used (9.5.11) to obtain

$$\int_{V_{\mathbf{x}}} P(r\mathbf{x}, \mathbf{y}) dS_{\mathbf{y}} \leq \int_{\partial B_1(\mathbf{0})} P(r\mathbf{x}, \mathbf{y}) dS_{\mathbf{y}} = 1 .$$

Hence

$$\begin{aligned} u_r : \partial B_1(\mathbf{0}) &\rightarrow \mathbb{R} \\ \mathbf{x} &\mapsto u(r\mathbf{x}) \end{aligned}$$

converges uniformly on $\partial B_1(\mathbf{0})$ to g as $r \rightarrow 1$. We may then extend u continuously to $\overline{B_1(\mathbf{0})}$ with $u(\mathbf{x}) = g(\mathbf{x})$ for $\mathbf{x} \in \partial B_1(\mathbf{0})$.

iv) Suppose that $g \in L^p(\partial B_1(\mathbf{0}))$ with $1 \leq p < \infty$. Given $\epsilon > 0$, choose $f \in C(\partial B_1(\mathbf{0}))$ such that

$$\|g - f\|_p < \epsilon/3 , \quad (9.5.12)$$

where $\|\cdot\|_p$ is the L^p -norm on $L^p(\partial B_1(\mathbf{0}))$. This is possible because $C(\partial B_1(\mathbf{0}))$ is dense in $L^p(\partial B_1(\mathbf{0}))$.

Let

$$v(\mathbf{x}) = \int_{\partial B_1(\mathbf{0})} f(\mathbf{y}) P(\mathbf{x}, \mathbf{y}) dS_{\mathbf{y}} , \quad \mathbf{x} \in B_1(\mathbf{0}) .$$

Then

$$\|g - u_r\|_p \leq \|g - f\|_p + \|f - v_r\|_p + \|v_r - u_r\|_p , \quad (9.5.13)$$

where v_r is defined by $v_r(\mathbf{x}) = v(r\mathbf{x})$ for all $\mathbf{x} \in \partial B_1(\mathbf{0})$ and $0 \leq r < 1$.

As seen in (iii), for r closed enough to 1, we have $\|f - v_r\|_{\infty} < \frac{\epsilon}{3B}$, where $B^p = \int_{\partial B_1(\mathbf{0})} dS$.

Hence

$$\|f - v_r\|_p < \frac{\epsilon}{3} . \quad (9.5.14)$$

The linear application

$$\begin{aligned} T_r : L^p(\partial B_1(\mathbf{0})) &\rightarrow L^p(\partial B_1(\mathbf{0})) \\ h &\mapsto \int_{\partial B_1(\mathbf{0})} h(\mathbf{y}) P(r\mathbf{x}, \mathbf{y}) dS_{\mathbf{y}} \end{aligned}$$

has norm less than or equal to 1 according to the Generalized Young's Inequality, Theorem 6.5.1, because

$$\int_{\partial B_1(\mathbf{0})} P(r\mathbf{x}, \mathbf{y}) dS_{\mathbf{y}} = 1$$

for $\mathbf{x} \in \partial B_1(\mathbf{0})$ and $0 \leq r < 1$, and

$$\int_{\partial B_1(\mathbf{0})} P(r\mathbf{x}, \mathbf{y}) dS_{\mathbf{x}} = \int_{\partial B_1(\mathbf{0})} P(r\mathbf{y}, \mathbf{x}) dS_{\mathbf{x}} = 1$$

for $\mathbf{y} \in \partial B_1(\mathbf{0})$ and $0 \leq r < 1$. Note that it follows from Lemma 9.5.14 that $P(r\mathbf{x}, \mathbf{y}) = P(r\mathbf{y}, \mathbf{x})$ for $\mathbf{x}$ and $\mathbf{y}$ in $\partial B_1(\mathbf{0})$. Therefore, $\|u_r - v_r\|_p = \|T_r(g) - T_r(f)\|_p \leq \|g - f\|_p$. It follows from (9.5.12) that

$$\|u_r - v_r\|_p < \frac{\epsilon}{3}. \quad (9.5.15)$$

Finally, (9.5.12), (9.5.13), (9.5.14) and (9.5.15) implies that $\|g - u_r\|_p < \epsilon$ for r closed enough to 1. ■

Some results about harmonic functions, which are given in a course on complex analysis, generalize to dimensions higher than two.

Theorem 9.5.19 (Harnack's Inequality)

Suppose that $\Omega \subset \mathbb{R}^n$ is an open set and that $u \in C^2(\Omega) \cap C^1(\overline{\Omega})$ is an harmonic function in Ω such that $u(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in \Omega$. Then, given $\mathbf{x}_0 \in \Omega$ and $\rho > 0$ such that $B_\rho(\mathbf{x}_0) \subset \Omega$, we have

$$\frac{\rho^{n-2}(\rho - r)}{(\rho + r)^{n-1}} u(\mathbf{x}_0) \leq u(\mathbf{x}) \leq \frac{\rho^{n-2}(\rho + r)}{(\rho - r)^{n-1}} u(\mathbf{x}_0)$$

for $\mathbf{x} \in B_\rho(\mathbf{x}_0)$ and $r = \|\mathbf{x} - \mathbf{x}_0\|$.

Proof.

Since $\|\mathbf{x} - \mathbf{y}\| \leq \rho + r$ for $\mathbf{x} \in B_\rho(\mathbf{x}_0)$ and $\mathbf{y} \in \partial B_\rho(\mathbf{x}_0)$, we get

$$\frac{\rho - r}{(\rho + r)^{n-1}} = \frac{\rho^2 - r^2}{(\rho + r)^n} \leq \frac{\rho^2 - r^2}{\|\mathbf{x} - \mathbf{y}\|^n}$$

for $\mathbf{x} \in B_\rho(\mathbf{x}_0)$ and $\mathbf{y} \in \partial B_\rho(\mathbf{x}_0)$. Similarly, since $\rho \leq \|\mathbf{x} - \mathbf{y}\| + r$ for $\mathbf{x} \in B_\rho(\mathbf{x}_0)$ and $\mathbf{y} \in \partial B_\rho(\mathbf{x}_0)$, we get

$$\frac{\rho^2 - r^2}{\|\mathbf{x} - \mathbf{y}\|^n} \leq \frac{\rho^2 - r^2}{(\rho - r)^n} = \frac{\rho + r}{(\rho - r)^{n-1}}$$

for $\mathbf{x} \in B_\rho(\mathbf{x}_0)$ and $\mathbf{y} \in \partial B_\rho(\mathbf{x}_0)$.

We therefore have

$$\begin{aligned} \int_{\partial B_\rho(\mathbf{x}_0)} u(\mathbf{y}) \frac{\rho - r}{\rho \omega_n (\rho + r)^{n-1}} dS_{\mathbf{y}} &\leq \int_{\partial B_\rho(\mathbf{x}_0)} u(\mathbf{y}) \frac{\rho^2 - r^2}{\rho \omega_n \|\mathbf{x} - \mathbf{y}\|^n} dS_{\mathbf{y}} \\ &\leq \int_{\partial B_\rho(\mathbf{x}_0)} u(\mathbf{y}) \frac{\rho + r}{\rho \omega_n (\rho - r)^{n-1}} dS_{\mathbf{y}} \end{aligned} \quad (9.5.16)$$

for $\mathbf{x} \in B_\rho(\mathbf{x}_0)$.

Moreover,

$$\begin{aligned} \int_{\partial B_\rho(\mathbf{x}_0)} u(\mathbf{y}) \frac{\rho - r}{\rho \omega_n (\rho + r)^{n-1}} dS_{\mathbf{y}} &= \frac{\rho^{n-2}(\rho - r)}{(\rho + r)^{n-1}} \left(\frac{1}{\omega_n \rho^{n-1}} \int_{\partial B_\rho(\mathbf{x}_0)} u(\mathbf{y}) dS_{\mathbf{y}} \right) \\ &= \frac{\rho^{n-2}(\rho - r)}{(\rho + r)^{n-1}} u(\mathbf{x}_0) \end{aligned} \quad (9.5.17)$$

because u is harmonic in Ω . Similarly,

$$\int_{\partial B_\rho(\mathbf{x}_0)} u(\mathbf{y}) \frac{\rho + r}{\rho \omega_n (\rho - r)^{n-1}} dS_{\mathbf{y}} = \frac{\rho^{n-2}(\rho + r)}{(\rho - r)^{n-1}} u(\mathbf{x}_0). \quad (9.5.18)$$

Finally, we have from (9.5.10) that

$$\int_{\partial B_\rho(\mathbf{x}_0)} u(\mathbf{y}) \frac{\rho^2 - r^2}{\rho \omega_n \|\mathbf{x} - \mathbf{y}\|^n} dS_{\mathbf{y}} = u(\mathbf{x}) \quad (9.5.19)$$

for $\mathbf{x} \in B_\rho(\mathbf{x}_0)$.

We obtain the conclusion of the theorem by substituting (9.5.17), (9.5.18) and (9.5.19) in (9.5.16). ■

9.6 Harmonic Polynomials

In this section, we study the harmonic polynomials on $\mathbb{R}^n$. They will be quite useful later. In particular, in the study of potential problem in the next chapter..

Let $P_k = \left\{ \sum_{\alpha \in \mathbb{N}^n, |\alpha|=k} a_\alpha \mathbf{x}^\alpha : a_\alpha \in \mathbb{R} \right\}$ and $H_k = \{p \in P_k : \Delta p = 0\}$.

Proposition 9.6.1

$$P_k = H_k \oplus \|\mathbf{x}\|_2^2 P_{k-2} \text{ and thus } P_k = \bigoplus_{j=0}^{\lfloor k/2 \rfloor} \|\mathbf{x}\|_2^{2j} H_{k-2j}.$$

Proof.

The direct sum in the statement of the proposition is based on a new scalar product on the space P_k . Given $p, q \in P_k$, we define the scalar product of p and q as $\langle p : q \rangle = p(D)q$, where for $p = \sum_{\alpha \in \mathbb{N}^n, |\alpha|=k} a_\alpha \mathbf{x}^\alpha \in P_k$, we set $p(D) = \sum_{\alpha \in \mathbb{N}^n, |\alpha|=k} a_\alpha D^\alpha$. It is easy to verify that this defines a scalar product on P_k . In particular, we have that

$$\langle \mathbf{x}^\alpha : \mathbf{x}^\beta \rangle = \begin{cases} 0 & \text{if } \alpha \neq \beta \\ \alpha! & \text{if } \alpha = \beta \end{cases}$$

Hence, for $p = \sum_{\alpha \in \mathbb{N}^n, |\alpha|=k} a_\alpha \mathbf{x}^\alpha$ and $q = \sum_{\alpha \in \mathbb{N}^n, |\alpha|=k} b_\alpha \mathbf{x}^\alpha$, we have that $\langle p : q \rangle = \sum_{\alpha \in \mathbb{N}^n, |\alpha|=k} a_\alpha b_\alpha \alpha!$.

We now show that the orthogonal complement of $\|\mathbf{x}\|_2^2 P_{k-2}$ in P_k is H_k . Since $q \in P_k$ is in the orthogonal complement of $\|\mathbf{x}\|_2^2 P_{k-2}$ if and only if $0 = \langle \|\mathbf{x}\|_2^2 p : q \rangle = p(D) \Delta q = \langle p : \Delta q \rangle$ for all $p \in P_{k-2}$, in particular for $p(\mathbf{x}) = x^\alpha$ with $|\alpha| = k-2$, we get that $\Delta q = 0$. Thus $q \in H_k$. ■

Since $\|\mathbf{x}\|_2 = 1$ on $\partial B_1(\mathbf{0}) \subset \mathbb{R}^n$, we get the following corollary.

Corollary 9.6.2

If $S = \partial B_1(\mathbf{0}) \subset \mathbb{R}^n$, then $P_k|_S = \bigoplus_{j=0}^{\lfloor k/2 \rfloor} H_{k-2j}|_S$.

Theorem 9.6.3

If $S = \partial B_1(\mathbf{0}) \subset \mathbb{R}^n$, then $L^2(S) = \sum_{k=0}^{\infty} H_k|_S$, where the sum is a direct sum with respect to the standard scalar product on $L^2(S)$.

Proof.

Consider $f \in L^2(S)$ and $\epsilon > 0$. Since $C(S)$ is dense in $L^2(S)$, there exists a continuous function $\tilde{f}$ on S such that

$$\|f - \tilde{f}\|_2 < \epsilon/2. \quad (9.6.1)$$

Using Weierstrass approximation theorem, there exists a polynomial p such that

$$\sup_{\mathbf{x} \in S} |\tilde{f}(\mathbf{x}) - p(\mathbf{x})| < \epsilon/(2\sqrt{\omega_n}). \quad (9.6.2)$$

It follows from (9.6.1) and (9.6.2) that $\|f - p\|_2 < \epsilon$. It follows from the previous corollary that $p|_S = \sum_{k=0}^K p_k$ with $p_k \in H_k$, where K is the degree of p .

Since f and ϵ were arbitrary, this proves that $L^2(S)$ is a sum of the $H_k|_S$. To prove that it is a direct sum, we have to show that $H_i \perp H_j$ if $i \neq j$. Before proving this result, we need to make the following observation.

Given $p = \sum_{\alpha \in \mathbb{N}^n, |\alpha|=k} a_\alpha \mathbf{x}^\alpha \in P_k$, we have that

$$\begin{aligned} \sum_{s=1}^n x_s \frac{\partial}{\partial x_s} p(\mathbf{x}) &= \sum_{\alpha \in \mathbb{N}^n, |\alpha|=k} a_\alpha \left(\sum_{s=1}^n x_s \frac{\partial}{\partial x_s} \mathbf{x}^\alpha \right) = \sum_{\alpha \in \mathbb{N}^n, |\alpha|=k} a_\alpha \left(\sum_{s=1}^n \alpha_s \mathbf{x}^\alpha \right) \\ &= \sum_{\alpha \in \mathbb{N}^n, |\alpha|=k} a_\alpha \underbrace{\left(\sum_{s=1}^n \alpha_s \right)}_{=k} \mathbf{x}^\alpha = k p(\mathbf{x}). \end{aligned}$$

The equation above is known as **Euler's relation**.

Given $p \in H_i$ and $q \in H_j$ with $i \neq j$, we get from the Green's identity (9.1.2) on $B_1(\mathbf{0})$ that

$$0 = \int_{B_1(\mathbf{0})} (p \Delta q - q \Delta p) d\mathbf{x} = \int_S \left(p \frac{\partial q}{\partial \nu} - q \frac{\partial p}{\partial \nu} \right) dS_{\mathbf{x}} = (j - i) \int_S p q dS_{\mathbf{x}} = (j - i) \langle p, q \rangle ,$$

where we have use the fact that

$$\frac{\partial p}{\partial \nu}(\mathbf{x}) = \nabla p(\mathbf{x}) \cdot \mathbf{x} = \sum_{s=1}^n x_s \frac{\partial}{\partial x_s} p(\mathbf{x}) = ip(\mathbf{x})$$

and similarly $\frac{\partial q}{\partial \nu}(\mathbf{x}) = jq(\mathbf{x})$ for $\mathbf{x} \in S$. Thus $\langle p, q \rangle = 0$. Since p and q are arbitrary, $H_i \perp H_j$. ■

A basis of P_k is formed of the monomials $\mathbf{x}^\alpha$ with $\alpha \in \mathbb{N}^n$ and $|\alpha| = k$. Suppose that you have a line of $n + k - 1$ balls. The number of possible monomials of degree k is the number of way to remove $n - 1$ balls among the $n + k - 1$ balls to create at most n groups of balls. The number of balls in the first group is the exponent of x_1 , the number of balls in the second group is the exponent of x_2 , and so on. Thus $\dim P_k = \frac{(n + k - 1)!}{(n - 1)! k!}$. It follows from Proposition 9.6.1 that

$$\begin{aligned} \dim H_k &= \dim P_k - \dim P_{k-2} = \frac{(n + k - 1)!}{(n - 1)! k!} - \frac{(n + k - 3)!}{(n - 1)! (k - 2)!} \\ &= ((n + k - 1)(n + k - 2) - k(k - 1)) \frac{(n + k - 3)!}{(n - 1)! k!} = (2k + n - 2) \frac{(n + k - 3)!}{(n - 2)! k!} . \end{aligned}$$

We get the following result.

Proposition 9.6.4

$$d_k \equiv \dim H_k = O(k^{n-2}) \text{ as } k \rightarrow \infty .$$

Proof.

For $k > 3$, we have that

$$\begin{aligned} \frac{d_k}{k^{n-2}} &= \frac{(2k + n - 2)}{(n - 2)! k^{n-2}} (n + k - 3)(n + k - 2) \dots (k + 1) \\ &= \left(2 + \frac{n - 2}{k} \right) \frac{1}{(n - 2)!} \left(1 + \frac{n - 3}{k} \right) \left(1 + \frac{n - 2}{k} \right) \dots \left(1 + \frac{1}{k} \right) \rightarrow \frac{2}{(n - 2)!} \neq 0 \end{aligned}$$

as $k \rightarrow \infty$. ■

Let $S = \partial B_1(\mathbf{0}) \subset \mathbb{R}^n$. For each $\mathbf{x} \in S$, the mapping

$$\begin{aligned} \phi_{\mathbf{x}} : H_k|_S &\rightarrow \mathbb{R} \\ p &\mapsto p(\mathbf{x}) \end{aligned}$$

is a linear functional on the finite dimensional vector subspace $H_k|_S$ of $L^2(S)$. From the finite dimensional version of the Riesz Representation Theorem, there exists $p_{\mathbf{x}}^{[k]} \in H_k|_S$ such that

$$p(\mathbf{x}) = \phi_{\mathbf{x}}(p) = \langle p, p_{\mathbf{x}}^{[k]} \rangle = \int_S p(\mathbf{y}) p_{\mathbf{x}}^{[k]}(\mathbf{y}) dS_{\mathbf{y}}$$

for all $p \in H_k|_S$.

Proposition 9.6.5

$$\|p_{\mathbf{x}}^{[k]}\|_2 = \sqrt{d_k/\omega_n} \text{ and } \|p_{\mathbf{x}}^{[k]}\|_{\infty} = d_k/\omega_n.$$

Proof.

i) Let $R: \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a rotation. Using R has a change of variable on S , we get

$$\begin{aligned} \langle p, p_{R(\mathbf{x})}^{[k]} \circ R \rangle &= \int_S p(\mathbf{y}) p_{R(\mathbf{x})}^{[k]}(R(\mathbf{y})) dS_{\mathbf{y}} = \int_S p(R^{-1}(\mathbf{y}), p_{R(\mathbf{x})}^{[k]}(\mathbf{y}) \underbrace{|\det R^{-1}(\mathbf{y})|}_{=1} dS_{\mathbf{y}} \\ &= \langle p \circ R^{-1}, p_{R(\mathbf{x})}^{[k]} \rangle = (p \circ R^{-1})(R(\mathbf{x})) = p(\mathbf{x}) \end{aligned}$$

for all $p \in H_k|_S$. Since $p_{\mathbf{x}}^{[k]}$ is the unique element of $H_k|_S$ such that $p(\mathbf{x}) = \langle p, p_{\mathbf{x}}^{[k]} \rangle$ for all $p \in H_k|_S$, we have that $p_{\mathbf{x}}^{[k]} = p_{R(\mathbf{x})}^{[k]} \circ R$ for all rotation R .

It follows that $p_{\mathbf{x}}^{[k]}(\mathbf{x}) = p_{R(\mathbf{x})}^{[k]}(R(\mathbf{x}))$ for all $\mathbf{x} \in S$. Thus $p_{\mathbf{x}}^{[k]}(\mathbf{x})$ is constant for $\mathbf{x} \in S$.

ii) Let $\{p_j\}_{j=1}^{d_k}$ be an orthonormal basis of $H_k|_S \subset L^2(S)$. We note that

$$p_{\mathbf{x}}^{[k]}(\mathbf{y}) = \sum_{j=1}^{d_k} \langle p_{\mathbf{x}}^{[k]}, p_j \rangle p_j(\mathbf{y}) = \sum_{j=1}^{d_k} \langle p_j, p_{\mathbf{x}}^{[k]} \rangle p_j(\mathbf{y}) = \sum_{j=1}^{d_k} p_j(\mathbf{x}) p_j(\mathbf{y}) \quad (9.6.3)$$

for all $\mathbf{y} \in S$. Thus,

$$p_{\mathbf{x}}^{[k]}(\mathbf{x}) = \sum_{j=1}^{d_k} (p_j(\mathbf{x}))^2$$

for all $\mathbf{x} \in S$. Hence,

$$d_k = \sum_{j=1}^{d_k} \|p_j\|_2^2 = \int_S \left(\sum_{j=1}^{d_k} (p_j(\mathbf{x}))^2 \right) dS_{\mathbf{x}} = \int_S (p_{\mathbf{x}}^{[k]}(\mathbf{x})) dS_{\mathbf{x}} = p_{\mathbf{x}}^{[k]}(\mathbf{x}) \int_S dS_{\mathbf{x}} = \omega_n p_{\mathbf{x}}^{[k]}(\mathbf{x}).$$

We deduce that $p_{\mathbf{x}}^{[k]}(\mathbf{x}) = d_k/\omega_n$.

iii) It follows from (9.6.3) that

$$\|p_{\mathbf{x}}^{[k]}\|_2^2 = \langle p_{\mathbf{x}}^{[k]}, p_{\mathbf{x}}^{[k]} \rangle = \sum_{i,j=1}^{d_k} p_j(\mathbf{x}) p_i(\mathbf{x}) \langle p_j, p_i \rangle = \sum_{j=1}^{d_k} (p_j(\mathbf{x}))^2 = p_{\mathbf{x}}^{[k]}(\mathbf{x}) = \frac{d_k}{\omega_n}.$$

iv) Since $p_{\mathbf{x}}^{[k]}(\mathbf{y}) = \langle p_{\mathbf{x}}^{[k]}, p_{\mathbf{y}}^{[k]} \rangle$ for all $\mathbf{x}, \mathbf{y} \in S$ by definition of $p_{\mathbf{y}}^{[k]}$, we have that

$$\left| p_{\mathbf{x}}^{[k]}(\mathbf{y}) \right| = \left| \langle p_{\mathbf{x}}^{[k]}, p_{\mathbf{y}}^{[k]} \rangle \right| \leq \|p_{\mathbf{x}}^{[k]}\|_2 \|p_{\mathbf{y}}^{[k]}\|_2 \leq \frac{d_k}{\omega_n}$$

for all $\mathbf{y} \in S$. Thus $\|p_{\mathbf{x}}^{[k]}\|_{\infty} = \sup_{\mathbf{y} \in S} |p_{\mathbf{x}}^{[k]}(\mathbf{y})| \leq \frac{d_k}{\omega_n}$. However, we have that $p_{\mathbf{x}}^{[k]}(\mathbf{x}) = \frac{d_k}{\omega_n}$. Thus $\|p_{\mathbf{x}}^{[k]}\|_{\infty} = \frac{d_k}{\omega_n}$. ■

Theorem 9.6.6

Let $S = \partial B_1(\mathbf{0}) \subset \mathbb{R}^n$. Given $g \in L^2(S)$. The function $u : \overline{B_1(\mathbf{0})} \rightarrow \mathbb{R}$ defined by

$$u(\mathbf{x}) = \sum_{k=0}^{\infty} \|\mathbf{x}\|^k \langle g, p_{\|\mathbf{x}\|^{-1}\mathbf{x}}^{[k]} \rangle \quad (9.6.4)$$

is an harmonic function on $B_1(\mathbf{0})$ and satisfies $u = g$ in $L^2(S)$. More specifically,

- i) The series in (9.6.4) is absolutely and uniformly convergent on $\overline{B_r(\mathbf{0})}$ for every $0 \leq r < 1$.
- ii) The series $\sum_{k=0}^{\infty} r^k \langle f, p_{\mathbf{x}}^{[k]} \rangle$ for $\mathbf{x} \in S$ is convergent in $L^2(S)$ for every $0 \leq r \leq 1$.
- iii) If we define

$$u_r(\mathbf{x}) \equiv u(r\mathbf{x}) = \sum_{k=0}^{\infty} r^k \langle g, p_{\mathbf{x}}^{[k]} \rangle$$

for all $\mathbf{x} \in S$, then $u_r \rightarrow u_1 = g$ as $r \rightarrow 1^-$ in $L^2(S)$.

Proof.

Let us first show that the orthogonal projection $\pi_k : L^2(S) \rightarrow H_k|_S$ is given by $\pi_k(g)(\mathbf{x}) = \langle g, p_{\mathbf{x}}^{[k]} \rangle$ for $\mathbf{x} \in S$. According to Theorem 9.6.3, we have that $g = \pi_k(g) + h$ where $h \in (H_k|_S)^{\perp}$. Thus, since $p_{\mathbf{x}}^{[k]} \in H_k|_S$, we have that

$$\langle g, p_{\mathbf{x}}^{[k]} \rangle = \langle \pi_k(g), p_{\mathbf{x}}^{[k]} \rangle + \langle h, p_{\mathbf{x}}^{[k]} \rangle = \langle \pi_k(g), p_{\mathbf{x}}^{[k]} \rangle = \pi_k(g)(\mathbf{x})$$

for all $\mathbf{x} \in S$. Therefore, there exists a unique $\phi_k \in H_k$ such that $\phi_k(\mathbf{x}) = \langle g, p_{\mathbf{x}}^{[k]} \rangle$ for all $\mathbf{x} \in S$. Note that $\phi_k(\mathbf{x}) = \|\mathbf{x}\|^k \phi_k(\|\mathbf{x}\|^{-1}\mathbf{x}) = r^k \phi_k(\mathbf{y})$ if $\mathbf{x} = r\mathbf{y}$ with $\mathbf{y} \in S$ because ϕ_k is homogeneous of degree k .

i) Choose $0 < r < 1$. It follows from Propositions 9.6.4 and 9.6.5 that

$$\|p_{\mathbf{x}}^{[k]}\|_2^2 = \int_S \left| p_{\mathbf{x}}^{[k]}(\mathbf{y}) \right|^2 dS_{\mathbf{y}} \leq C(1+k)^{n-2}$$

for a constant C ⁴ and all $\mathbf{x} \in S$. Hence, for all $\mathbf{x} \in \overline{B_r(\mathbf{0})} \setminus \{\mathbf{0}\}$, we have that

$$\begin{aligned} \sum_{k=0}^{\infty} |\phi_k(\mathbf{x})| &= \sum_{k=0}^{\infty} \|\mathbf{x}\|^k \phi_k(\|\mathbf{x}\|^{-1}\mathbf{x}) = \sum_{k=0}^{\infty} \|\mathbf{x}\|^k \left\langle g, p_{\|\mathbf{x}\|^{-1}\mathbf{x}}^{[k]} \right\rangle \leq \sum_{k=0}^{\infty} r^k \|g\|_2 \|p_{\|\mathbf{x}\|^{-1}\mathbf{x}}^{[k]}\|_2 \\ &\leq \sqrt{C} \|g\|_2 \sum_{k=0}^{\infty} r^k (1+k)^{(n-2)/2} < \infty \end{aligned}$$

because the last series converges according to the ratio test. We therefore have that the series $\sum_{k=0}^{\infty} \phi_k$ is absolutely and uniformly convergent on $\overline{B_r(\mathbf{0})}$. It follows from Corollary 9.1.6

that $u = \sum_{k=0}^{\infty} \phi_k$ is harmonic on $B_r(\mathbf{0})$.

ii) Since $\phi_k = \pi_k(g)$ on S , it follows from Theorem 9.6.3 that $\sum_{k=0}^{\infty} \phi_k$ converges to g in $L^2(S)$.

Moreover, $\|g\|_2^2 = \sum_{j=0}^{\infty} \|\phi_k\|_2^2 < \infty$ where $\|\phi_k\|_2 = \left(\int_S |\phi_k(\mathbf{x})|^2 dS_{\mathbf{x}} \right)^{1/2}$. Since $r^k \phi_k$ is in $H_k|_S$ and $\sum_{j=0}^{\infty} \|r^k \phi_k\|_2^2 \leq \sum_{j=0}^{\infty} \|\phi_k\|_2^2 < \infty$ for $0 \leq r \leq 1$, it follows again from Theorem 9.6.3 that $\sum_{j=0}^{\infty} r^k \phi_k$ converges in $L^2(S)$ for $0 \leq r \leq 1$.

iii) We have that $u_r = \sum_{j=0}^{\infty} r^k \phi_k$ on S . It follows from (ii) that $u_r \in L^2(S)$ for $0 \leq r \leq 1$ because the series that defined u_r converges in $L^2(S)$. Hence,

$$\|u_r - u_1\|_2 \leq \left\| u_r - \sum_{k=0}^K r^k \phi_k \right\|_2 + \left\| \sum_{k=0}^K r^k \phi_k - \sum_{k=0}^K \phi_k \right\|_2 + \left\| \sum_{k=0}^K \phi_k - u_1 \right\|_2$$

for $K > 0$. Given $\epsilon > 0$, choose K large enough to have the first and third norms smaller than $\epsilon/3$. This is possible because of (ii). Since S is a compact set, there exists $0 < R < 1$ such that

$$(1 - r^k) \sup_{\mathbf{x} \in S} \sum_{k=0}^K |\phi_k(\mathbf{x})| < \frac{\epsilon}{3\sqrt{\omega_n}}$$

for $R < r \leq 1$. Thus, for $R < r \leq 1$, the second norm is smaller than $\epsilon/3$ and it follows that $\|u_r - u_1\| \leq \epsilon$ for $R < r \leq 1$. ■

Corollary 9.6.7

Let $S = \partial B_1(\mathbf{0}) \subset \mathbb{R}^n$. The Poisson kernel P for the Dirichlet problem in the ball $B_1(\mathbf{0})$ given in (9.5.9) can be expressed as

$$P(\mathbf{x}, \mathbf{y}) = \sum_{k=0}^{\infty} \|\mathbf{x}\|^k p_{\|\mathbf{x}\|^{-1}\mathbf{x}}^{[k]}(\mathbf{y})$$

for $\mathbf{x} \in B_1(\mathbf{0})$ and $\mathbf{y} \in S$. The series is uniformly convergent on $\overline{B_r(\mathbf{x})}$ for every

⁴We have to use $(1+k)^{n-2}$ instead of k^{n-2} as predicted by (9.6.4) to handle the case $k = 0$.

$0 \leq r < 1$.

If $g \in C(S)$, then u defined in (9.6.4) is the unique harmonic function on $B_1(\mathbf{0})$ which is continuous on $\overline{B_1(\mathbf{0})}$ and satisfies $u = g$ on S .

Proof.

Given $0 < r < 1$, it follows from Propositions 9.6.4 and 9.6.5 that $\|p_{\mathbf{x}}^{[k]}\|_{\infty} \leq C(1+k)^{n-2}$ for a constant C ⁵ and all $\mathbf{x} \in S$. Hence, for all $\mathbf{x} \in \overline{B_r(\mathbf{0})} \setminus \{\mathbf{0}\}$, we have that

$$\sum_{k=0}^{\infty} \|\mathbf{x}\|^k p_{\|\mathbf{x}\|^{-1}\mathbf{x}}^{[k]}(\mathbf{y})\| \leq \sum_{k=0}^{\infty} r^k \|p_{\|\mathbf{x}\|^{-1}\mathbf{x}}^{[k]}\|_{\infty} \leq C \sum_{k=0}^{\infty} r^k (1+k)^{n-2} < \infty$$

because the last series converges according to the ratio test. We therefore have that the series $\sum_{k=0}^{\infty} \|\mathbf{x}\|^k p_{\|\mathbf{x}\|^{-1}\mathbf{x}}^{[k]}(\mathbf{y})$ is absolutely and uniformly convergent on $\overline{B_r(\mathbf{0})}$, and the limit is a continuous function on $\overline{B_r(\mathbf{0})}$ which is harmonic on $B_r(\mathbf{0})$.

Suppose that $g \in L^2(\Omega)$. It follows from Theorems 9.5.16 and 9.6.6 that the harmonic function u in $B_1(\mathbf{0})$ with $u = g$ in $L^2(S)$ satisfies

$$\begin{aligned} \int_S g(\mathbf{y}) P(s\mathbf{x}, \mathbf{y}) dS_{\mathbf{y}} &= u(s\mathbf{x}) = \sum_{k=0}^{\infty} s^k \left\langle g, p_{\mathbf{x}}^{[k]} \right\rangle = \sum_{k=0}^{\infty} s^k \int_S g(\mathbf{y}) p_{\mathbf{x}}^{[k]}(\mathbf{y}) dS_{\mathbf{y}} \\ &= \int_S g(\mathbf{y}) \left(\sum_{k=0}^{\infty} s^k p_{\mathbf{x}}^{[k]}(\mathbf{y}) \right) dS_{\mathbf{y}} \end{aligned} \quad (9.6.5)$$

for all $\mathbf{x} \in S$ and $0 \leq s \leq r$. We have used the uniform convergence of $\sum_{k=0}^{\infty} \|\mathbf{x}\|^k p_{\|\mathbf{x}\|^{-1}\mathbf{x}}^{[k]}(\mathbf{y})$ on $\overline{B_r(\mathbf{0})}$ to interchange the sum and the integral in the equation above.

Since (9.6.5) is true for all functions $g \in L^2(S)$, we have that $P(s\mathbf{x}, \mathbf{y}) = \sum_{k=0}^{\infty} s^k p_{\mathbf{x}}^{[k]}(\mathbf{y})$ for almost all $\mathbf{y} \in S$. Moreover, since $\mathbf{y} \mapsto P(s\mathbf{x}, \mathbf{y})$ and $\mathbf{y} \mapsto \sum_{k=0}^{\infty} s^k p_{\mathbf{x}}^{[k]}(\mathbf{y})$ are continuous on $\mathbf{y} \in S$, we get that $P(s\mathbf{x}, \mathbf{y}) = \sum_{k=0}^{\infty} s^k p_{\mathbf{x}}^{[k]}(\mathbf{y})$ for all $\mathbf{y} \in S$. This is true for all $\mathbf{x} \in S$ and all $0 \leq s \leq r$.

If $g \in C(S)$, then it follows from Theorem 9.5.16 that

$$u(\mathbf{x}) = \int_{\partial B_1(\mathbf{0})} g(\mathbf{y}) P(\mathbf{x}, \mathbf{y}) dS_{\mathbf{y}} = \sum_{k=0}^{\infty} \|\mathbf{x}\|^k \left\langle g, p_{\|\mathbf{x}\|^{-1}\mathbf{x}}^{[k]} \right\rangle$$

u can be continuously extended to $\overline{B_1(\mathbf{0})}$ by $u(\mathbf{x}) = g(\mathbf{x})$ for $\mathbf{x} \in S$. The uniqueness of u then comes from Theorem 9.2.6. ■

⁵We have to use $(1+k)^{n-2}$ instead of k^{n-2} as predicted by Proposition 9.6.4 to handle the case $k = 0$.

9.7 Reflection Principle

In complex analysis, we have a reflection principle to expand an harmonic function defined for the complex numbers with non-negative imaginary part and null on the real axis to an harmonic function on the entire complex plan. There is an equivalent result for harmonic functions in $\mathbb{R}^n$.

Theorem 9.7.1 (Reflection Principle)

Let $\Omega \subset \mathbb{R}^n$ be an open subset, symmetric with respect to the hyperplane $x_n = 0$; namely, $(\check{\mathbf{x}}, x_n) \in \Omega$ if and only if $(\check{\mathbf{x}}, -x_n) \in \Omega$. Let $\Omega_+ = \{(\check{\mathbf{x}}, x_n) \in \Omega : x_n > 0\}$, $\Omega_- = \{(\check{\mathbf{x}}, x_n) \in \Omega : x_n < 0\}$ and $\Omega_0 = \{(\check{\mathbf{x}}, 0) \in \Omega\}$. Suppose that $u : \Omega_+ \cup \Omega_0 \rightarrow \mathbb{R}$ is continuous, harmonic in Ω_+ and $u(\mathbf{x}) = 0$ for $\mathbf{x} \in \Omega_0$. Then u can be extended to an harmonic function on Ω by $u(\check{\mathbf{x}}, x_n) = -u(\check{\mathbf{x}}, -x_n)$ for $(\mathbf{x}, x_n) \in \Omega_-$.

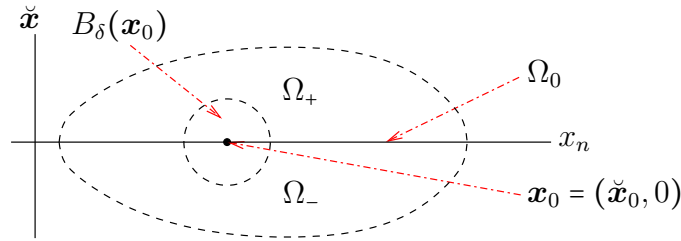


Figure 9.5: Figure used in the proof of the reflection principle, Theorem 9.7.1, for harmonic functions.

Proof.

In this proof, we refer to Figure 9.5. The extension is obviously continuous on Ω because u is continuous on $\Omega_+ \cup \Omega_0$ and null on Ω_0 . Moreover, the extension of u is harmonic on $\Omega \setminus \Omega_0$. We need only prove that u is harmonic in a neighbourhood of an arbitrary point $\mathbf{x}_0 = (\check{\mathbf{x}}_0, 0) \in \Omega_0$.

Choose $\delta > 0$ such that $\overline{B_\delta(\mathbf{x}_0)} \subset \Omega$. According to Remark 9.5.17, the solution of the Dirichlet problem $\Delta v = 0$ on $B_\delta(\mathbf{x}_0)$ with the boundary condition $v = u$ on $\partial B_\delta(\mathbf{x}_0)$ is

$$v(\mathbf{x}) = \int_{\partial B_\delta(\mathbf{x}_0)} u(\mathbf{y}) \frac{\delta^2 - \|\mathbf{x} - \mathbf{x}_0\|^2}{\delta \omega_n \|\mathbf{x} - \mathbf{y}\|^n} dS_{\mathbf{y}}$$

for $\mathbf{x} \in B_\delta(\mathbf{x}_0)$. For $(\check{\mathbf{x}}, x_n) \in B_\delta(\mathbf{x}_0)$, we have $\|(\check{\mathbf{x}}, x_n) - \mathbf{x}_0\| = \|(\check{\mathbf{x}}, -x_n) - \mathbf{x}_0\|$ and $\|(\check{\mathbf{x}}, x_n) - (\check{\mathbf{y}}, y_n)\| = \|(\check{\mathbf{x}}, -x_n) - (\check{\mathbf{y}}, -y_n)\|$. Hence

$$v(\check{\mathbf{x}}, x_n) = \int_{\partial B_\delta(\mathbf{x}_0)} u(\check{\mathbf{y}}, y_n) \frac{\delta^2 - \|(\check{\mathbf{x}}, x_n) - \mathbf{x}_0\|^2}{\delta \omega_n \|(\check{\mathbf{x}}, x_n) - (\check{\mathbf{y}}, y_n)\|^n} dS_{\mathbf{y}}$$

$$\begin{aligned}
&= \int_{\partial B_\delta(\mathbf{x}_0)} u(\check{\mathbf{y}}, y_n) \frac{\delta^2 - \|(\check{\mathbf{x}}, -x_n) - \mathbf{x}_0\|^2}{\delta \omega_n \|(\check{\mathbf{x}}, -x_n) - (\check{\mathbf{y}}, -y_n)\|^n} dS_{\mathbf{y}} \\
&= \int_{\partial B_\delta(\mathbf{x}_0)} u(\check{\mathbf{y}}, -y_n) \frac{\delta^2 - \|(\check{\mathbf{x}}, -x_n) - \mathbf{x}_0\|^2}{\delta \omega_n \|(\check{\mathbf{x}}, -x_n) - (\check{\mathbf{y}}, y_n)\|^n} dS_{\mathbf{y}} \\
&= - \int_{\partial B_\delta(\mathbf{x}_0)} u(\check{\mathbf{y}}, y_n) \frac{\delta^2 - \|(\check{\mathbf{x}}, -x_n) - \mathbf{x}_0\|^2}{\delta \omega_n \|(\check{\mathbf{x}}, -x_n) - (\check{\mathbf{y}}, y_n)\|^n} dS_{\mathbf{y}} = -v(\check{\mathbf{x}}, -x_n)
\end{aligned}$$

for all $(\check{\mathbf{x}}, x_n) \in B_\delta(\mathbf{x}_0)$. We have used the symmetry of $\partial B_\delta(\mathbf{x}_0)$ with respect to the hyperplane $y_n = 0$ for the third equality and $u(\check{\mathbf{y}}, -y_n) = -u(\check{\mathbf{y}}, y_n)$ for the last equality. It follows from this identity that $v(\check{\mathbf{x}}, 0) = -v(\check{\mathbf{x}}, 0)$ for all $(\check{\mathbf{x}}, 0) \in \Omega_0 \cap B_\delta(\mathbf{x}_0)$. Thus $v(\check{\mathbf{x}}, 0) = u(\check{\mathbf{x}}, 0) = 0$ for all $(\check{\mathbf{x}}, 0) \in \Omega_0 \cap B_\delta(\mathbf{x}_0)$.

Let $B_+ = B_\delta(\mathbf{x}_0) \cap \Omega_+$. Since $u = v$ on ∂B_+ and both u and v are harmonic in B_+ , we have by uniqueness of solutions that $u = v$ in B_+ . For similar reasons, $u = v$ on $B_- = B_\delta(\mathbf{x}_0) \cap \Omega_-$. Thus, since $u = v$ on $B_\delta(\mathbf{x}_0) \cap \Omega_0$, we get that $u = v$ on $B_\delta(\mathbf{x}_0)$. In particular, u is harmonic in $B_\delta(\mathbf{x}_0)$. ■

9.8 Perron's Method

In Section 9.9, we “give” the solution of the Dirichlet problem using potential theory. In this section, we prove the existence of the solution of the Dirichlet problem on a bounded open subset of $\mathbb{R}^n$ using subharmonic functions. This method is known as Perron's method. It is valid for a more general type of boundaries than the method used in Section 9.9. However, both methods relies upon the knowledge of the existence of the solution for the Dirichlet problem on $B_1(\mathbf{0})$ given before. Perron's method only proves the existence of the solution but does not provide a formula for the solution as it is done in Section 9.9.

In this section, we follow the presentation in [16].

Definition 9.8.1

Suppose that Ω is an open subset of $\mathbb{R}^n$ and that $u \in C(\Omega)$. Given $\mathbf{z} \in \Omega$ and $\rho > 0$ such that $\overline{B_\rho(\mathbf{z})} \subset \Omega$, we define a new continuous function $u_{\mathbf{z}, \rho} : \Omega \rightarrow \mathbb{R}$ by

$$u_{\mathbf{z}, \rho} = \begin{cases} u(\mathbf{x}) & \text{if } \mathbf{x} \in \Omega \setminus B_\rho(\mathbf{z}) \\ v(\mathbf{x}) & \text{if } \mathbf{x} \in B_\rho(\mathbf{z}) \end{cases}$$

where $v : B_\rho(\mathbf{z}) \rightarrow \mathbb{R}$ is the solution of the Dirichlet problem $\Delta v = 0$ on $B_\rho(\mathbf{z})$ with $v(\mathbf{x}) = u(\mathbf{x})$ for $\mathbf{x} \in \partial B_\rho(\mathbf{z})$.

Lemma 9.8.2

Let Ω be an open subset of $\mathbb{R}^n$ and $u \in \Omega$ be a subharmonic function. Suppose that $\mathbf{z} \in \Omega$ and $\overline{B_\rho(\mathbf{z})} \subset \Omega$ for some $\rho > 0$. Then $u(\mathbf{x}) \leq u_{\mathbf{z}, \rho}(\mathbf{x})$ for all $\mathbf{x} \in \Omega$, and $u_{\mathbf{z}, \rho} : \Omega \rightarrow \mathbb{R}$ is subharmonic.

Proof.

i) We have that $u(\mathbf{x}) \leq u_{\mathbf{z},\rho}(\mathbf{x})$ for all $\mathbf{x} \in \Omega \setminus B_\rho(\mathbf{z})$ by definition of $u_{\mathbf{z},\rho}$. Since $u - u_{\mathbf{z},\rho}$ is a subharmonic function on $B_\rho(\mathbf{z})$, is continuous on $\overline{B_\rho(\mathbf{z})}$, and satisfies $(u - u_{\mathbf{z},\rho})(\mathbf{x}) = 0$ for all $\mathbf{x} \in \partial B_\rho(\mathbf{z})$, it follows from the Maximum Principle Theorem, Theorem 9.2.3, that $(u - u_{\mathbf{z},\rho})(\mathbf{x}) \leq 0$ for all $\mathbf{x} \in \overline{B_\rho(\mathbf{z})}$; namely, $u(\mathbf{x}) \leq u_{\mathbf{z},\rho}(\mathbf{x})$ for all $\mathbf{x} \in \overline{B_\rho(\mathbf{z})}$.

ii) Since $u_{\mathbf{z},\rho} = u$ on the open set $\Omega \setminus \overline{B_\rho(\mathbf{z})}$, we have that $u_{\mathbf{z},\rho}$ is subharmonic on $\Omega \setminus \overline{B_\rho(\mathbf{z})}$. Since $u_{\mathbf{z},\rho}$ is harmonic on $B_\rho(\mathbf{z})$, then $u_{\mathbf{z},\rho}$ is subharmonic on $B_\rho(\mathbf{z})$.

To prove that $u_{\mathbf{z},\rho}$ is subharmonic on Ω , we need to show that for every $\mathbf{x} \in \partial B_\rho(\mathbf{z})$, we can find $R > 0$ such that $B_r(\mathbf{x}) \subset \Omega$ and $u_{\mathbf{z},\rho}(\mathbf{x}) \leq M_{u_{\mathbf{z},\rho}}(\mathbf{x}, r)$ for all $0 < r < R$.

Consider $\mathbf{x} \in \partial B_\rho(\mathbf{z})$. Since u is subharmonic on Ω , there exists $R > 0$ such that $B_r(\mathbf{x}) \subset \Omega$ and $u(\mathbf{x}) \leq M_u(\mathbf{x}, r)$ for all $0 < r < R$. But $u(\mathbf{x}) = u_{\mathbf{z},\rho}(\mathbf{x})$ because $\mathbf{x} \in \partial B_\rho(\mathbf{z})$, and $u(\mathbf{x}) \leq u_{\mathbf{z},\rho}(\mathbf{x})$ for all $\mathbf{x} \in \Omega$. Thus, $u_{\mathbf{z},\rho}(\mathbf{x}) = u(\mathbf{x}) \leq M_u(\mathbf{x}, r) \leq M_{u_{\mathbf{z},\rho}}(\mathbf{x}, r)$ for all $0 < r < R$. ■

Lemma 9.8.3

Let Ω be an open subset of $\mathbb{R}^n$ and $u \in \Omega$ is a subharmonic function. Suppose that $\mathbf{z} \in \Omega$ and $\overline{B_\rho(\mathbf{z})} \subset \Omega$ for some $\rho > 0$. Then $u(\mathbf{z}) \leq M_u(\mathbf{z}, \rho)$

Proof.

We have that $u(\mathbf{z}) \leq u_{\mathbf{z},\rho}(\mathbf{z}) = M_{u_{\mathbf{z},\rho}}(\mathbf{z}, \rho) = M_u(\mathbf{z}, \rho)$, where the first inequality is a consequence of the previous lemma, the first equality is true because $u_{\mathbf{z},\rho}$ is harmonic on $B_\rho(\mathbf{z})$, and the last equality is true because $u_{\mathbf{z},\rho}(\mathbf{x}) = u(\mathbf{x})$ for all $\mathbf{x} \in \partial B_\rho(\mathbf{z})$. ■

Lemma 9.8.4

Let Ω be an open subset of $\mathbb{R}^n$ and $u \in C(\Omega)$. u is harmonic on Ω if and only if u and $-u$ are subharmonic on Ω .

Proof.

⇒) If u is harmonic on Ω then u and $-u$ are harmonic functions and so also subharmonic on Ω .

⇐) Suppose that u and $-u$ are subharmonic functions on Ω . Given $\mathbf{z} \in \Omega$ and $\rho > 0$ such that $\overline{B_\rho(\mathbf{z})} \subset \Omega$, we have from Lemma 9.8.2 that $u(\mathbf{x}) \leq u_{\mathbf{z},\rho}(\mathbf{x})$ and $-u(\mathbf{x}) \leq -u_{\mathbf{z},\rho}(\mathbf{x})$ for all $\mathbf{x} \in B_\rho(\mathbf{z})$. Thus, $u = u_{\mathbf{z},\rho}$ on $B_\rho(\mathbf{z})$. Since $u_{\mathbf{z},\rho}$ is harmonic on $B_\rho(\mathbf{z})$, we have that u is harmonic on $B_\rho(\mathbf{z})$. Since this is true for all $\mathbf{z} \in \Omega$ and $\rho > 0$ such that $\overline{B_\rho(\mathbf{z})} \subset \Omega$, then u is harmonic on Ω . ■

Lemma 9.8.5

Let Ω be an open subset of $\mathbb{R}^n$ and $u \in C(\Omega)$. If, for each $\mathbf{x} \in \Omega$, there exists $R > 0$ (that may depends on $\mathbf{x}$) such that $B_R(\mathbf{x}) \subset \Omega$ and $u(\mathbf{x}) = M_u(\mathbf{x}, r)$ for $0 < r < R$, then u is harmonic on Ω .

Proof.

The assumption of the lemma implies that u and $-u$ are subharmonic on Ω . The conclusion of the lemma follows from the previous lemma. ■

Suppose that Ω is an bounded open subset of $\mathbb{R}^n$ and that $f : \partial\Omega \rightarrow \mathbb{R}$ is a continuous function. We define the set

$$\Sigma_f = \{u \in C(\overline{\Omega}) : u \text{ is subharmonic on } \Omega \text{ and } u(\mathbf{x}) \leq f(\mathbf{x}) \text{ for all } \mathbf{x} \in \partial\Omega\}$$

and the function

$$u_f : \overline{\Omega} \rightarrow \mathbb{R} \\ \mathbf{x} \mapsto \sup_{u \in \Sigma_f} u(\mathbf{x})$$

We obviously need to justify why u_f is well defined. Since f is continuous on the compact set $\partial\Omega$, we have that $m = \min_{\mathbf{x} \in \partial\Omega} f(\mathbf{x}) > -\infty$ and $M = \max_{\mathbf{x} \in \partial\Omega} f(\mathbf{x}) < \infty$. The function $u_m : \overline{\Omega} \rightarrow \mathbb{R}$ defined by $u_m(\mathbf{x}) = m$ for all $\mathbf{x} \in \overline{\Omega}$ belongs to Σ_f . So, $\Sigma_f \neq \emptyset$. It also follows from the Maximum Principle Theorem that $\sup_{\mathbf{x} \in \overline{\Omega}} u(\mathbf{x}) \leq M$ for all $u \in \Sigma_f$. Thus, $u_f(\mathbf{x}) \leq M$ for all $\mathbf{x} \in \overline{\Omega}$.

Lemma 9.8.6

If $\{u_j\}_{j=1}^k \subset \Sigma_f$, then $u = \max\{u_1, u_2, \dots, u_j\} \in \Sigma_f$.

Proof.

Note that $u = \max\{u_1, u_2, \dots, u_j\}$ is defined by $u(\mathbf{x}) = \max\{u_1(\mathbf{x}), u_2(\mathbf{x}), \dots, u_j(\mathbf{x})\}$ for all $\mathbf{x} \in \overline{\Omega}$.

We leave it to the reader to verify that $u \in C(\overline{\Omega})$. We have that $u(\mathbf{x}) \leq f(\mathbf{x})$ for $\mathbf{x} \in \partial\Omega$ because $u_j(\mathbf{x}) \leq f(\mathbf{x})$ for $\mathbf{x} \in \partial\Omega$ and all j . It follows from Lemma 9.8.3 that

$$u(\mathbf{x}) = \max\{u_1(\mathbf{x}), u_2(\mathbf{x}), \dots, u_j(\mathbf{x})\} \leq \max\{M_{u_1}(\mathbf{x}, r), M_{u_2}(\mathbf{x}, r), \dots, M_{u_j}(\mathbf{x}, r)\} \\ \leq M_u(\mathbf{x}, r)$$

for all $\mathbf{x} \in \Omega$ with r such that $\overline{B_r(\mathbf{x})} \subset \Omega$. Thus, u is subharmonic on Ω . ■

Lemma 9.8.7

Suppose that $u \in C(\overline{B_r(\mathbf{x})})$ and u is harmonic in $B_r(\mathbf{x})$, then

$$\left| \frac{\partial u}{\partial x_j}(\mathbf{x}) \right| \leq \frac{n}{r} \sup_{\mathbf{w} \in \partial B_r(\mathbf{x})} |u(\mathbf{w})|$$

for $1 \leq j \leq n$.

Proof.

It follows from Remark 9.5.17 that

$$u(\mathbf{z}) = \int_{\partial B_r(\mathbf{x})} u(\mathbf{y}) \frac{r^2 - \|\mathbf{z} - \mathbf{x}\|^2}{r \omega_n \|\mathbf{z} - \mathbf{y}\|^n} dS_{\mathbf{y}}$$

for $\mathbf{z} \in B_r(\mathbf{x})$. Since

$$(\mathbf{z}, \mathbf{y}) \rightarrow \frac{\partial}{\partial z_j} \left(\frac{r^2 - \|\mathbf{z} - \mathbf{x}\|^2}{\|\mathbf{z} - \mathbf{y}\|^n} \right) = \frac{-2(z_j - x_j) \|\mathbf{z} - \mathbf{y}\|^n - n(z_j - y_j)(r^2 - \|\mathbf{z} - \mathbf{x}\|^2) \|\mathbf{z} - \mathbf{y}\|^{n-2}}{\|\mathbf{z} - \mathbf{y}\|^{2n}}$$

is uniformly continuous on $\overline{B_s(\mathbf{x})} \times \partial B_r(\mathbf{x})$ for any $0 < s < r$, we get

$$\frac{\partial u}{\partial z_j}(\mathbf{z}) = \frac{1}{r \omega_n} \int_{\partial B_r(\mathbf{x})} u(\mathbf{y}) \left(\frac{-2(z_j - x_j) \|\mathbf{z} - \mathbf{y}\|^n - n(z_j - y_j)(r^2 - \|\mathbf{z} - \mathbf{x}\|^2) \|\mathbf{z} - \mathbf{y}\|^{n-2}}{\|\mathbf{z} - \mathbf{y}\|^{2n}} \right) dS_{\mathbf{y}}$$

for $\mathbf{z} \in B_r(\mathbf{x})$. Since $\|\mathbf{x} - \mathbf{y}\| = r$ for $\mathbf{y} \in \partial B_r(\mathbf{x})$, we obtain

$$\frac{\partial u}{\partial z_j}(\mathbf{x}) = -\frac{n}{r^{n+1} \omega_n} \int_{\partial B_r(\mathbf{x})} (x_j - y_j) u(\mathbf{y}) dS_{\mathbf{y}} = \frac{n}{r \omega_n} \int_{\partial B_1(\mathbf{0})} y_j u(\mathbf{x} + r\mathbf{y}) dS_{\mathbf{y}}.$$

Finally, since $|y_j u(\mathbf{x} + r\mathbf{y})| \leq |u(\mathbf{x} + r\mathbf{y})| \leq \sup_{\mathbf{w} \in \partial B_r(\mathbf{x})} |u(\mathbf{w})|$ for all $\mathbf{y} \in \partial B_1(\mathbf{0})$, we get

$$\begin{aligned} \left| \frac{\partial u}{\partial z_j}(\mathbf{x}) \right| &\leq \frac{n}{r \omega_n} \int_{\partial B_1(\mathbf{0})} |y_j u(\mathbf{x} + r\mathbf{y})| dS_{\mathbf{y}} \leq \frac{n}{r \omega_n} \sup_{\mathbf{w} \in \partial B_r(\mathbf{x})} |u(\mathbf{w})| \int_{\partial B_1(\mathbf{0})} dS_{\mathbf{y}} \\ &= \frac{n}{r} \sup_{\mathbf{w} \in \partial B_r(\mathbf{x})} |u(\mathbf{w})|. \end{aligned}$$

Lemma 9.8.8

u_f is harmonic on Ω .

Proof.

Choose $\mathbf{z} \in \Omega$ and $\rho > 0$ such that $\overline{B_\rho(\mathbf{z})} \subset \Omega$, and $0 < \eta < \rho$.

i) Given a sequence $\{\mathbf{x}_j\}_{j=1}^\infty \subset B_\eta(\mathbf{z})$. By definition of u_f , we can find for each j a sequence $\{v_{i,j}\}_{i=1}^\infty \subset \Sigma_f$ such that $\lim_{i \rightarrow \infty} v_{i,j}(\mathbf{x}_j) = u_f(\mathbf{x}_j)$ for $j \geq 1$.

Let $v_i = \max_{1 \leq j \leq i} v_{i,j}$. It follows from Lemma 9.8.6 that $v_i \in \Sigma_f$. Since $v_{i,j}(\mathbf{x}_j) \leq v_i(\mathbf{x}_j) \leq u_f(\mathbf{x}_j)$ for all $i \geq j$ and $\lim_{i \rightarrow \infty} v_{i,j}(\mathbf{x}_j) = u_f(\mathbf{x}_j)$, we get that $\lim_{i \rightarrow \infty} v_i(\mathbf{x}_j) = u_f(\mathbf{x}_j)$ for all j .

Substituting v_i by $\max\{u_m, v_i\}$, we get that $m \leq v_i(\mathbf{x}) \leq M$ for all $\mathbf{x} \in \Omega$. We still preserve $v_i \in \Sigma_f$ according to Lemma 9.8.6, and $\lim_{i \rightarrow \infty} v_i(\mathbf{x}_j) = u_f(\mathbf{x}_j)$ for all j .

Let $v_{i,\mathbf{z},\rho} = (v_i)_{\mathbf{z},\rho}$. We have that $v_{i,\mathbf{z},\rho}$ is harmonic on $B_\rho(\mathbf{z})$ for all i . We still have that $v_{i,\mathbf{z},\rho} \in \Sigma_f$ due to Lemma 9.8.2 and $v_{i,\mathbf{z},\rho} = v_i$ on $\partial\Omega$. Moreover, from Lemma 9.8.2, we also get that $v_i(\mathbf{x}_j) \leq v_{i,\mathbf{z},\rho}(\mathbf{x}_j) \leq u_f(\mathbf{x}_j)$ for all i and j . Thus $\lim_{i \rightarrow \infty} v_{i,\mathbf{z},\rho}(\mathbf{x}_j) = u_f(\mathbf{x}_j)$ for all j .

ii) Given $\mathbf{x} \in \overline{B_\eta(\mathbf{z})}$, let $r = \rho - \|\mathbf{x} - \mathbf{z}\|$. We have that $\overline{B_r(\mathbf{x})} \subset \overline{B_\rho(\mathbf{z})}$ and $r \geq \rho - \eta > 0$. Since $v_{i,\mathbf{z},\rho}$ is continuous on $\overline{B_\rho(\mathbf{z})}$ and harmonic on $B_\rho(\mathbf{z})$, it follows from Lemma 9.8.7 that

$$\left| \frac{\partial v_{i,\mathbf{z},\rho}}{\partial x_j}(\mathbf{x}) \right| \leq \frac{n}{r} \sup_{\mathbf{y} \in \partial B_r(\mathbf{x})} |v_{i,\mathbf{z},\rho}(\mathbf{y})| \leq \frac{n}{\rho - \eta} \sup_{\mathbf{y} \in \partial \Omega} |v_{i,\mathbf{z},\rho}(\mathbf{y})| = \frac{nM}{\rho - \eta}$$

for $1 \leq j \leq n$, where we have used the Maximum Principle Theorem for subharmonic functions on Ω to obtain the second inequality. Since $\mathbf{x} \in \overline{B_\eta(\mathbf{z})}$ is arbitrary, we have that

$$\left| \frac{\partial v_{i,\mathbf{z},\rho}}{\partial x_j}(\mathbf{x}) \right| \leq \frac{nM}{\rho - \eta}$$

for all $\mathbf{x} \in \overline{B_\eta(\mathbf{z})}$ and $1 \leq j \leq n$. Thus $\|D v_{i,\mathbf{z},\rho}(\mathbf{x})\| \leq n\sqrt{n}M/(\rho - \eta)$ for all $\mathbf{x} \in \overline{B_\eta(\mathbf{z})}$ and all i . Combining this relation with the Mean Value Theorem for functions of several variables, it is easy to show that $\{v_{i,\mathbf{z},\rho}\}_{j=1}^\infty$ is equicontinuous on $\overline{B_\eta(\mathbf{z})}$. Since $m \leq v_{i,\mathbf{z},\rho}(\mathbf{x}) \leq M$ for all $\mathbf{x} \in \overline{B_\eta(\mathbf{z})}$, we may use Arzela-Ascoli Theorem⁶ to conclude that there exists a subsequence $\{v_{i_k,\mathbf{z},\rho}\}_{k=1}^\infty$ that converge uniformly on $\overline{B_\eta(\mathbf{z})}$.

iii) Let $v(\mathbf{x}) \equiv \lim_{k \rightarrow \infty} v_{i_k,\mathbf{z},\rho}(\mathbf{x})$ for all $\mathbf{x} \in \Omega$. Since $\{v_{i_k,\mathbf{z},\rho}\}_{k=1}^\infty$ is a sequence of harmonic functions on $B_\eta(\mathbf{z})$ and converges uniformly to v on $\overline{B_\eta(\mathbf{z})}$, we get from Corollary 9.1.6 that v is harmonic on $B_\eta(\mathbf{z})$ and continuous on $\overline{B_\eta(\mathbf{z})}$.

We also have that $v(\mathbf{x}_j) = u_f(\mathbf{x}_j)$ for all j because $\lim_{i \rightarrow \infty} v_{i,\mathbf{z},\rho}(\mathbf{x}_j) = u_f(\mathbf{x}_j)$ for all j .

iv) Let us show that u_f is continuous on $B_\eta(\mathbf{z})$. Given $\mathbf{w} \in B_\eta(\mathbf{z})$, choose a sequence $\{\mathbf{x}_j\}_{j=1}^\infty \subset B_\eta(\mathbf{z})$ such that $\mathbf{x}_1 = \mathbf{w}$ and $\{\mathbf{x}_j\}_{j=2}^\infty$ converges to $\mathbf{w}$. Since v is continuous⁷ on $\overline{B_\eta(\mathbf{z})}$, we then have that $\lim_{j \rightarrow \infty} u_f(\mathbf{x}_j) = \lim_{j \rightarrow \infty} v(\mathbf{x}_j) = v(\mathbf{w}) = v(\mathbf{x}_1) = u_f(\mathbf{x}_1) = u_f(\mathbf{w})$. Since this is true for all sequences converging to $\mathbf{w}$, we have that u_f is continuous at $\mathbf{w}$. Since $\mathbf{w} \in B_\eta(\mathbf{z})$ is arbitrary, u_f is continuous on $B_\eta(\mathbf{z})$.

v) Let $\{\mathbf{x}_j\}_{j=1}^\infty$ be a dense sequence in $B_\eta(\mathbf{z})$. Since u_f and the function v associated to this sequence are continuous functions that agree on the dense set $\{\mathbf{x}_j\}_{j=1}^\infty$, we have that $u_f = v$ on $\overline{B_\eta(\mathbf{z})}$. Thus u_f is harmonic on $B_\eta(\mathbf{z})$. Since this is true for all $\mathbf{z} \in \Omega$ and all $\eta < \rho$ such that $\overline{B_\rho(\mathbf{z})} \subset \Omega$, we finally get that u_f is harmonic on Ω . ■

Since the beginning, we have assumed that the boundary of Ω was a manifold of class C^2 . Using Perron's method to solve the Dirichlet problem, we may impose less restrictive conditions on the boundary of Ω .

Definition 9.8.9 (Barrier Condition)

The boundary $\partial\Omega$ of an open subset $\Omega \subset \mathbb{R}^n$ satisfies the **barrier condition** if, for every $\mathbf{z} \in \partial\Omega$, there exists a function $g_{\mathbf{z}} \in C(\overline{\Omega})$ such that $g_{\mathbf{z}}$ is subharmonic on Ω , $g_{\mathbf{z}}(\mathbf{z}) = 0$ and $g_{\mathbf{z}}(\mathbf{x}) < 0$ for all $\mathbf{x} \in \partial\Omega \setminus \{\mathbf{z}\}$.

⁶In some textbook, it is referred as Arzela Theorem only and in others as Ascoli Theorem only. Why?

⁷Note that the sequence $\{v_{i,j}\}_{i=1}^\infty$ will change if we change the sequence $\{\mathbf{x}_j\}_{j=1}^\infty \subset B_\eta(\mathbf{z})$ and so may v .

We now show that a bounded open set Ω with a boundary of class C^2 satisfies the barrier condition. For each $\mathbf{z} \in \partial\Omega$, there exist $\mathbf{w}$ on the line $\{\mathbf{z} + t\nu(\mathbf{z}) : t > 0\}$ and $s > 0$ small enough such that $\overline{B_s(\mathbf{w})} \cap \overline{\Omega} = \{\mathbf{z}\}$ because $\partial\Omega$ is a C^2 -manifold⁸ (Figure 9.6). As usual, $\nu(\mathbf{z})$ is the outward unit normal to $\partial\Omega$ at $\mathbf{z} \in \partial\Omega$. Let

$$g_{\mathbf{z}}(\mathbf{x}) = \begin{cases} \left\| \frac{1}{s}(\mathbf{x} - \mathbf{w}) \right\|^{2-n} - 1 & \text{if } n > 2 \\ -\ln \left(\left\| \frac{1}{s}(\mathbf{x} - \mathbf{w}) \right\| \right) & \text{if } n = 2 \end{cases}$$

We have from Corollary 8.3.3 that $g_{\mathbf{z}}$ is harmonic on $\mathbb{R}^n \setminus \{\mathbf{w}\}$. We have also by construction that $|g_{\mathbf{z}}(\mathbf{x})| < 0$ for $\mathbf{x} \notin \overline{B_s(\mathbf{w})}$ because $\|\mathbf{x} - \mathbf{w}\| > s$ in that case, and $g_{\mathbf{z}}(\mathbf{x}) = 0$ for $\mathbf{x} \in \partial B_s(\mathbf{w})$ because $\|\mathbf{x} - \mathbf{w}\| = s$ in that case. Hence, $g_{\mathbf{z}} \in C(\overline{\Omega})$ is harmonic on Ω , and satisfies $g_{\mathbf{z}}(\mathbf{z}) = 0$ because $\mathbf{z} \in \partial B_s(\mathbf{w})$ and $g_{\mathbf{z}}(\mathbf{x}) < 0$ for $\mathbf{x} \in \partial\Omega \setminus \{\mathbf{z}\}$ because $\mathbf{x} \notin \overline{B_s(\mathbf{w})}$.

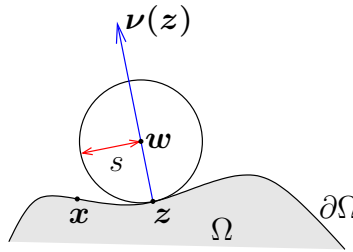


Figure 9.6: Figure associated to the construction of the function $g_{\mathbf{z}}$ used to demonstrate that $\partial\Omega$ of class C^2 satisfies the barrier condition.

There are sets $\Omega \subset \mathbb{R}^n$ that do not have a boundary of class C^2 but satisfy the barrier condition. The simplest examples that come to mind are the sets Ω which are strictly convex. Therefore, for each point $\mathbf{z} \in \partial\Omega$, there exists a hyperplane $P_{\mathbf{z}} = \{\mathbf{y} \in \mathbb{R}^n : \langle \mathbf{y}, \mathbf{w} \rangle = 0\}$ for some $\mathbf{w}$ such that $(\mathbf{z} + P_{\mathbf{z}}) \cap \overline{\Omega} = \{\mathbf{z}\}$. We may assume that $\mathbf{w}$ is a unit vector pointing outside of Ω . Then $g_{\mathbf{z}} : \mathbb{R}^n \rightarrow \mathbb{R}$ defined by $g_{\mathbf{z}}(\mathbf{x}) = \langle \mathbf{x} - \mathbf{z}, \mathbf{w} \rangle$ is an harmonic function, since it is linear, such that $g_{\mathbf{z}}(\mathbf{z}) = 0$ and $g_{\mathbf{z}}(\mathbf{x}) < 0$ for $\mathbf{x} \in \Omega \setminus \{\mathbf{z}\}$ because $\mathbf{x} - \mathbf{z}$ is on the opposite site of $P_{\mathbf{z}}$ than $\mathbf{w}$ (Figure 9.7). Note that $g_{\mathbf{z}}(\mathbf{x})$ is the orthogonal projection of the vector $\mathbf{x} - \mathbf{z}$ along $\mathbf{w}$.

For the rest of this section, we assume that Ω is a bounded open subset of $\mathbb{R}^n$ and satisfies the barrier condition.

Lemma 9.8.10

$$\liminf_{\mathbf{x} \rightarrow \mathbf{z}} u_f(\mathbf{x}) \geq f(\mathbf{z}) \text{ for all } \mathbf{z} \in \partial\Omega.$$

Proof.

Given $\mathbf{z} \in \partial\Omega$ and $\epsilon > 0$, we show that there exists $v_{\epsilon} \in \Sigma_f$ such that $\lim_{\mathbf{x} \rightarrow \mathbf{z}} v_{\epsilon}(\mathbf{x}) \geq f(\mathbf{z}) - \epsilon$.

⁸This can be proved using Lemma 9.9.32 in the next chapter.

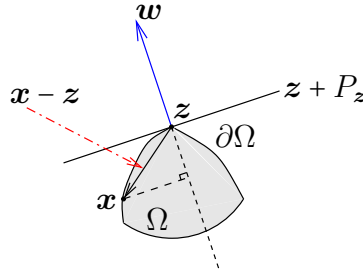


Figure 9.7: Figure associated to the construction of the function g_z used to demonstrate that $\partial\Omega$ satisfies the barrier condition if Ω is strictly convex. $\partial\Omega$ may not be differentiable everywhere. For a clearer visualization, the vectors $\mathbf{w}$ and $\mathbf{x} - \mathbf{z}$ have been translated by $\mathbf{z}$ in the figure above to have then start at the point $\mathbf{z}$.

We will then get from $u_f(\mathbf{x}) \geq v_\epsilon(\mathbf{x})$ for all $\mathbf{x} \in \overline{\Omega}$ that $\liminf_{\mathbf{x} \rightarrow \mathbf{z}} u_f(\mathbf{x}) \geq \lim_{\mathbf{x} \rightarrow \mathbf{z}} v_\epsilon(\mathbf{x}) \geq f(\mathbf{z}) - \epsilon$. Thus, $\liminf_{\mathbf{x} \rightarrow \mathbf{z}} u_f(\mathbf{x}) \geq f(\mathbf{z})$ because $\epsilon > 0$ is arbitrary.

Since $f : \partial\Omega \rightarrow \mathbb{R}$ is continuous, there exists $\delta > 0$ such that $|f(\mathbf{x}) - f(\mathbf{z})| < \epsilon$ for $\mathbf{x} \in \partial\Omega \cap B_\delta(\mathbf{z})$. Thus $f(\mathbf{z}) - \epsilon < f(\mathbf{x})$ for $\mathbf{x} \in \partial\Omega \cap B_\delta(\mathbf{z})$.

Moreover, since g_z given in the definition of the barrier condition is continuous on the compact set $\partial\Omega \setminus B_\delta(\mathbf{z})$, we have that $b_g = \max_{\mathbf{x} \in \partial\Omega \setminus B_\delta(\mathbf{z})} g_z(\mathbf{x}) = g_z(\mathbf{y})$ for some $\mathbf{y} \in \partial\Omega \setminus B_\delta(\mathbf{z})$. Thus, $b_g < 0$. Similarly, since f is continuous on the compact set $\partial\Omega \setminus B_\delta(\mathbf{z})$, we have that $b_f = \min_{\mathbf{x} \in \partial\Omega \setminus B_\delta(\mathbf{z})} f(\mathbf{x}) > -\infty$. Choose C such that $f(\mathbf{z}) - \epsilon + Cb_g \leq b_f$ and consider

$$\begin{aligned} u_\epsilon : \overline{\Omega} &\rightarrow \mathbb{R} \\ \mathbf{x} &\mapsto f(\mathbf{z}) - \epsilon + Cg_z(\mathbf{x}) \end{aligned}$$

Since $g_z \in C(\overline{\Omega})$ is subharmonic on Ω , so is u_ϵ . Moreover, $u_\epsilon(\mathbf{x}) \leq f(\mathbf{z}) - \epsilon \leq f(\mathbf{x})$ for all $\mathbf{x} \in \partial\Omega \cap B_\delta(\mathbf{z})$, and $u_\epsilon(\mathbf{x}) = f(\mathbf{z}) - \epsilon + Cg_z(\mathbf{x}) \leq f(\mathbf{z}) - \epsilon + Cb_g \leq b_f \leq f(\mathbf{x})$ for all $\mathbf{x} \in \partial\Omega \setminus B_\delta(\mathbf{z})$. We therefore have that $u_\epsilon \in \Sigma_f$. ■

Lemma 9.8.11

$$\lim_{\mathbf{x} \rightarrow \mathbf{z}} u_f(\mathbf{x}) = f(\mathbf{z}) \text{ for all } \mathbf{z} \in \partial\Omega.$$

Proof.

It follows from the previous lemma that we only have to prove that $\limsup_{\mathbf{x} \rightarrow \mathbf{z}} u_f(\mathbf{x}) \leq f(\mathbf{z})$ for all $\mathbf{z} \in \partial\Omega$.

We have that $u_{-f}(\mathbf{x}) = \sup_{u \in \Sigma_{-f}} u(\mathbf{x}) = - \inf_{u \in \Sigma_{-f}} -u(\mathbf{x})$. Note that $u_{-f}(\mathbf{x}) \leq -f(\mathbf{x})$ for all $\mathbf{x} \in \partial\Omega$ by definition of Σ_{-f} .

If $v \in \Sigma_{-f}$, then $v(\mathbf{x}) \leq -f(\mathbf{x})$ for all $\mathbf{x} \in \partial\Omega$. If $u \in \Sigma_f$, then $u(\mathbf{x}) \leq f(\mathbf{x})$ for all $\mathbf{x} \in \partial\Omega$. Hence, $u(\mathbf{x}) + v(\mathbf{x}) \leq 0$ for all $\mathbf{x} \in \partial\Omega$ whatever $v \in \Sigma_{-f}$ and $u \in \Sigma_f$. Since $u + v \in C(\overline{\Omega})$ is subharmonic on Ω , we get from the Maximum Principle Theorem, Theorem 9.2.3, that $u(\mathbf{x}) \leq -v(\mathbf{x})$ for all $\mathbf{x} \in \overline{\Omega}$ whatever $v \in \Sigma_{-f}$ and $u \in \Sigma_f$. Thus, $u_f(\mathbf{x}) = \sup_{u \in \Sigma_f} u(\mathbf{x}) \leq -v(\mathbf{x})$ for all $\mathbf{x} \in \overline{\Omega}$ and all $v \in \Sigma_{-f}$. It follows that $u_f(\mathbf{x}) = \sup_{u \in \Sigma_f} u(\mathbf{x}) \leq \inf_{v \in \Sigma_{-f}} -v(\mathbf{x}) = -u_{-f}(\mathbf{x})$ for all $\mathbf{x} \in \overline{\Omega}$.

Hence, for $\mathbf{z} \in \partial\Omega$, $\limsup_{\mathbf{x} \rightarrow \mathbf{z}} u_f(\mathbf{x}) \leq \limsup_{\mathbf{x} \rightarrow \mathbf{z}} -u_{-f}(\mathbf{x}) = -\liminf_{\mathbf{x} \rightarrow \mathbf{z}} u_{-f}(\mathbf{x}) \leq f(\mathbf{z})$ because $\liminf_{\mathbf{x} \rightarrow \mathbf{z}} u_{-f}(\mathbf{x}) \geq -f(\mathbf{z})$ according to the previous lemma applied to $-f$. ■

We have finally proved the following result.

Theorem 9.8.12

If $\Omega \subset \mathbb{R}^n$ is a bounded open set such that $\partial\Omega$ satisfies the barrier condition, and if $f : \partial\Omega \rightarrow \mathbb{R}$ is a continuous function, then there exists a unique function $u \in C(\overline{\Omega})$ which is harmonic on Ω and satisfies $u(\mathbf{x}) = f(\mathbf{x})$ for all $\mathbf{x} \in \partial\Omega$.

Proof.

According to Lemmas 9.8.8 and 9.8.11, the function $u_f : \overline{\Omega} \rightarrow \mathbb{R}$ is a continuous function on $\overline{\Omega}$, harmonic on Ω , and satisfying $u_f(\mathbf{x}) = f(\mathbf{x})$ for all $\mathbf{x} \in \partial\Omega$. According to Theorem 9.2.6, u_f is the unique function with such properties. ■

9.9 Layer Potentials

We follow closely the content of [10] on the subject of layer potentials.

In this chapter, Ω is a bounded open subset of $\mathbb{R}^n$. We assume that $\Omega = \bigcup_{i=1}^m \Omega_i$, where the Ω_i are the connected components of Ω . In particular, $\Omega_i \cap \Omega_j = \emptyset$ for $i \neq j$. Moreover, we assume that $\Omega' = \mathbb{R}^n \setminus \overline{\Omega} = \bigcup_{i=0}^{m'} \Omega'_i$, where the Ω'_i are the connected components of Ω' with Ω'_0 being the unbounded component. We assume that $\partial\Omega$ is C^2 and observe that $\partial\Omega = \partial\Omega'$.

As usual, let $\nu(\mathbf{x})$ denote the outward unit normal to $\partial\Omega$ at $\mathbf{x} \in \partial\Omega$. The **interior normal derivative** of $u \in C^1(\Omega) \cap C(\overline{\Omega})$ at $\mathbf{x} \in \partial\Omega$ is

$$\frac{\partial u}{\partial \nu^-}(\mathbf{x}) = \lim_{t \rightarrow 0^-} \frac{1}{t} (u(\mathbf{x} + t\nu(\mathbf{x})) - u(\mathbf{x}))$$

for $\mathbf{x} \in \partial\Omega$ if the limit exists. Similarly, the **exterior normal derivative** of $u \in C^1(\Omega') \cap C(\overline{\Omega}')$ at $\mathbf{x} \in \partial\Omega'$ is

$$\frac{\partial u}{\partial \nu^+}(\mathbf{x}) = \lim_{t \rightarrow 0^+} \frac{1}{t} (u(\mathbf{x} + t\nu(\mathbf{x})) - u(\mathbf{x}))$$

for $\mathbf{x} \in \partial\Omega$ if the limit exists.

The following spaces of functions will be used for the studies of Neumann problems. $C_\nu(\Omega)$ is the space of functions $u \in C^1(\Omega) \cap C(\bar{\Omega})$ such that the interior normal derivative exists at all points $\mathbf{x} \in \partial\Omega$ and

$$\max_{\mathbf{x} \in \partial\Omega} \left| \nabla u(\mathbf{x} + t\nu(\mathbf{x})) \cdot \nu(\mathbf{x}) - \frac{\partial u}{\partial \nu^-}(\mathbf{x}) \right| \rightarrow 0 \quad \text{as } t \rightarrow 0^- . \quad (9.9.1)$$

Similarly, $C_\nu(\Omega')$ is the space of functions $u \in C^1(\Omega') \cap C(\bar{\Omega}')$ such that the exterior normal derivative exists at all points $\mathbf{x} \in \partial\Omega'$ and

$$\max_{\mathbf{x} \in \partial\Omega'} \left| \nabla u(\mathbf{x} + t\nu(\mathbf{x})) \cdot \nu(\mathbf{x}) - \frac{\partial u}{\partial \nu^+}(\mathbf{x}) \right| \rightarrow 0 \quad \text{as } t \rightarrow 0^+ . \quad (9.9.2)$$

The main goal of this chapter is to solve the following four problems.

D1 - Interior Dirichlet Problem: Given $f \in C(\partial\Omega)$, find $u \in C(\bar{\Omega})$ such that $\Delta u = 0$ in Ω (in the sense of distributions) and $u|_{\partial\Omega} = f$.

D2 - Exterior Dirichlet Problem: Given $f \in C(\partial\Omega')$, find $u \in C(\bar{\Omega}')$ such that $\Delta u = 0$ in $\Omega' \cup \{\infty\}$ (in the sense of distributions) and $u|_{\partial\Omega} = f$.

N1 - Interior Neumann Problem: Given $f \in C(\partial\Omega)$, find $u \in C_\nu(\Omega)$ such that $\Delta u = 0$ in Ω (in the sense of distributions) and $\frac{\partial u}{\partial \nu^-} = f$ on $\partial\Omega$.

N2 - Exterior Dirichlet Problem: Given $f \in C(\partial\Omega')$, find $u \in C_\nu(\Omega')$ such that $\Delta u = 0$ in $\Omega' \cup \{\infty\}$ (in the sense of distributions) and $\frac{\partial u}{\partial \nu^+} = f$ on $\partial\Omega$.

We need to explain what it means for a function to be harmonic at infinity in (D2) and (N2).

Suppose that $g: V \rightarrow W$ is a C^∞ diffeomorphism between two open subsets V and W of $\mathbb{R}^n$. Let $G(\mathbf{x}) = (\mathbf{D}g(\mathbf{x}))(\mathbf{D}g(\mathbf{x}))^\top$ for $\mathbf{x} \in V$. The components of $G(\mathbf{x})$ are

$$G_{i,j}(\mathbf{x}) = \sum_{k=1}^n \frac{\partial g_i}{\partial x_k}(\mathbf{x}) \frac{\partial g_j}{\partial x_k}(\mathbf{x}) \quad , \quad 1 \leq i, j \leq n .$$

Let $h = g^{-1}$ and $H(\mathbf{y}) = (\mathbf{D}h(\mathbf{y}))^\top (\mathbf{D}h(\mathbf{y}))$ for $\mathbf{y} \in W$. The components of $H(\mathbf{y})$ are

$$H_{i,j}(\mathbf{y}) = \sum_{k=1}^n \frac{\partial h_k}{\partial y_i}(\mathbf{y}) \frac{\partial h_k}{\partial y_j}(\mathbf{y}) \quad , \quad 1 \leq i, j \leq n .$$

We have that $G(h(\mathbf{y})) = (H(\mathbf{y}))^{-1}$ for all $\mathbf{y} \in W$ because

$$\text{Id} = \mathbf{D}(g(h(\mathbf{y}))) = \mathbf{D}g(h(\mathbf{y}))(\mathbf{D}h(\mathbf{y}))$$

for all $\mathbf{y} \in W$ implies that $Dg(h(\mathbf{y})) = (Dh(\mathbf{y}))^{-1}$ for all $\mathbf{y} \in W$. Thus,

$$G(h(\mathbf{y})) = (Dg(h(\mathbf{y}))(Dg(h(\mathbf{y})))^\top = (Dh(\mathbf{y}))^{-1} \left((Dh(\mathbf{y}))^{-1} \right)^\top = (H(\mathbf{y}))^{-1}$$

for all $\mathbf{y} \in W$

Let $d_h(\mathbf{y}) = |\det(Dh(\mathbf{y}))|$ for $\mathbf{y} \in W$. We have that

$$d_h(g(\mathbf{x})) = |\det(Dh(g(\mathbf{x})))| = |\det(Dg(\mathbf{x}))^{-1}| = |\det(Dg(\mathbf{x}))|^{-1}$$

for all $\mathbf{x} \in V$.

Suppose that $u \in C^2(V)$ and let $\check{u}(\mathbf{y}) = u(h(\mathbf{y}))$ for all $\mathbf{y} \in W$. We have the following result.

Proposition 9.9.1

Suppose that $u, g, h = g^{-1}$ and G are as defined above. Then

$$\Delta u(\mathbf{x}) = \left(\frac{1}{d_h(\mathbf{y})} \sum_{i,j=1}^n \frac{\partial}{\partial y_j} \left(d_h(\mathbf{y}) G_{i,j}(h(\mathbf{y})) \frac{\partial \check{u}}{\partial y_i}(\mathbf{y}) \right) \right) \Big|_{\mathbf{y}=g(\mathbf{x})} \quad (9.9.3)$$

in the sense of distributions for $\mathbf{x} \in V$.

Proof.

Given $\phi \in \mathcal{D}(V)$, let $\check{\phi}(\mathbf{y}) = \phi(h(\mathbf{y}))$ for $\mathbf{y} \in W$. We have

$$\begin{aligned} \int_V (\Delta u) \phi \, d\mathbf{x} &= \sum_{k=1}^n \int_V \frac{\partial^2 u}{\partial x_k^2} \phi \, d\mathbf{x} = - \sum_{k=1}^n \int_V \frac{\partial u}{\partial x_k} \frac{\partial \phi}{\partial x_k} \, d\mathbf{x} \\ &= - \sum_{k=1}^n \int_W \left(\sum_{i=1}^n \frac{\partial \check{u}}{\partial y_i}(\mathbf{y}) \frac{\partial g_i}{\partial x_k}(h(\mathbf{y})) \right) \left(\sum_{j=1}^n \frac{\partial \check{\phi}}{\partial y_j}(\mathbf{y}) \frac{\partial g_j}{\partial x_k}(h(\mathbf{y})) \right) d_h(\mathbf{y}) \, d\mathbf{y} \\ &= - \sum_{j,i=1}^n \int_W \frac{\partial \check{u}}{\partial y_i}(\mathbf{y}) \frac{\partial \check{\phi}}{\partial y_j}(\mathbf{y}) \left(\sum_{k=1}^n \frac{\partial g_i}{\partial x_k}(h(\mathbf{y})) \frac{\partial g_j}{\partial x_k}(h(\mathbf{y})) \right) d_h(\mathbf{y}) \, d\mathbf{y} \\ &= - \sum_{j,i=1}^n \int_W \frac{\partial \check{u}}{\partial y_i}(\mathbf{y}) \frac{\partial \check{\phi}}{\partial y_j}(\mathbf{y}) G_{i,j}(h(\mathbf{y})) d_h(\mathbf{y}) \, d\mathbf{y} \\ &= - \sum_{j,i=1}^n \int_W \left(\frac{\partial \check{u}}{\partial y_i}(\mathbf{y}) G_{i,j}(h(\mathbf{y})) d_h(\mathbf{y}) \right) \frac{\partial \check{\phi}}{\partial y_j}(\mathbf{y}) \, d\mathbf{y} \\ &= \sum_{j,i=1}^n \int_W \frac{\partial}{\partial y_j} \left(\frac{\partial \check{u}}{\partial y_i}(\mathbf{y}) G_{i,j}(h(\mathbf{y})) d_h(\mathbf{y}) \right) \check{\phi}(\mathbf{y}) \, d\mathbf{y} \\ &= \int_V \left(\sum_{j,i=1}^n \frac{\partial}{\partial y_j} \left(\frac{\partial \check{u}}{\partial y_i}(\mathbf{y}) G_{i,j}(h(\mathbf{y})) d_h(\mathbf{y}) \right) \right) \Big|_{\mathbf{y}=g(\mathbf{x})} \phi(\mathbf{x}) |\det(Dg(\mathbf{x}))| \, d\mathbf{x} \\ &= \int_V \left(\frac{1}{d_h(\mathbf{y})} \sum_{j,i=1}^n \frac{\partial}{\partial y_j} \left(\frac{\partial \check{u}}{\partial y_i}(\mathbf{y}) G_{i,j}(h(\mathbf{y})) d_h(\mathbf{y}) \right) \right) \Big|_{\mathbf{y}=g(\mathbf{x})} \phi(\mathbf{x}) \, d\mathbf{x} . \quad \blacksquare \end{aligned}$$

Consider the inversion with respect to the unit ball in $\mathbb{R}^n$,

$$T: \mathbb{R}^n \setminus \{\mathbf{0}\} \rightarrow \mathbb{R}^n \setminus \{\mathbf{0}\}$$

$$\mathbf{x} \mapsto \frac{1}{\|\mathbf{x}\|_2^2} \mathbf{x}$$

We have that $T = T^{-1}$. Without loss of generality, we may assume that $\mathbf{0} \in \Omega$. The **Kelvin transform** of a function $u \in C(\Omega')$, is the function

$$\tilde{u}: T(\Omega') \rightarrow \mathbb{R}$$

$$\mathbf{y} \mapsto \|\mathbf{y}\|_2^{2-n} u(T(\mathbf{y}))$$

The function T is well defined on $T(\Omega')$ because $\mathbf{0} \notin \Omega'$.

Let $h(\mathbf{y}) = T(\mathbf{y})$ for all $\mathbf{y} \in T(\Omega')$. So, the set W in the previous discussion is now $T(\Omega')$ and the set V is now Ω' . As before, $\tilde{u}(\mathbf{y}) = u(h(\mathbf{y})) = u(T(\mathbf{u}))$.

We have that

$$\frac{\partial h_k}{\partial y_i} = \frac{\partial}{\partial y_i} \left(y_k \left(\sum_{j=1}^n y_j^2 \right)^{-1} \right) = -\frac{2y_k y_i}{\|\mathbf{y}\|_2^4} + \frac{\delta_{i,k}}{\|\mathbf{y}\|_2^2}$$

for $1 \leq i, k \leq n$. Hence,

$$\begin{aligned} H_{i,j}(\mathbf{y}) &= \sum_{k=1}^n \left(-\frac{2y_i y_k}{\|\mathbf{y}\|_2^4} + \frac{\delta_{i,k}}{\|\mathbf{y}\|_2^2} \right) \left(-\frac{2y_j y_k}{\|\mathbf{y}\|_2^4} + \frac{\delta_{j,k}}{\|\mathbf{y}\|_2^2} \right) \\ &= \frac{4y_i y_j}{\|\mathbf{y}\|_2^8} \sum_{k=1}^n y_k^2 - \sum_{k=1}^n \frac{2y_i y_k \delta_{j,k}}{\|\mathbf{y}\|_2^6} - \sum_{k=1}^n \frac{2y_j y_k \delta_{i,k}}{\|\mathbf{y}\|_2^6} + \sum_{k=1}^n \frac{\delta_{i,k} \delta_{j,k}}{\|\mathbf{y}\|_2^4} \\ &= \frac{4y_i y_j}{\|\mathbf{y}\|_2^6} - \frac{2y_i y_j}{\|\mathbf{y}\|_2^6} - \frac{2y_j y_i}{\|\mathbf{y}\|_2^6} + \frac{\delta_{i,j}}{\|\mathbf{y}\|_2^4} = \frac{\delta_{i,j}}{\|\mathbf{y}\|_2^4} \end{aligned}$$

for $1 \leq i, j \leq n$. It follows that

$$d_h(\mathbf{y}) = |\det(Dh)(\mathbf{y})| = |\det H(\mathbf{y})|^{1/2} = \frac{1}{\|\mathbf{y}\|_2^{2n}}.$$

We also have that $G_{i,j}(h(\mathbf{y})) = \delta_{i,j} \|\mathbf{y}\|_2^4$ because $G(h(\mathbf{y})) = (H(\mathbf{y}))^{-1}$ for all $\mathbf{y} \in T(\Omega')$.

It follows from (9.9.3) with $h = T$ and $u \in C^2(\Omega')$ that

$$\begin{aligned} \Delta u(\mathbf{x}) &= \|\mathbf{y}\|_2^{2n} \sum_{i,j=1}^n \frac{\partial}{\partial y_j} \left(\frac{\delta_{i,j}}{\|\mathbf{y}\|_2^{2n-4}} \frac{\partial \tilde{u}}{\partial y_i}(\mathbf{y}) \right) = \|\mathbf{y}\|_2^{2n} \sum_{i=1}^n \frac{\partial}{\partial y_i} \left(\frac{1}{\|\mathbf{y}\|_2^{2n-4}} \frac{\partial \tilde{u}}{\partial y_i}(\mathbf{y}) \right) \\ &= \|\mathbf{y}\|_2^{2n} \sum_{i=1}^n \left((2-n) \frac{2y_i}{\|\mathbf{y}\|_2^{2n-2}} \frac{\partial \tilde{u}}{\partial y_i}(\mathbf{y}) + \frac{1}{\|\mathbf{y}\|_2^{2n-4}} \frac{\partial^2 \tilde{u}}{\partial y_i^2}(\mathbf{y}) \right) \\ &= \|\mathbf{y}\|_2^{n+2} \sum_{i=1}^n \left((2-n) \frac{2y_i}{\|\mathbf{y}\|_2^n} \frac{\partial \tilde{u}}{\partial y_i}(\mathbf{y}) + \frac{1}{\|\mathbf{y}\|_2^{n-2}} \frac{\partial^2 \tilde{u}}{\partial y_i^2}(\mathbf{y}) \right) \\ &= \|\mathbf{y}\|_2^{n+2} \sum_{i=1}^n \left(\frac{\partial^2}{\partial y_i^2} \left(\frac{1}{\|\mathbf{y}\|_2^{n-2}} \right) \tilde{u}(\mathbf{y}) + 2 \frac{\partial}{\partial y_i} \left(\frac{1}{\|\mathbf{y}\|_2^{n-2}} \right) \frac{\partial \tilde{u}}{\partial y_i}(\mathbf{y}) + \frac{1}{\|\mathbf{y}\|_2^{n-2}} \frac{\partial^2 \tilde{u}}{\partial y_i^2}(\mathbf{y}) \right) \end{aligned}$$

$$= \|\mathbf{y}\|_2^{n+2} \sum_{i=1}^n \frac{\partial^2}{\partial y_i^2} \left(\frac{1}{\|\mathbf{y}\|_2^{n-2}} \check{u}(\mathbf{y}) \right)$$

for $\mathbf{y} = T(\mathbf{x})$, we here we have used $\Delta \|\mathbf{y}\|_2^{2-n} = 0$ for $\mathbf{y} \neq \mathbf{0}$ for the second to last equality.

We have proved that

$$\Delta u(\mathbf{x}) = \left(\|\mathbf{y}\|_2^{n+2} \Delta_{\mathbf{y}} \check{u}(\mathbf{y}) \right) \Big|_{\mathbf{y}=T(\mathbf{x})}$$

for $\mathbf{x} \in \Omega'$. The following result follows from this relation.

Proposition 9.9.2

If $u \in C^2(\Omega')$, then u is harmonic on Ω' if and only if $\check{u}$ is harmonic on $T(\Omega')$.

Definition 9.9.3

We say that $u \in C(\Omega')$ is harmonic at infinity if $\check{u}$ has a removable singularity at $\mathbf{0}$. Namely, $\check{u}$ can be extended to $\mathbf{x} = \mathbf{0}$ in such a way that $\check{u}$ is harmonic on a neighbourhood of $\mathbf{0}$.

We will need conditions later to determine when a function is harmonic at infinity. To obtain this conditions, we first need the following lemma.

Lemma 9.9.4

Suppose that $U \subset \mathbb{R}^n$ is an open set and that $\mathbf{p} \in \mathbb{R}^n$. If $u : U \rightarrow \mathbb{R}$ is an harmonic function on $U \setminus \{\mathbf{p}\}$ such that $|u(\mathbf{x})| = o(\|\mathbf{x} - \mathbf{p}\|^{2-n})$ as $\|\mathbf{x} - \mathbf{p}\| \rightarrow 0$ when $n > 2$ or $|u(\mathbf{x})| = o(\ln(\|\mathbf{x} - \mathbf{p}\|^{-1}))$ as $\|\mathbf{x} - \mathbf{p}\| \rightarrow 0$ when $n = 2$, then u has a removable singularity at $\mathbf{p}$; namely, we may define u at $\mathbf{p}$ such that u is harmonic on U .

Proof.

Since U is open, there exists $R > 0$ such that $\overline{B_R(\mathbf{p})} \subset U$. Let $V = R^{-1}(U - \{\mathbf{p}\})$. The lemma will be proved if we prove that $v : V \rightarrow \mathbb{R}$ defined by $v(\mathbf{x}) = u(R(\mathbf{x} + \mathbf{p}))$ has a removable singularity at $\mathbf{0}$. We may therefore assume that $\overline{B_1(\mathbf{0})} \subset U$.

Since v is continuous on $\partial B_1(\mathbf{0})$, it follows from Theorem 9.5.16 that there exists a continuous function $w : \overline{B_1(\mathbf{0})} \rightarrow \mathbb{R}$ such that w is harmonic on $B_1(\mathbf{0})$ and $w(\mathbf{x}) = v(\mathbf{x})$ for $\mathbf{x} \in \partial B_1(\mathbf{0})$.

We now prove that $v = w$ on $\overline{B_1(\mathbf{0})} \setminus \{\mathbf{0}\}$. We will therefore be able to remove the singularity at the origin by setting $v(\mathbf{0}) = w(\mathbf{0})$.

Given $\epsilon > 0$ and $0 < \delta < 1$, let $f : \overline{B_1(\mathbf{0})} \setminus B_\delta(\mathbf{0}) \rightarrow \mathbb{R}$ be the function defined by

$$f(\mathbf{x}) = \begin{cases} v(\mathbf{x}) - w(\mathbf{x}) - \epsilon(\|\mathbf{x}\|^{2-n} - 1) & \text{if } n > 2 \\ v(\mathbf{x}) - w(\mathbf{x}) - \epsilon \ln(\|\mathbf{x}\|^{-1}) & \text{if } n = 2 \end{cases}$$

We have that f is continuous on $\overline{B_1(\mathbf{0})} \setminus B_\delta(\mathbf{0})$ and harmonic on $B_1(\mathbf{0}) \setminus \overline{B_\delta(\mathbf{0})}$ because v , w , $\|\mathbf{x}\|^{2-n} - 1$ when $n > 2$, and $\ln(\|\mathbf{x}\|^{-1})$ when $n = 2$ are continuous on $\overline{B_1(\mathbf{0})} \setminus B_\delta(\mathbf{0})$ and harmonic on $B_1(\mathbf{0}) \setminus \overline{B_\delta(\mathbf{0})}$. Corollary 8.3.3 is used to justify the previous statement.

We have that $f(\mathbf{x}) = 0$ for $\mathbf{x} \in \partial B_1(\mathbf{0})$. Let $m = \min_{\mathbf{x} \in \overline{B_1(\mathbf{0})}} |w(\mathbf{x})|$. Such a minimum exists because w is continuous on the compact set $\overline{B_1(\mathbf{0})}$. By taking δ small enough, it follows from the assumption of the lemma that

$$f(\mathbf{x}) \leq \begin{cases} \underbrace{\left(\frac{v(\mathbf{x})}{\|\mathbf{x}\|^{2-n}} - \epsilon \right)}_{<0} \underbrace{\|\mathbf{x}\|^{2-n}}_{\rightarrow +\infty} - m + \epsilon < 0 & \text{if } n > 2 \\ \underbrace{\left(\frac{v(\mathbf{x})}{\ln(\|\mathbf{x}\|^{-1})} - \epsilon \right)}_{<0} \underbrace{\ln(\|\mathbf{x}\|^{-1})}_{\rightarrow +\infty} - m < 0 & \text{if } n = 2 \end{cases}$$

for all $\mathbf{x} \in B_\delta(\mathbf{0})$. Thus $f(\mathbf{x}) \leq 0$ on $\partial B_\delta(\mathbf{0})$.

It follows from the maximum principle, Theorem 9.2.3, applied to f on $\overline{B_1(\mathbf{0})} \setminus B_\delta(\mathbf{0})$ that $f(\mathbf{x}) \leq 0$ for $\mathbf{x} \in \overline{B_1(\mathbf{0})} \setminus B_\delta(\mathbf{0})$. Since this is true for all δ small enough, we get that $f(\mathbf{x}) \leq 0$ for $\mathbf{x} \in \overline{B_1(\mathbf{0})} \setminus \{\mathbf{0}\}$. Letting ϵ converges to 0, we get that $v(\mathbf{x}) - w(\mathbf{x}) \leq 0$ for $\mathbf{x} \in \overline{B_1(\mathbf{0})} \setminus \{\mathbf{0}\}$.

Proceeding similarly with $-f$ instead of f , we get that $w(\mathbf{x}) - v(\mathbf{x}) \geq 0$ for $\mathbf{x} \in \overline{B_1(\mathbf{0})} \setminus \{\mathbf{0}\}$. Completing the proof that $w(\mathbf{x}) - v(\mathbf{x}) = 0$ for $\mathbf{x} \in \overline{B_1(\mathbf{0})} \setminus \{\mathbf{0}\}$. ■

Proposition 9.9.5

Suppose that $u \in C^2(\Omega')$ is harmonic on Ω' . The following statements are equivalent.

1. u is harmonic at infinity.
2. $u(\mathbf{x}) = O(\|\mathbf{x}\|^{2-n})$ as $\|\mathbf{x}\| \rightarrow \infty$.
3. $|u(\mathbf{x})| \rightarrow 0$ as $\|\mathbf{x}\| \rightarrow \infty$ when $n > 2$, or $|u(\mathbf{x})| = o(\ln(\|\mathbf{x}\|))$ as $\|\mathbf{x}\| \rightarrow \infty$ when $n = 2$.

Proof.

1 $\Rightarrow$ 2) If u is harmonic at infinity, then the Kelvin transform $\tilde{u}$ has a removable singularity at the origin. Thus, there exist $\delta > 0$ such that $B_\delta(\mathbf{0}) \subset T(\Omega')$ and a harmonic function $w : B_\delta(\mathbf{0}) \rightarrow \mathbb{R}$ such that

$$w(\mathbf{x}) = \tilde{u}(\mathbf{x}) = \|\mathbf{x}\|_2^{2-n} u(T(\mathbf{x}))$$

for all $\mathbf{x} \in B_\delta(\mathbf{0}) \setminus \{\mathbf{0}\}$. Let $M = \max_{\mathbf{x} \in \overline{B_{\delta/2}(\mathbf{0})}} |w(\mathbf{x})|$.

We get that

$$u(\mathbf{x}) = \|\mathbf{x}\|_2^{2-n} w(T(\mathbf{x}))$$

for $\mathbf{x} \in T(B_\delta(\mathbf{0})) = \{\mathbf{x} : \|\mathbf{x}\| > 1/\delta\} \subset \Omega'$. Hence, $|u(\mathbf{x})| \leq M\|\mathbf{x}\|^{2-n}$ for $\|\mathbf{x}\| > 2/\delta$.

2 $\Rightarrow$ 3) If $u(\mathbf{x}) = O(\|\mathbf{x}\|^{2-n})$ as $\|\mathbf{x}\| \rightarrow \infty$, then there exist $C, R > 0$ such that $|u(\mathbf{x})| \leq C\|\mathbf{x}\|^{2-n}$ for $\|\mathbf{x}\| > R$. If $n > 0$, then it automatically follows that $|u(\mathbf{x})| \rightarrow 0$ as $\|\mathbf{x}\| \rightarrow \infty$. If $n = 2$, then $|u(\mathbf{x})| \leq C$ for $\|\mathbf{x}\| > R$. Hence,

$$\left| \frac{u(\mathbf{x})}{\ln(\|\mathbf{x}\|)} \right| \leq \frac{C}{\ln(\|\mathbf{x}\|)} \rightarrow 0$$

as $\|\mathbf{x}\| \rightarrow \infty$.

3 $\Rightarrow$ 1) It follows from the definition of the Kelvin transform $\tilde{u}$ of u that

$$\left| \frac{\tilde{u}(\mathbf{x})}{\|\mathbf{x}\|_2^{2-n}} \right| = |u(T(\mathbf{x}))| = |u(\|\mathbf{x}\|^{-2}\mathbf{x})| \rightarrow 0$$

as $\|\mathbf{x}\| \rightarrow 0$ when $n > 2$, and

$$\left| \frac{\tilde{u}(\mathbf{x})}{\ln(\|\mathbf{x}\|^{-1})} \right| = \left| \frac{u(T(\mathbf{x}))}{\ln(\|\mathbf{x}\|^{-1})} \right| = \left| \frac{u(\|\mathbf{x}\|^{-2}\mathbf{x})}{\ln(\|\mathbf{x}\|^{-1})} \right| \rightarrow 0$$

as $\|\mathbf{x}\| \rightarrow 0$ when $n = 2$ by assumption. Hence, according to Lemma 9.9.4, $\tilde{u}$ has a removable singularity at the origin in both cases. ■

To prove the next proposition, we will use some of the results about the homogeneous polynomials of degree k in $\mathbb{R}^n$ presented in Section 9.6.

Proposition 9.9.6

Suppose that $u \in C^2(\Omega')$ is harmonic on $\Omega' \cup \{\infty\}$. Let $\nu(\mathbf{x})$ be the outward unit normal to $\partial B_r(\mathbf{0}) \subset \Omega'$ at $\mathbf{x} \in \partial B_r(\mathbf{0})$. If $n > 2$, then $\left| \frac{\partial u}{\partial \nu}(\mathbf{x}) \right| = O(\|\mathbf{x}\|^{1-n})$ as $\|\mathbf{x}\| \rightarrow \infty$. If $n = 2$, then $\left| \frac{\partial u}{\partial \nu}(\mathbf{x}) \right| = O(\|\mathbf{x}\|^{-2})$ as $\|\mathbf{x}\| \rightarrow \infty$.

Proof.

We may assume that there exists $0 < R < 1$ such that $\{\mathbf{x} : \|\mathbf{x}\| > R\} \subset \Omega'$. If it is not so, choose $q > 0$ such that $\{\mathbf{x} : \|\mathbf{x}\| > q/2\} \subset \Omega'$ and set $W = q^{-1}\Omega'$. The function $w : W \rightarrow \mathbb{R}$ defined by $w(\mathbf{x}) = u(q\mathbf{x})$ for $\mathbf{x} \in W$ is of class C^2 and is harmonic on $W \cup \{\infty\}$. Since $\{\mathbf{x} : \|\mathbf{x}\| > 1/2\} \subset W$, we may choose $0 < R < 1$ such that $\{\mathbf{x} : \|\mathbf{x}\| > R\} \subset W$. It is then enough to prove the proposition for $w : W \rightarrow \mathbb{R}$.

For the rest of this proof, we will refer to the notation introduced in Section 9.6.

We consider the Kelvin function $\tilde{u}$ of u . Since u is harmonic at infinity, we may assume that $\tilde{u}$ is harmonic on $\overline{B_1(\mathbf{0})} \subset T(B_R(\mathbf{0}))$. Let $S = \partial B_1(\mathbf{0})$ and consider the functions $f : S \rightarrow \mathbb{R}$ and $\phi_k \in H_k|_S$ defined by $f(\mathbf{x}) = \tilde{u}(\mathbf{x})$ and $\phi_k(\mathbf{x}) = \left\langle f, p_{\mathbf{x}}^{[k]} \right\rangle$ for all $\mathbf{x} \in S$ respectively.

Recall that $p_{\mathbf{x}}^{[k]}$ is the homogeneous polynomial of degree k defined in Section 9.6. It follows from Corollary 9.6.7 that

$$\tilde{u}(\mathbf{x}) = \sum_{k=0}^{\infty} \|\mathbf{x}\|^k \phi_k(\|\mathbf{x}\|^{-1}\mathbf{x})$$

for $\mathbf{x} \in B_1(\mathbf{0})$. We also have that the series converges absolutely and uniformly on $B_R(\mathbf{0})$ for every $0 \leq R < 1$. We get that

$$u(\mathbf{x}) = \|T(\mathbf{x})\|^{n-2} \tilde{u}(T(\mathbf{x})) = \sum_{k=0}^{\infty} \|\mathbf{x}\|^{2-n-k} \phi_k(\|\mathbf{x}\|^{-1}\mathbf{x})$$

for $\|\mathbf{x}\| > 1$. If $\mathbf{x} = r\mathbf{y}$ with $r > 1$ and $\|\mathbf{y}\| = 1$, we have that

$$u(r\mathbf{y}) = \sum_{k=0}^{\infty} r^{2-n-k} \phi_k(\mathbf{y}) ,$$

where the convergence of the series is uniform for $r > 1/R$ and $\mathbf{y} \in S$ whatever $0 < R < 1$. We may therefore interchange the derivative with respect to r with the summation to get

$$\frac{\partial u}{\partial \nu}(r\mathbf{y}) = \sum_{k=0}^{\infty} (2-n-k) r^{1-n-k} \phi_k(\mathbf{y}) = r^{1-n} \sum_{k=0}^{\infty} (2-n-k) r^{-k} \phi_k(\mathbf{y}) \quad (9.9.4)$$

for $r > 1$.

i) We consider first the case $n > 2$. For $r > 3$ and p large enough such that $(n+p-2) < 2^p$, we have that

$$\left| \sum_{k=p}^{\infty} (2-n-k) r^{-k} \phi_k(\mathbf{y}) \right| \leq \sum_{k=p}^{\infty} (n+k-2) 3^{-k} |\phi_k(\mathbf{y})| \leq \sum_{k=p}^{\infty} \left(\frac{2}{3}\right)^k |\phi_k(\mathbf{y})|$$

for all $\mathbf{y} \in S$. Since $\sum_{k=p}^{\infty} (2/3)^k |\phi_k|$ converges uniformly on S to a continuous function, there

exists a constant $C > 0$ such that $\sum_{k=p}^{\infty} (2/3)^k |\phi_k(\mathbf{y})| < C$ for all $\mathbf{y} \in S$. It then follows from

(9.9.4) that $\left| \frac{\partial u}{\partial \nu}(r\mathbf{y}) \right| = O(r^{1-n})$ for $n > 2$ as $r \rightarrow \infty$.

ii) If $n = 2$, the expression in (9.9.4) becomes

$$\frac{\partial u}{\partial \nu}(r\mathbf{y}) = r^{-2} \sum_{k=1}^{\infty} (-k) r^{-k-1} \phi_k(\mathbf{y}) .$$

For $r > 3$ and p large enough such that $p < 2^{p+1}$, we have that

$$\left| \sum_{k=p}^{\infty} (-k) r^{-k-1} \phi_k(\mathbf{y}) \right| \leq \sum_{k=p}^{\infty} k 3^{-k-1} |\phi_k(\mathbf{y})| \leq \sum_{k=p}^{\infty} \left(\frac{2}{3}\right)^{k+1} |\phi_k(\mathbf{y})|$$

for all $\mathbf{y} \in S$. As before, since $\sum_{k=p}^{\infty} (2/3)^k |\phi_k|$ converges uniformly on S to a continuous function,

there exists a constant $C > 0$ such that $\sum_{k=p}^{\infty} (2/3)^k |\phi_k(\mathbf{y})| < C$ for all $\mathbf{y} \in S$. It then follows

from (9.9.4) that $\left| \frac{\partial u}{\partial \nu}(r\mathbf{y}) \right| = O(r^{-2})$ for $n = 2$ as $r \rightarrow \infty$. ■

Remark 9.9.7

The method used in this chapter to solve (D1), (D2), (N1) and (N2) is called **Layer Potentials**. In Section 12.7, we use the results of the present chapter to prove Theorem 9.5.7. ♠

9.9.1 Uniqueness of solutions**9.9.1.1 Problem D1**

The uniqueness of the solution of (D1) is given by Theorem 9.2.6.

9.9.1.2 Problem D2

To prove the uniqueness of the solution of (D2), suppose that $u \in C(\overline{\Omega'})$ is a solution of (D2) with $u = f = 0$ on $\partial\Omega'$. According to Proposition 9.9.2, The Kelvin transform $\tilde{u}$ of u is harmonic in $T(\Omega')$ and satisfies $\tilde{u} = \tilde{f} = 0$ on $T(\partial\Omega') = \partial T(\Omega')$. Because $T(\Omega')$ is a bounded open subset of $\mathbb{R}^n$, we have that $\tilde{u} = 0$ on $T(\Omega')$ by uniqueness of solution for the problem (D1). This implies that $u = 0$ on Ω' .

9.9.1.3 Problem N1

To prove the uniqueness of the solution of (N1) modulo a function which is constant on each connected component of Ω , suppose that $u \in C_\nu(\Omega)$ is a solution of (N1) with $\frac{\partial u}{\partial \nu^-} = f = 0$ on $\partial\Omega$.

We will need the following famous result.

Theorem 9.9.8 (Tubular Neighbourhood Theorem)

Suppose that $S \subset \mathbb{R}^n$ is a $(n-1)$ -dimensional compact oriented manifold of class C^k with $k > 1$. There exists $\mu > 0$ such that

$$O_\mu = \{\mathbf{x} + t\nu(\mathbf{x}) : \mathbf{x} \in S, |t| < \mu\}$$

is an open neighbourhood of S . As usual, $\nu(\mathbf{x})$ is the outward unit normal to S at $\mathbf{x} \in M$. Moreover, the mapping

$$\begin{aligned} F : S \times]-\mu, \mu[&\rightarrow \mathbb{R}^n \\ (\mathbf{x}, t) &\mapsto \mathbf{x} + t\nu(\mathbf{x}) \end{aligned}$$

is a diffeomorphism of class C^{k-1} onto O_μ .

Proof.

The proof can be found in any good textbook on differential geometry like [1, 37]. ■

Let $\Omega_\epsilon \subset \Omega$ be the set bounded by $\partial\Omega_\epsilon = \{\mathbf{x} + \epsilon\nu(\mathbf{x}) : \mathbf{x} \in \partial\Omega\}$ for $\epsilon < 0$, where $|\epsilon|$ is small enough so that $O_{|\epsilon|} = \{\mathbf{x} + t\nu(\mathbf{x}) : \mathbf{x} \in \partial\Omega, |t| < |\epsilon|\}$ is a tubular neighbourhood of $\partial\Omega$.

From the Green's identity (9.1.1) with $v = u$, we have

$$\int_{\Omega_\epsilon} \|\nabla u\|^2 d\mathbf{x} = \int_{\partial\Omega_\epsilon} u \frac{\partial u}{\partial \nu} dS_{\mathbf{x}} - \int_{\Omega_\epsilon} u \Delta u d\mathbf{x} = \int_{\partial\Omega_\epsilon} u \frac{\partial u}{\partial \nu} dS_{\mathbf{x}} \quad (9.9.5)$$

because $\Delta u = 0$ on Ω . From the Lebesgue's Monotone Convergence Theorem, we may conclude that

$$\int_{\Omega_\epsilon} \|\nabla u\|_2^2 d\mathbf{x} = \int_{\Omega} \chi_{\Omega_\epsilon} \|\nabla u\|_2^2 d\mathbf{x} \rightarrow \int_{\Omega} \|\nabla u\|^2 d\mathbf{x} \quad (9.9.6)$$

as $\epsilon \rightarrow 0^-$, where

$$\chi_{\Omega_\epsilon}(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{x} \in \Omega_\epsilon \\ 0 & \text{if } \mathbf{x} \notin \Omega_\epsilon \end{cases}$$

Moreover, we have that

$$\begin{aligned} \int_{\partial\Omega_\epsilon} u \frac{\partial u}{\partial \nu} dS_{\mathbf{x}} &= \int_{\partial\Omega} u(\mathbf{x} + \epsilon\nu(\mathbf{x})) \nabla u(\mathbf{x} + \epsilon\nu(\mathbf{x})) \cdot \nu(\mathbf{x}) dS_{\mathbf{x}} \\ &\rightarrow \int_{\partial\Omega} u(\mathbf{x}) \frac{\partial u}{\partial \nu^-}(\mathbf{x}) dS_{\mathbf{x}} = 0 \end{aligned} \quad (9.9.7)$$

as $\epsilon \rightarrow 0^-$ because $\nabla u(\mathbf{x} + \epsilon\nu(\mathbf{x})) \cdot \nu(\mathbf{x}) \rightarrow \frac{\partial u}{\partial \nu^-}(\mathbf{x})$ uniformly on $\partial\Omega$ as $\epsilon \rightarrow 0^-$ since $u \in C_\nu(\Omega)$, and $u(\mathbf{x} + \epsilon\nu(\mathbf{x})) \rightarrow u(\mathbf{x})$ uniformly on $\partial\Omega$ as $\epsilon \rightarrow 0^-$ since u is uniformly continuous on the compact set $\bar{\Omega}$. We have also used the fact that the outward unit normal to $\partial\Omega_\epsilon$ at $\mathbf{x} + \epsilon\nu(\mathbf{x})$ is $\nu(\mathbf{x})$ because of our tubular neighbourhood.

It follows from (9.9.5), (9.9.6) and (9.9.7) that $\int_{\Omega} \|\nabla u\|^2 d\mathbf{x} = 0$. Since $u \in C^1(\Omega)$, this implies that $\nabla u = \mathbf{0}$ everywhere on Ω . Thus u is constant on each connected component of Ω .

9.9.1.4 Problem N2

We prove the uniqueness of the solution of (N2) modulo a function which is constant on each connected component of Ω . Moreover, we prove that $u = 0$ on the unbounded component if $n > 2$.

Suppose that $u \in C_\nu(\Omega')$ is a solution of (N2) with $\frac{\partial u}{\partial \nu^+} = f = 0$ on $\partial\Omega$.

Let $\Omega'_\epsilon \subset \Omega'$ be the set bounded by $\partial\Omega'_\epsilon = \{\mathbf{x} + \epsilon\nu(\mathbf{x}) : \mathbf{x} \in \partial\Omega\}$ for $\epsilon > 0$ small enough so that $O_\epsilon = \{\mathbf{x} + t\nu(\mathbf{x}) : \mathbf{x} \in \partial\Omega, |t| < \epsilon\}$ is a tubular neighbourhood of $\partial\Omega = \partial\Omega'$.

Choose $r > 0$ such that $\overline{\mathbb{R}^n \setminus \Omega'_\epsilon} \subset B_r(\mathbf{0})$. From the Green's identity (9.1.1) with $v = u$, we have

$$\int_{B_r(\mathbf{0}) \cap \Omega'_\epsilon} \|\nabla u\|^2 d\mathbf{x} = \int_{\partial B_r(\mathbf{0})} u \frac{\partial u}{\partial \nu} dS_{\mathbf{x}} - \int_{\partial\Omega'_\epsilon} u \frac{\partial u}{\partial \nu} dS_{\mathbf{x}} - \int_{B_r(\mathbf{0}) \cap \Omega'_\epsilon} u \Delta u d\mathbf{x}$$

$$= \int_{\partial B_r(\mathbf{0})} u \frac{\partial u}{\partial \nu} dS_{\mathbf{x}} - \int_{\partial \Omega'_\epsilon} u \frac{\partial u}{\partial \nu} dS_{\mathbf{x}} \quad (9.9.8)$$

because $\Delta u = 0$ in Ω' .

Using the Lebesgue's Monotone Convergence Theorem as we did to prove the uniqueness of the solution for the problem (N1), we have that

$$\int_{B_r(\mathbf{0}) \cap \Omega'_\epsilon} \|\nabla u\|^2 d\mathbf{x} \rightarrow \int_{B_r(\mathbf{0}) \cap \Omega'} \|\nabla u\|^2 d\mathbf{x} \quad (9.9.9)$$

as $\epsilon \rightarrow 0^+$. Moreover, it follows from (9.9.2) that

$$\begin{aligned} \int_{\partial \Omega'_\epsilon} u \frac{\partial u}{\partial \nu} dS_{\mathbf{x}} &= \int_{\partial \Omega} u(\mathbf{x} + \epsilon \nu(\mathbf{x})) \nabla u(\mathbf{x} + \epsilon \nu(\mathbf{x})) \cdot \nu(\mathbf{x}) dS_{\mathbf{x}} \\ &\rightarrow \int_{\partial \Omega} u(\mathbf{x}) \frac{\partial u}{\partial \nu^+}(\mathbf{x}) dS_{\mathbf{x}} = 0 \end{aligned} \quad (9.9.10)$$

as $\epsilon \rightarrow 0^+$ because $\nabla u(\mathbf{x} + \epsilon \nu(\mathbf{x})) \cdot \nu(\mathbf{x}) \rightarrow \frac{\partial u}{\partial \nu^+}(\mathbf{x})$ uniformly on $\partial \Omega$ as $\epsilon \rightarrow 0^+$ since $u \in C_\nu(\Omega')$, and $u(\mathbf{x} + \epsilon \nu(\mathbf{x})) \rightarrow u(\mathbf{x})$ uniformly on $\partial \Omega$ as $\epsilon \rightarrow 0^+$ since u is uniformly continuous on the compact set $\overline{B_r(\mathbf{0}) \cap \Omega'}$. It follows from (9.9.8), (9.9.9) and (9.9.10) that

$$\int_{B_r(\mathbf{0}) \cap \Omega'} \|\nabla u\|^2 d\mathbf{x} = \int_{\partial B_r(\mathbf{0})} u \frac{\partial u}{\partial \nu} dS_{\mathbf{x}}. \quad (9.9.11)$$

n > 2: From Proposition 9.9.5 and Proposition 9.9.6, we have that $|u(\mathbf{x})| = O(\|\mathbf{x}\|^{2-n})$ and $\left| \frac{\partial u}{\partial \nu}(\mathbf{x}) \right| = O(\|\mathbf{x}\|^{1-n})$ as $\|\mathbf{x}\| \rightarrow \infty$. Hence, there exists a constant C such that

$$\left| \int_{\partial B_r(\mathbf{0})} u \frac{\partial u}{\partial \nu} dS_{\mathbf{x}} \right| \leq \int_{\partial B_r(\mathbf{0})} |u| \left| \frac{\partial u}{\partial \nu} \right| dS_{\mathbf{x}} \leq C r^{2-n} r^{1-n} \underbrace{\int_{\partial B_r(\mathbf{0})} dS_{\mathbf{x}}}_{= \omega_n r^{n-1}} = C \omega_n r^{2-n} \rightarrow 0$$

as $r \rightarrow \infty$. It follows from (9.9.11) and the Lebesgue's Monotone Convergence Theorem that

$$\int_{B_r(\mathbf{0}) \cap \Omega'} \|\nabla u\|_2^2 d\mathbf{x} \rightarrow \int_{\Omega'} \|\nabla u\|_2^2 d\mathbf{x} = 0$$

as $r \rightarrow \infty$. Since $u \in C^1(\Omega')$, we find that $\nabla u = \mathbf{0}$ everywhere on Ω' . Thus u is constant on each connected component of Ω' .

Moreover, $u = 0$ on Ω'_0 , the unbounded component of Ω' , because $u(\mathbf{x}) \rightarrow 0$ as $\|\mathbf{x}\| \rightarrow \infty$ according to Proposition 9.9.5. This is possible for a constant function u on Ω'_0 only if $u = 0$ on Ω'_0 .

n = 2: From Proposition 9.9.5 and Proposition 9.9.6, we have that $|u(\mathbf{x})| = o(\ln(\|\mathbf{x}\|))$ and $\left| \frac{\partial u}{\partial \nu}(\mathbf{x}) \right| = O(\|\mathbf{x}\|^{-2})$ as $\|\mathbf{x}\| \rightarrow \infty$. Hence, there exists a constant C such that

$$\left| \int_{\partial B_r(\mathbf{0})} u \frac{\partial u}{\partial \nu} dS_{\mathbf{x}} \right| \leq \int_{\partial B_r(\mathbf{0})} |u| \left| \frac{\partial u}{\partial \nu} \right| dS_{\mathbf{x}} \leq C r^{-2} \ln(r) \underbrace{\int_{\partial B_r(\mathbf{0})} dS_{\mathbf{x}}}_{= 2\pi r} = 2\pi C \frac{\ln(r)}{r} \rightarrow 0$$

as $r \rightarrow \infty$. As before, it follows from (9.9.11) and the Lebesgue Monotone Convergence Theorem that

$$\int_{B_r(\mathbf{0}) \cap \Omega'} \|\nabla u\|_2^2 d\mathbf{x} \rightarrow \int_{\Omega'} \|\nabla u\|^2 d\mathbf{x} = 0$$

as $r \rightarrow \infty$. Since $u \in C^1(\Omega')$, we find that $\nabla u = \mathbf{0}$ everywhere on Ω' . Thus u is constant on each connected component of Ω' .

9.9.2 Necessary condition for the Neumann Problem

In this section, we give necessary conditions for the existence of the solutions of the Interior and Exterior Neumann problem. The first proposition has been proved in the previous section.

Proposition 9.9.9

If u is a solution of (N2) with $f = 0$ on $\partial\Omega'$, the following statements are true.

1. For $n = 2$, u is constant on each connected component Ω'_j , $0 \leq j \leq m'$, of Ω' .
2. For $n > 2$, u is constant on each bounded connected component Ω'_j , $1 \leq j \leq m'$, of Ω' and $u = 0$ on the unbounded connected component Ω'_0 .

Proposition 9.9.10

1. If u is a solution of (N1), then $\int_{\partial\Omega_i} f = 0$ for $1 \leq i \leq m$.
2. If u is a solution of (N2), then $\int_{\partial\Omega'_i} f = 0$ for $1 \leq i \leq m'$. Moreover, $\int_{\partial\Omega'_0} f = 0$ for $n = 2$.

Proof.

1) For $1 \leq i \leq m$, let $\Omega_{i,\epsilon} \subset \Omega_i$ be the set bounded by $\partial\Omega_{i,\epsilon} = \{\mathbf{x} + \epsilon\nu(\mathbf{x}) : \mathbf{x} \in \partial\Omega_i\}$ for $\epsilon < 0$ and $|\epsilon|$ small enough so that $O_{|\epsilon|} = \{\mathbf{x} + t\nu(\mathbf{x}) : \mathbf{x} \in \partial\Omega_i, |t| < |\epsilon|\}$ is a tubular neighbourhood of $\partial\Omega_i$.

From the Green's identity (9.1.1) with $v = 1$, we have

$$\int_{\partial\Omega_{i,\epsilon}} \frac{\partial u}{\partial \nu} dS_{\mathbf{x}} = \int_{\Omega_{i,\epsilon}} \Delta u d\mathbf{x} = 0 \quad (9.9.12)$$

because $\Delta u = 0$ on Ω_i . However, we have that

$$\int_{\partial\Omega_{i,\epsilon}} \frac{\partial u}{\partial \nu} dS_{\mathbf{x}} = \int_{\partial\Omega_i} \nabla u(\mathbf{x} + \epsilon\nu(\mathbf{x})) \cdot \nu(\mathbf{x}) dS_{\mathbf{x}} \rightarrow \int_{\partial\Omega_i} \frac{\partial u}{\partial \nu^-} dS_{\mathbf{x}} = \int_{\partial\Omega_i} f dS_{\mathbf{x}}$$

as $\epsilon \rightarrow 0^-$ because $\nabla u(\mathbf{x} + \epsilon\nu(\mathbf{x})) \cdot \nu(\mathbf{x}) \rightarrow \frac{\partial u}{\partial \nu^-}(\mathbf{x})$ uniformly on $\partial\Omega_i$ as $\epsilon \rightarrow 0^-$ since $u \in C_\nu(\Omega)$.

It follows from (9.9.12) that $\int_{\partial\Omega_i} f dS_{\mathbf{x}} = 0$.

2) To prove that $\int_{\partial\Omega'_i} f dS_{\mathbf{x}} = 0$ for $1 \leq i \leq m'$ is identical to the proof given in (i) if Ω_i is replaced by Ω'_i and $\epsilon > 0$ is used.

If $n = 2$, let $\Omega'_{0,\epsilon} \subset \Omega'_0$ be the set bounded by $\partial\Omega'_{0,\epsilon} = \{\mathbf{x} + \epsilon\nu(\mathbf{x}) : \mathbf{x} \in \partial\Omega'_0\}$ for $\epsilon > 0$ small enough so that $O_\epsilon = \{\mathbf{x} + t\nu(\mathbf{x}) : \mathbf{x} \in \partial\Omega'_0, |t| < \epsilon\}$ is a tubular neighbourhood of $\partial\Omega = \partial\Omega'_0$.

Choose $r > 0$ large enough such that $\partial\Omega'_{0,\epsilon} \subset B_r(\mathbf{0})$.

From the Green's identity (9.1.1) with $v = 1$, we have

$$\int_{\partial\Omega'_{0,\epsilon}} \frac{\partial u}{\partial \nu} dS_{\mathbf{x}} = \int_{\partial B_r(\mathbf{0})} \frac{\partial u}{\partial \nu} dS_{\mathbf{x}} - \int_{B_r(\mathbf{0}) \cap \Omega'_{0,\epsilon}} \Delta u d\mathbf{x} = \int_{\partial B_r(\mathbf{0})} \frac{\partial u}{\partial \nu} dS_{\mathbf{x}} \quad (9.9.13)$$

because $\Delta u = 0$ on Ω'_0 . We have that

$$\int_{\partial\Omega'_{0,\epsilon}} \frac{\partial u}{\partial \nu} dS_{\mathbf{x}} = \int_{\partial\Omega'_0} \nabla u(\mathbf{x} + \epsilon\nu(\mathbf{x})) \cdot \nu(\mathbf{x}) dS_{\mathbf{x}} \rightarrow \int_{\partial\Omega'_0} \frac{\partial u}{\partial \nu^+} dS_{\mathbf{x}} = \int_{\partial\Omega'_0} f dS_{\mathbf{x}} \quad (9.9.14)$$

as $\epsilon \rightarrow 0^+$ because $\nabla u(\mathbf{x} + \epsilon\nu(\mathbf{x})) \cdot \nu(\mathbf{x}) \rightarrow \frac{\partial u}{\partial \nu^+}(\mathbf{x})$ uniformly on $\partial\Omega'_0$ as $\epsilon \rightarrow 0^+$ since $u \in C_\nu(\Omega')$.

Moreover, from Proposition 9.9.6, we have that $\left| \frac{\partial u}{\partial \nu}(\mathbf{x}) \right| = O(\|\mathbf{x}\|^{-2})$ as $\|\mathbf{x}\| \rightarrow \infty$. Therefore, there exists a constant C such that

$$\left| \int_{\partial B_r(\mathbf{0})} \frac{\partial u}{\partial \nu} dS_{\mathbf{x}} \right| \leq \int_{\partial B_r(\mathbf{0})} \left| \frac{\partial u}{\partial \nu} \right| dS_{\mathbf{x}} \leq Cr^{-2} \int_{\partial B_r(\mathbf{0})} dS_{\mathbf{x}} = 2\pi Cr^{-1} \rightarrow 0 \quad (9.9.15)$$

as $r \rightarrow \infty$. It follows from (9.9.13), (9.9.14) and (9.9.15), that $\int_{\partial\Omega'_0} f dS_{\mathbf{x}} = 0$. ■

9.9.3 Note on the Integration on Manifolds

Before diving into the study of the integral operators which are of interest to us, we need to review integration on manifolds. Good references on the subject are [1, 37].

Definition 9.9.11

Let Ω be an open subset of $\mathbb{R}^n$ and S be a subset of Ω . Let $\{V_i\}_{i=1}^\infty$ be a collection of open subsets of Ω such that $S \subset \bigcup_{i=1}^\infty V_i$. A **partition of unity** for S subordinated to $\{V_i\}_{i=1}^\infty$ is a set of functions $\{\phi_i\}_{i=1}^\infty$ such that:

1. The ϕ_i are C^∞ functions on Ω .
2. $0 \leq \phi_i(\mathbf{x}) \leq 1$ for all i and all $\mathbf{x} \in \Omega$.
3. $\text{supp } \phi_i \subset V_i$ for all i .
4. For any compact set $K \subset S$, we have that $\text{supp } \phi_i \cap K = \emptyset$ for all but a finite number of indices.

$$5. \sum_{i=1}^{\infty} \phi_i(\mathbf{x}) = 1 \text{ for all } \mathbf{x} \in S.$$

The definition of partition of unity subordinated to a finite collection of open subsets $\{V_i\}_{i=1}^N$ is identical. Note that the support of ϕ_i may not be compact if the open set V_i is not bounded.

The summation in the previous definition is finite for every $\mathbf{x} \in S$ because, given any compact set $K \subset S$, $K \cap \text{supp } \phi_i = \emptyset$ for all but a finite number of indices.

Proposition 9.9.12

Let Ω be an open subset of $\mathbb{R}^n$ and S be a subset of Ω . If $\{V_i\}_{i=1}^{\infty}$ is a collection of open sets in Ω such that $S \subset \bigcup_{i=1}^{\infty} V_i$. Then, there exists a partition of unity for S subordinate to $\{V_i\}_{i=1}^{\infty}$.

Proof.

This proposition is proved in any good functional analysis or differential geometry textbook. For instance, see [2]. ■

There is another definition of partition of unity that will be useful in Section 11.3 and that we introduce right now.

Definition 9.9.13

Let Ω be an open subset of $\mathbb{R}^n$ and $\{V_i\}_{i=1}^{\infty}$ be a collection of open subsets of $\mathbb{R}^n$ such that $\Omega = \bigcup_{i=1}^{\infty} V_i$. A **partition of unity** for Ω subordinated to $\{V_i\}_{i=1}^{\infty}$ is a set of functions $\{\phi_i\}_{i=1}^{\infty}$ such that:

1. The ϕ_i are C^∞ functions on Ω .
2. $0 \leq \phi_i(\mathbf{x}) \leq 1$ for all i and all $\mathbf{x} \in \Omega$.
3. $\text{supp } \phi_i \subset V_i$ for all i .
4. For any compact set $K \subset \Omega$, there exists an open set $W \supset K$ such that $\text{supp } \phi_i \cap W = \emptyset$ for all but a finite number of indices.
5. $\sum_{i=1}^{\infty} \phi_i(\mathbf{x}) = 1$ for all $\mathbf{x} \in \Omega$.

Note that the support of ϕ_i may not be compact if the open set V_i is not bounded.

As for the previous definition of partition of unity, the summation in the previous definition is finite for every $\mathbf{x} \in \Omega$ because, given any compact set $K \subset \Omega$, there exist an open set $W \supset K$ such that $W \cap \text{supp } \phi_i = \emptyset$ for all but a finite number of indices.

Note that Definition 9.9.13 implies Definition 9.9.11.

Proposition 9.9.14

Let Ω be an open subset of $\mathbb{R}^n$ and $\{V_i\}_{i=1}^\infty$ is be collection of open sets in Ω such that $S \subset \bigcup_{i=1}^\infty V_i$. Then, there exists a partition of unity for Ω subordinate to $\{V_i\}_{i=1}^\infty$.

Proof.

See [31]. ■

Suppose that S is an oriented differentiable manifold and $f : S \rightarrow \mathbb{R}$ is a continuous function with compact support. Let $\mathcal{B} = \{(U_j, \psi_j)\}_{j \in J}$ be an atlas for S , where $J \subset \mathbb{N}$ is an index set. Moreover, let $\{\phi_j\}_{j \in J}$ be a partition of unity for S subordinated to $\mathcal{B}$; namely, $\{\phi_j\}_{j \in J}$ is a partition of unity for S subordinated to $\{V_j\}_{j \in J}$, where $U_j = V_j \cap S$ and V_j is an open subset of $\mathbb{R}^n$. Note that we assume that the topology on S is the induced topology from $\mathbb{R}^n$.

The integral $\int_S f(\mathbf{x}) dS$ is defined by

$$\begin{aligned} \int_S f(\mathbf{x}) dS &= \int_S \sum_{j \in J} \phi_j(\mathbf{x}) f(\mathbf{x}) dS = \sum_{j \in J} \int_{U_j} \phi_j(\mathbf{x}) f(\mathbf{x}) dS \\ &= \sum_{j \in J} \int_{\psi_j(U_j)} \phi_j(\psi_j^{-1}(\mathbf{w})) f(\psi_j^{-1}(\mathbf{w})) |\det D\psi_j^{-1}(\mathbf{w})| d\mathbf{w} . \end{aligned}$$

The integral is independent of the chosen atlas and partition of unity. Moreover, since f has a compact support, the summation is finite. The sum is over all j such that $U_j \cap \text{supp } f \neq \emptyset$.

The manifold that we consider in this chapter is $S = \partial\Omega$ where $\Omega \subset \mathbb{R}^n$ is a bounded open set. We also assume S is of class at least C^2 . The orientation on S is determined by the outward unit normal $\nu(\mathbf{x})$ to S at $\mathbf{x} \in S$.

To simplify the computations later, we select a particular atlas of S as it follows. At each point $\mathbf{x} \in M$, we select a chart $(U_{\mathbf{x}}, \psi_{\mathbf{x}})$ such that:

- $U_{\mathbf{x}}$ is a connected set containing $\mathbf{x}$ and $U_{\mathbf{x}} = S \cap V_{\mathbf{x}}$ for some bounded open set $V_{\mathbf{x}} \subset \mathbb{R}^n$. This is possible because S is the boundary of an open set.
- $\psi_{\mathbf{x}}$ is the projection of $U_{\mathbf{x}}$ onto $W_{\mathbf{x}} \subset T_{\mathbf{x}}$ along the vector $\nu(\mathbf{x})$, where $T_{\mathbf{x}}$ is the tangent space to $U_{\mathbf{x}}$ at the point $\mathbf{x}$. We have that $\psi_{\mathbf{x}}^{-1} : W_{\mathbf{x}} \rightarrow \mathbb{R}^n$ is of class C^2 . Using translation and rotation, we may consider that $U_{\mathbf{x}}$ is the graph of a function of class C^2 from $\mathbb{R}^{n-1}$ to $\mathbb{R}$ (Figure 9.8).
- By shrinking $V_{\mathbf{x}}$ and $U_{\mathbf{x}}$ if necessary, we may assume that $\psi_{\mathbf{x}}^{-1}$ is of class C^2 on the compact set $\overline{W_{\mathbf{x}}}$.
- It is also convenient to assume that the diameter of $V_{\mathbf{x}}$, and so of $W_{\mathbf{x}}$, is less than 1.

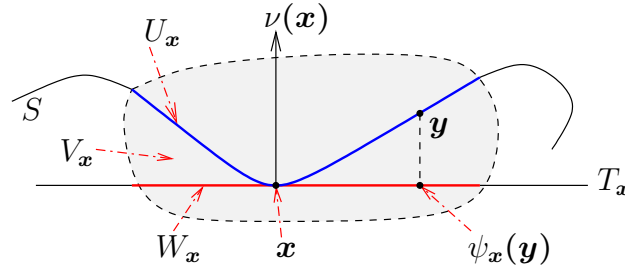


Figure 9.8: Chart (U_x, ψ_x) of the manifold S in Section 9.9.3.

Since S is compact, there exists a finite cover $\{V_{x_j}\}_{j \in J}$ of S and so a finite atlas $\mathcal{B} = \{(U_{x_j}, \psi_{x_j})\}_{j \in J}$ of S .

Let $Q_{x_j} = \max_{w \in \overline{W_{x_j}}} |\det D \psi_{x_j}^{-1}(w)|$. We have that $Q_{x_j} < \infty$ because we assume that $\phi_{x_j}^{-1}$ is of class C^2 on the compact set $\overline{W_{x_j}}$. If $Q = \max_{j \in J} Q_{x_j}$, then

$$|\det D \psi_{x_j}^{-1}(w)| < Q \quad (9.9.16)$$

for all $w \in W_{x_j}$ and all $j \in J$.

Given $x \in S$ and $\eta < 1$, let $S_{x,\eta} = S \cap B_\eta(x)$. In the next sections, we will need an upper bound for the integrals

$$\int_{S_{x,\eta}} \|x - y\|^{-\alpha} dS_y \quad \text{and} \quad \int_{S_{x,\eta}} |\ln(\|x - y\|)| dS_y$$

which depends only on η but not on $x \in S$, where $0 \leq \alpha < n - 1$.

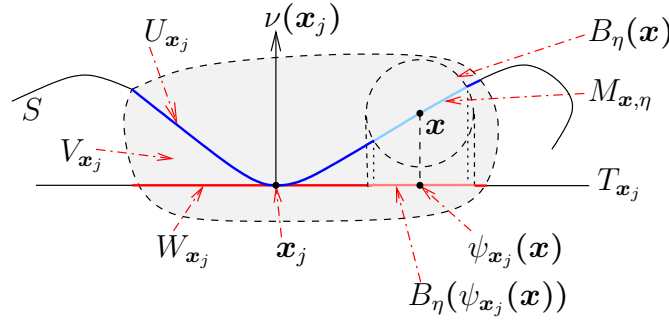
If $\eta/2$ is smaller than the Lebesgue number associated to the open covering $\{V_{x_j}\}_{j \in J}$ of S ⁹, then $\overline{B_\eta(x)} \subset V_{x_j}$ for some $j \in J$. Thus $\overline{S_{x,\eta}} \subset U_{x_j}$ (Figure 9.9).

n > 2 Since $\|x - y\| \geq \|\psi_{x_j}(x) - \psi_{x_j}(y)\|$ for all $x, y \in U_{x_j}$ and $\psi_{x_j}(S_{x,\eta}) \subset B_\eta(\psi_{x_j}(x))$, we get that

$$\begin{aligned} \int_{S_{x,\eta}} \|x - y\|^{-\alpha} dS_y &= \int_{\psi_{x_j}(S_{x,\eta})} \|x - \psi_{x_j}^{-1}(w)\|^{-\alpha} |\det D \psi_{x_j}^{-1}(w)| dw \\ &\leq \int_{\psi_{x_j}(S_{x,\eta})} \|\psi_{x_j}(x) - w\|^{-\alpha} |\det D \psi_{x_j}^{-1}(w)| dw \leq Q \int_{B_\eta(\psi_{x_j}(x))} \|\psi_{x_j}(x) - w\|^{-\alpha} dw \\ &= Q \int_{\partial B_1(0)} \int_0^\eta r^{-\alpha} r^{n-2} dr dS = \frac{Q \omega_{n-1} \eta^{n-\alpha-1}}{n - \alpha - 1}, \end{aligned}$$

where we have use the change of variables $w = \psi_{x_j}(x) + rz$ for $0 \leq r < \eta$ and $z \in \partial B_1(0) \subset \mathbb{R}^{n-1}$.

⁹[23] is a good reference on the subject of the Lebesgue number.

Figure 9.9: Region $S_{x,\eta}$ of the manifold S in Section 9.9.3.

Similarly, we get

$$\begin{aligned} \int_{S_{x,\eta}} |\ln(\|\mathbf{x} - \mathbf{y}\|)| \, dS_{\mathbf{y}} &= \int_{\psi_{x_j}(S_{x,\eta})} |\ln(\|\mathbf{x} - \psi_{x_j}^{-1}(\mathbf{w})\|)| |\det D\psi_{x_j}^{-1}(\mathbf{w})| \, d\mathbf{w} \\ &\leq \int_{\psi_{x_j}(S_{x,\eta})} |\ln(\|\psi_{x_j}(\mathbf{x}) - \mathbf{w}\|)| |\det D\psi_{x_j}^{-1}(\mathbf{w})| \, d\mathbf{w} \leq Q \int_{B_\eta(\psi_{x_j}(\mathbf{x}))} |\ln(\|\psi_{x_j}(\mathbf{x}) - \mathbf{w}\|)| \, d\mathbf{w} \\ &= Q \int_{\partial B_1(\mathbf{0})} \int_0^\eta |\ln(r)| r^{n-2} \, dr \, dS \leq Q\omega_{n-1} \left(\frac{|\ln(\eta)|\eta^{n-1}}{n-1} + \frac{\eta^{n-1}}{(n-1)^2} \right), \end{aligned}$$

where again we have use the change of variables $\mathbf{w} = \psi_{x_j}(\mathbf{x}) + r\mathbf{z}$ for $0 \leq r < \eta < 1$ and $\mathbf{z} \in \partial B_1(\mathbf{0}) \subset \mathbb{R}^{n-1}$.

n = 2) The computations are very similar to those of the case $n > 2$ with a single difference as we will see. Since $n = 2$, we have that $\alpha = 0$. Hence, we get

$$\int_{S_{x,\eta}} \|\mathbf{x} - \mathbf{y}\|^{-\alpha} \, dS_{\mathbf{y}} = \int_{S_{x,\eta}} dS_{\mathbf{y}} \leq Q \int_{B_\eta(\psi_{x_j}(\mathbf{x})) \subset \mathbb{R}} d\mathbf{w} = Q \int_{\psi_{x_j}(\mathbf{x})-\eta}^{\psi_{x_j}(\mathbf{x})+\eta} d\mathbf{w} = 2Q\eta.$$

Similarly, we get

$$\begin{aligned} \int_{S_{x,\eta}} |\ln(\|\mathbf{x} - \mathbf{y}\|)| \, dS_{\mathbf{y}} &\leq Q \int_{B_\eta(\psi_{x_j}(\mathbf{x})) \subset \mathbb{R}} |\ln(\|\psi_{x_j}(\mathbf{x}) - \mathbf{w}\|)| \, d\mathbf{w} \\ &= 2Q \int_{\psi_{x_j}(\mathbf{x})}^{\psi_{x_j}(\mathbf{x})+\eta} |\ln(\|\psi_{x_j}(\mathbf{x}) - \mathbf{w}\|)| \, d\mathbf{w} = 2Q \int_0^\eta |\ln(r)| \, dr \leq 2Q (\eta |\ln(\eta)| + \eta). \end{aligned}$$

In conclusion,

$$\int_{S_{x,\eta}} \|\mathbf{x} - \mathbf{y}\|^{-\alpha} \, dS_{\mathbf{y}} \leq \frac{Q\omega_{n-1}\eta^{n-\alpha-1}}{n-\alpha-1} \quad (9.9.17)$$

and

$$\int_{S_{x,\eta}} |\ln(\|\mathbf{x} - \mathbf{y}\|)| \, dS_{\mathbf{y}} \leq Q\omega_{n-1} \left(\frac{|\ln(\eta)|\eta^{n-1}}{n-1} + \frac{\eta^{n-1}}{(n-1)^2} \right) \quad (9.9.18)$$

for $0 \leq \alpha < n-1$ and $n \geq 2$ if we set $\omega_1 = 2$.

9.9.4 Integral Operators

In this section, we set $S = \partial\Omega$. Recall that S is a compact C^2 -manifold of dimension $n - 1$ in $\mathbb{R}^n$.

Definition 9.9.15

Let $K : S \times S \rightarrow \mathbb{R}$ be a measurable function. K is a **kernel of order** α with $0 < \alpha < n - 1$ if

$$K(\mathbf{x}, \mathbf{y}) = A(\mathbf{x}, \mathbf{y}) \|\mathbf{x} - \mathbf{y}\|^{-\alpha} ,$$

where $A : S \times S \rightarrow \mathbb{R}$ is a bounded measurable function. K is a **kernel of order** 0 if

$$K(\mathbf{x}, \mathbf{y}) = A(\mathbf{x}, \mathbf{y}) \ln(\|\mathbf{x} - \mathbf{y}\|) + B(\mathbf{x}, \mathbf{y}) ,$$

where $A, B : S \times S \rightarrow \mathbb{R}$ are bounded measurable functions. Moreover, if

$$K : \{(\mathbf{x}, \mathbf{y}) \in S \times S : \mathbf{x} \neq \mathbf{y}\} \rightarrow \mathbb{R}$$

is also continuous, we say that K is a **continuous kernel of order** α

Proposition 9.9.16

Suppose that $K : S \times S \rightarrow \mathbb{R}$ is a kernel of order α with $0 \leq \alpha < n - 1$, and $\text{supp } K \subset \{(\mathbf{x}, \mathbf{y}) \in S \times S : \|\mathbf{x} - \mathbf{y}\| < \epsilon\}$ for $\epsilon > 0$. Then

$$(T_K f)(\mathbf{x}) = \int_S K(\mathbf{x}, \mathbf{y}) f(\mathbf{y}) dS_{\mathbf{y}} \quad , \quad \mathbf{x} \in S ,$$

defines a bounded linear operator $T_K : L^p(S) \rightarrow L^p(S)$ for $1 \leq p \leq \infty$. If $\alpha > 0$, then there exists a constant C such that

$$\|T_K f\|_p \leq C \epsilon^{n-\alpha-1} \|A\|_{\infty} \|f\|_p \quad (9.9.19)$$

for $f \in L^p(S)$, where A is given in Definition 9.9.15. If $\alpha = 0$, then there exists a constant C such that

$$\|T_K f\|_p \leq C \epsilon^{n-1} (\|A\|_{\infty} (1 + |\ln(\epsilon)|) + \|B\|_{\infty}) \|f\|_p \quad (9.9.20)$$

for $f \in L^p(S)$, where A and B are given in Definition 9.9.15.

Proof.

That $T_K : L^p(S) \rightarrow L^p(S)$ defines a bounded linear operator follows from (9.9.19) and (9.9.20).

Let $\mathcal{B} = \{(U_{\mathbf{x}_j}, \psi_{\mathbf{x}_j})\}_{j \in J}$ be the finite atlas for S defined in Section 9.9.3, and let $\{\phi_{\mathbf{x}_j}\}_{j \in J}$ be a partition of unity for S subordinated to $\mathcal{B}$.

$\alpha > 0$) We have that

$$\begin{aligned}
\int_S |K(\mathbf{x}, \mathbf{y})| dS_{\mathbf{y}} &= \int_{S \cap B_\epsilon(\mathbf{x})} |A(\mathbf{x}, \mathbf{y})| \|\mathbf{x} - \mathbf{y}\|^{-\alpha} dS_{\mathbf{y}} \leq \|A\|_\infty \int_{S \cap B_\epsilon(\mathbf{x})} \|\mathbf{x} - \mathbf{y}\|^{-\alpha} dS_{\mathbf{y}} \\
&= \|A\|_\infty \sum_{j \in J} \int_{U_{\mathbf{x}_j} \cap B_\epsilon(\mathbf{x})} \phi_{\mathbf{x}_j}(\mathbf{y}) \|\mathbf{x} - \mathbf{y}\|^{-\alpha} dS_{\mathbf{y}} \leq \|A\|_\infty \sum_{j \in J} \int_{U_{\mathbf{x}_j} \cap B_\epsilon(\mathbf{x})} \|\mathbf{x} - \mathbf{y}\|^{-\alpha} dS_{\mathbf{y}} \\
&= \|A\|_\infty \sum_{j \in J} \int_{\psi_{\mathbf{x}_j}^{-1}(U_{\mathbf{x}_j} \cap B_\epsilon(\mathbf{x}))} \|\mathbf{x} - \psi_{\mathbf{x}_j}^{-1}(\mathbf{w})\|^{-\alpha} |\det D \psi_{\mathbf{x}_j}^{-1}(\mathbf{w})| d\mathbf{w} \\
&\leq \|A\|_\infty Q \sum_{j \in J} \int_{\psi_{\mathbf{x}_j}^{-1}(U_{\mathbf{x}_j} \cap B_\epsilon(\mathbf{x}))} \|\mathbf{x} - \psi_{\mathbf{x}_j}^{-1}(\mathbf{w})\|^{-\alpha} d\mathbf{w} \\
&\leq \|A\|_\infty Q \sum_{j \in J} \int_{\psi_{\mathbf{x}_j}^{-1}(U_{\mathbf{x}_j} \cap B_\epsilon(\mathbf{x}))} \|\psi_{\mathbf{x}_j}(\mathbf{x}) - \mathbf{w}\|^{-\alpha} d\mathbf{w} \\
&\leq \|A\|_\infty Q \sum_{j \in J} \int_{\partial B_1(\mathbf{0})} \int_0^\epsilon r^{-\alpha} r^{n-2} dr dS = \frac{\|A\|_\infty Q \omega_{n-1} |J| \epsilon^{n-\alpha-1}}{n - \alpha - 1} < \infty
\end{aligned}$$

for all $\mathbf{x} \in S$, where $|J|$ is the number of elements in J . Some of the inequalities in the relation above will benefit from some explanations. The second inequality is a consequence of $0 \leq \phi_{\mathbf{x}_j}(\mathbf{y}) \leq 1$ for all $\mathbf{y}$ and all functions in the partition of unity. The third inequality comes from (9.9.16). The fourth inequality comes $\|\mathbf{x} - \mathbf{y}\| \geq \|\psi_{\mathbf{x}_j}(\mathbf{x}) - \psi_{\mathbf{x}_j}(\mathbf{y})\|$ for all $\mathbf{x}, \mathbf{y} \in U_{\mathbf{x}_j}$. The last inequality is true because $\psi_{\mathbf{x}_j}^{-1}(U_{\mathbf{x}_j} \cap B_\epsilon(\mathbf{x})) \subset B_\epsilon(\psi_{\mathbf{x}_j}^{-1}(\mathbf{x})) \subset \mathbb{R}^{n-1}$. We have used the change of variables $\mathbf{w} = \psi_{\mathbf{x}_j}(\mathbf{x}) + r\mathbf{z}$ with $0 \leq r < \epsilon$ and $\mathbf{z} \in \partial B_1(\mathbf{0}) \subset \mathbb{R}^{n-1}$.

Similarly,

$$\int_S |K(\mathbf{x}, \mathbf{y})| dS_{\mathbf{x}} \leq \frac{\|A\|_\infty Q \omega_{n-1} |J| \epsilon^{n-\alpha-1}}{n - \alpha - 1}$$

for all $\mathbf{y} \in S$.

Thus, (9.9.19) with $C = \frac{Q \omega_{n-1} |J|}{n - \alpha - 1}$ follows from the generalized Young's inequality, Theorem 6.5.1.

$\alpha = 0$) Proceeding as we did for $\alpha > 0$, we find that

$$\begin{aligned}
\int_S |K(\mathbf{x}, \mathbf{y})| dS_{\mathbf{y}} &= \int_{S \cap B_\epsilon(\mathbf{x})} |A(\mathbf{x}, \mathbf{y}) \ln(\|\mathbf{x} - \mathbf{y}\|) + B(\mathbf{x}, \mathbf{y})| dS_{\mathbf{y}} \\
&\leq \|A\|_\infty \int_{S \cap B_\epsilon(\mathbf{x})} |\ln(\|\mathbf{x} - \mathbf{y}\|)| dS_{\mathbf{y}} + \|B\|_\infty \int_{S \cap B_\epsilon(\mathbf{x})} dS_{\mathbf{y}} \\
&= \|A\|_\infty \sum_{j \in J} \int_{U_{\mathbf{x}_j} \cap B_\epsilon(\mathbf{x})} \phi_{\mathbf{x}_j}(\mathbf{y}) |\ln(\|\mathbf{x} - \mathbf{y}\|)| dS_{\mathbf{y}} + \|B\|_\infty \sum_{j \in J} \int_{U_{\mathbf{x}_j} \cap B_\epsilon(\mathbf{x})} \phi_{\mathbf{x}_j}(\mathbf{y}) dS_{\mathbf{y}} \\
&\leq \|A\|_\infty \sum_{j \in J} \int_{U_{\mathbf{x}_j} \cap B_\epsilon(\mathbf{x})} |\ln(\|\mathbf{x} - \mathbf{y}\|)| dS_{\mathbf{y}} + \|B\|_\infty \sum_{j \in J} \int_{U_{\mathbf{x}_j} \cap B_\epsilon(\mathbf{x})} dS_{\mathbf{y}} \\
&= \|A\|_\infty \sum_{j \in J} \int_{\psi_{\mathbf{x}_j}^{-1}(U_{\mathbf{x}_j} \cap B_\epsilon(\mathbf{x}))} |\ln(\|\mathbf{x} - \psi_{\mathbf{x}_j}^{-1}(\mathbf{w})\|)| |\det D \psi_{\mathbf{x}_j}^{-1}(\mathbf{w})| d\mathbf{w} \\
&\quad + \|B\|_\infty \sum_{j \in J} \int_{\psi_{\mathbf{x}_j}^{-1}(U_{\mathbf{x}_j} \cap B_\epsilon(\mathbf{x}))} |\det D \psi_{\mathbf{x}_j}^{-1}(\mathbf{w})| d\mathbf{w} \\
&\leq \|A\|_\infty Q \sum_{j \in J} \int_{\psi_{\mathbf{x}_j}^{-1}(U_{\mathbf{x}_j} \cap B_\epsilon(\mathbf{x}))} |\ln(\|\mathbf{x} - \psi_{\mathbf{x}_j}^{-1}(\mathbf{w})\|)| d\mathbf{w} + \|B\|_\infty Q \sum_{j \in J} \int_{\psi_{\mathbf{x}_j}^{-1}(U_{\mathbf{x}_j} \cap B_\epsilon(\mathbf{x}))} d\mathbf{w}
\end{aligned}$$

$$\begin{aligned}
&\leq \|A\|_\infty Q \sum_{j \in J} \int_{\psi_{\mathbf{x}_j}^{-1}(U_{\mathbf{x}_j} \cap B_\epsilon(\mathbf{x}))} |\ln(\|\psi_{\mathbf{x}_j}(\mathbf{x}) - \mathbf{w}\|)| \, d\mathbf{w} + \|B\|_\infty Q \sum_{j \in J} \int_{\psi_{\mathbf{x}_j}^{-1}(U_{\mathbf{x}_j} \cap B_\epsilon(\mathbf{x}))} d\mathbf{w} \\
&\leq \|A\|_\infty Q \sum_{j \in J} \int_{\partial B_1(\mathbf{0})} \int_0^\epsilon |\ln(r)| r^{n-2} \, dr \, dS + \|B\|_\infty Q \sum_{j \in J} \int_{\partial B_1(\mathbf{0})} \int_0^\epsilon r^{n-2} \, dr \, dS \\
&\leq \frac{\|A\|_\infty Q \omega_{n-1} |J| (|\ln(\epsilon)| + 1) \epsilon^{n-1}}{n-1} + \frac{\|B\|_\infty Q \omega_{n-1} |J| \epsilon^{n-1}}{n-1} < \infty
\end{aligned}$$

for all $\mathbf{x} \in S$. To obtain the fourth inequality, we have the fact that the diameter of $V_{\mathbf{x}_j}$ is less than 1 to get $\ln(\|\mathbf{x} - \psi_{\mathbf{x}_j}^{-1}(\mathbf{w})\|) \leq \ln(\|\psi_{\mathbf{x}_j}(\mathbf{x}) - \mathbf{w}\|)$ for $\|\psi_{\mathbf{x}_j}(\mathbf{x}) - \mathbf{w}\| \leq \|\mathbf{x} - \psi_{\mathbf{x}_j}^{-1}(\mathbf{w})\| \leq 1$.

Similarly,

$$\int_S |K(\mathbf{x}, \mathbf{y})| \, dS_{\mathbf{x}} \leq \frac{\|A\|_\infty Q \omega_{n-1} |J| (|\ln(\epsilon)| + 1) \epsilon^{n-1}}{n-1} + \frac{\|B\|_\infty Q \omega_{n-1} |J| \epsilon^{n-1}}{n-1} < \infty$$

for all $\mathbf{y} \in S$.

Thus, (9.9.20) with $C = Q \omega_{n-1} |J| / (n-1)$ follows from the generalized Young's inequality, Theorem 6.5.1.

Note that

$$\int_0^\epsilon |\ln(r)| r^{n-2} \, dr = \frac{|\ln(\epsilon)| \epsilon^{n-1}}{n-1} + \frac{\epsilon^{n-1}}{(n-1)^2} \leq \frac{(|\ln(\epsilon)| + 1) \epsilon^{n-1}}{n-1}$$

for $0 < \epsilon < 1$, and

$$\int_0^\epsilon |\ln(r)| r^{n-2} \, dr = \frac{|\ln(\epsilon)| \epsilon^{n-1}}{n-1} - \frac{\epsilon^{n-1}}{(n-1)^2} + \frac{2}{(n-1)^2} \leq \frac{|\ln(\epsilon)| \epsilon^{n-1}}{n-1} + \frac{\epsilon^{n-1}}{(n-1)^2} \leq \frac{(|\ln(\epsilon)| + 1) \epsilon^{n-1}}{n-1}$$

for $\epsilon > 1$. ■

The proof of the previous proposition contains in some way the following results.

Proposition 9.9.17

If K is a kernel of order $0 \leq \alpha < n-1$, then there exists a constant C such that $\int_S |K(\mathbf{x}, \mathbf{y})| \, dS_{\mathbf{y}} < C$ for all $\mathbf{x} \in S$. In particular, the function defined by $\mathbf{y} \mapsto K(\mathbf{x}, \mathbf{y})$ for $\mathbf{y} \in S$ is in $L^1(S)$ for all $\mathbf{x} \in S$.

Proof.

We may assume that $0 < \alpha < n$. If K is a kernel of order 0, we have

$$\begin{aligned}
|K(\mathbf{x}, \mathbf{y})| &\leq \|A\|_\infty |\ln(\|\mathbf{x} - \mathbf{y}\|)| + \|B\|_\infty \\
&= (\|A\|_\infty \|\mathbf{x} - \mathbf{y}\|^\alpha |\ln(\|\mathbf{x} - \mathbf{y}\|)| + \|B\|_\infty \|\mathbf{x} - \mathbf{y}\|^\alpha) \|\mathbf{x} - \mathbf{y}\|^{-\alpha} \leq D \|\mathbf{x} - \mathbf{y}\|^{-\alpha}
\end{aligned}$$

for $\mathbf{x}, \mathbf{y} \in S$ and $\alpha > 0$, where

$$D = \max_{\mathbf{x}, \mathbf{y} \in S} (\|A\|_\infty \|\mathbf{x} - \mathbf{y}\|^\alpha |\ln(\|\mathbf{x} - \mathbf{y}\|)| + \|B\|_\infty \|\mathbf{x} - \mathbf{y}\|^\alpha) < \infty$$

because S is compact and $\lim_{r \rightarrow 0} r^\alpha \ln(r) = 0$ whatever $\alpha > 0$. Thus K is a kernel of order $\alpha > 0$.

Choose $\epsilon > 0$ such that $\text{supp } K \subset \{(\mathbf{x}, \mathbf{y}) \in S \times S : \|\mathbf{x} - \mathbf{y}\| < \epsilon\}$. Note that we may not have $\epsilon < 1$ but this is not necessary for $\alpha > 0$.

We can proceed exactly as in the case $\alpha > 0$ of the proof of Proposition 9.9.16 to show that

$$\int_S |K(\mathbf{x}, \mathbf{y})| \, dS_{\mathbf{y}} \leq C \equiv \frac{\|A\|_\infty Q \omega_{n-1} |J| \epsilon^{n-\alpha-1}}{n - \alpha - 1}$$

for all $\mathbf{x} \in S$. ■

Proposition 9.9.18

If $K : S \times S \rightarrow \mathbb{R}$ is a kernel of order α with $0 \leq \alpha < n - 1$, then $T_K : L^2(S) \rightarrow L^2(S)$ is compact.

Proof.

Given $0 < \epsilon < 1$, defined $K_\epsilon^{[o]} : S \times S \rightarrow \mathbb{R}$ by

$$K_\epsilon^{[o]}(\mathbf{x}, \mathbf{y}) = \begin{cases} K(\mathbf{x}, \mathbf{y}) & \text{if } \|\mathbf{x} - \mathbf{y}\| \geq \epsilon \\ 0 & \text{otherwise} \end{cases}$$

and $K_\epsilon^{[i]} : S \times S \rightarrow \mathbb{R}$ by $K_\epsilon^{[i]} = K - K_\epsilon^{[o]}$. Choose $r > 0$ such that $S \subset \overline{B_r(\mathbf{0})}$.

i) $K_\epsilon^{[o]} \in L^2(S \times S)$ because $K_\epsilon^{[o]}$ is bounded on the compact set $S \times S$. More precisely, for $\alpha > 0$, we have $|K_\epsilon^{[o]}(\mathbf{x}, \mathbf{y})| = |A(\mathbf{x}, \mathbf{y})| \|\mathbf{x} - \mathbf{y}\|^{-\alpha} \leq \|A\|_\infty \epsilon^{-\alpha}$ for $\|\mathbf{x} - \mathbf{y}\| \geq \epsilon$ and $|K_\epsilon^{[o]}(\mathbf{x}, \mathbf{y})| = 0$ for $\|\mathbf{x} - \mathbf{y}\| < \epsilon$. For $\alpha = 0$, we have

$$|K_\epsilon^{[o]}(\mathbf{x}, \mathbf{y})| \leq |A(\mathbf{x}, \mathbf{y}) \ln(\|\mathbf{x} - \mathbf{y}\|) + B(\mathbf{x}, \mathbf{y})| \leq \|A\|_\infty \max\{|\ln(\epsilon)|, |\ln(2r)|\} + \|B\|_\infty$$

for $\|\mathbf{x} - \mathbf{y}\| \geq \epsilon$ and $|K_\epsilon^{[o]}(\mathbf{x}, \mathbf{y})| = 0$ for $\|\mathbf{x} - \mathbf{y}\| < \epsilon$.

Thus $K_\epsilon^{[o]}$ is an Hilbert-Schmidt kernel according to Theorem 4.4.46. This implies in particular that $T_{K_\epsilon^{[o]}} : L^2(S) \rightarrow L^2(S)$ is compact.

ii) Since $K_\epsilon^{[i]}$ is a kernel of order α , it follows from Proposition 9.9.16 that $\|T_K - T_{K_\epsilon^{[o]}}\|_2 = \|T_{K_\epsilon^{[i]}}\|_2 \leq C \|A\|_\infty \epsilon^{n-\alpha-1}$ for a constant C when $\alpha > 0$, and $\|T_K - T_{K_\epsilon^{[o]}}\|_2 = \|T_{K_\epsilon^{[i]}}\|_2 \leq C (\|A\|_\infty (1 + |\ln(\epsilon)|) + \|B\|_\infty) \epsilon^n$ for a constant C when $\alpha = 0$.

Hence $T_{K_\epsilon^{[o]}} \rightarrow T$ as $\epsilon \rightarrow 0$. Since the operators $T_{K_\epsilon^{[o]}}$ are compact, it follows from Proposition 4.4.21 that T_K is compact. ■

Proposition 9.9.19

If $K : S \times S \rightarrow \mathbb{R}$ is a continuous kernel of order α with $0 \leq \alpha < n - 1$ and $f : S \rightarrow \mathbb{R}$ is a bounded function, then $T_K f \in C(S)$.

Proof.

As we did at the beginning of the proof of Proposition 9.9.17, we may assume that $0 < \alpha < n$; namely, that K is a kernel of order $\alpha > 0$.

Choose $\mathbf{x} \in S$ and $\epsilon > 0$. We prove that $T_K f$ is continuous at $\mathbf{x}$. For $r > 0$ and $\mathbf{w} \in S$, let

$$S_{\mathbf{w},r} = \{\mathbf{z} \in S : \|\mathbf{w} - \mathbf{z}\| < r\} = S \cap B_r(\mathbf{w}) .$$

Given $\delta > 0$ and $\mathbf{y} \in S_{\mathbf{x},\delta}$, we have that

$$\begin{aligned} |(T_K f)(\mathbf{x}) - (T_K f)(\mathbf{y})| &= \left| \int_S (K(\mathbf{x}, \mathbf{z}) - K(\mathbf{y}, \mathbf{z})) f(\mathbf{z}) \, dS_{\mathbf{z}} \right| \\ &\leq \int_S |K(\mathbf{x}, \mathbf{z}) - K(\mathbf{y}, \mathbf{z})| |f(\mathbf{z})| \, dS_{\mathbf{z}} \leq I_1(\mathbf{x}, \mathbf{y}) + I_2(\mathbf{x}, \mathbf{y}) , \end{aligned}$$

where

$$I_1(\mathbf{x}, \mathbf{y}) = \int_{S_{\mathbf{x},2\delta}} (|K(\mathbf{x}, \mathbf{z})| + |K(\mathbf{y}, \mathbf{z})|) |f(\mathbf{z})| \, dS_{\mathbf{z}}$$

and

$$I_2(\mathbf{x}, \mathbf{y}) = \int_{S \setminus S_{\mathbf{x},2\delta}} |K(\mathbf{x}, \mathbf{z}) - K(\mathbf{y}, \mathbf{z})| |f(\mathbf{z})| \, dS_{\mathbf{z}} .$$

i) We have that

$$I_1(\mathbf{x}, \mathbf{y}) \leq \|A\|_{\infty} \|f\|_{\infty} \int_{S_{2\delta}} (\|\mathbf{x} - \mathbf{z}\|^{-\alpha} + \|\mathbf{y} - \mathbf{z}\|^{-\alpha}) \, dS_{\mathbf{z}} .$$

If we assume that $3\delta/2$ is smaller than the Lebesgue number associated to the atlas that we have introduced in Section 9.9.3, then we get from (9.9.17) that

$$\int_{S_{\mathbf{x},2\delta}} \|\mathbf{x} - \mathbf{z}\|^{-\alpha} \, dS_{\mathbf{z}} \leq \frac{Q\omega_{n-1}(2\delta)^{n-\alpha-1}}{n-\alpha-1} .$$

Moreover, since $S_{\mathbf{x},2\delta} \subset S_{\mathbf{y},3\delta}$, we get that

$$\int_{S_{\mathbf{x},2\delta}} \|\mathbf{y} - \mathbf{z}\|^{-\alpha} \, dS_{\mathbf{z}} \leq \int_{S_{\mathbf{y},3\delta}} \|\mathbf{y} - \mathbf{z}\|^{-\alpha} \, dS_{\mathbf{z}} \leq \frac{Q\omega_{n-1}(3\delta)^{n-\alpha-1}}{n-\alpha-1} .$$

Thus

$$I_1(\mathbf{x}, \mathbf{y}) \leq \|A\|_{\infty} \|f\|_{\infty} Q\omega_{n-1} \frac{2^{n-\alpha} + 3^{n-\alpha}}{n-\alpha-1} \delta^{n-\alpha-1} \rightarrow 0 \quad \text{as } \delta \rightarrow 0 .$$

Choose δ smaller if needed to get $I_1(\mathbf{x}, \mathbf{y}) < \epsilon/2$ for $\mathbf{y} \in S_{\mathbf{x},\delta}$.

ii) K is continuous on the compact set $\overline{S_{\mathbf{x},\delta}} \times (S \setminus S_{\mathbf{x},2\delta})$ because $(\mathbf{z}, \mathbf{z}) \notin \overline{S_{\mathbf{x},\delta}} \times (S \setminus S_{\mathbf{x},2\delta})$ for all $\mathbf{z} \in S$. Hence K is uniformly continuous on $\overline{S_{\mathbf{x},\delta}} \times (S \setminus S_{\mathbf{x},2\delta})$. Choose a positive number $\delta_0 < \delta$ such that

$$|K(\mathbf{x}, \mathbf{z}) - K(\mathbf{y}, \mathbf{z})| < \frac{\epsilon}{2\|f\|_{\infty} \int_{S \setminus S_{2\delta}} dS}$$

for $\mathbf{y} \in S_{\mathbf{x}, \delta_0}$ and $\mathbf{z} \in S \setminus S_{\mathbf{x}, 2\delta}$. Then,

$$I_2(\mathbf{x}, \mathbf{y}) = \int_{S \setminus S_{\mathbf{x}, 2\delta}} |K(\mathbf{x}, \mathbf{z}) - K(\mathbf{y}, \mathbf{z})| |f(\mathbf{z})| \, dS_{\mathbf{z}} \leq \|f\|_{\infty} \left(\frac{\epsilon}{2\|f\|_{\infty} \int_{S \setminus S_{2\delta}} dS} \right) \int_{S \setminus S_{2\delta}} dS < \frac{\epsilon}{2}$$

for $\mathbf{y} \in S_{\mathbf{x}, \delta_0}$.

iii) Hence, $|(T_K f)(\mathbf{x}) - (T_K f)(\mathbf{y})| < \epsilon$ for $\mathbf{y} \in S_{\mathbf{x}, \delta_0}$. This proves that $T_K f$ is continuous at an arbitrary point $\mathbf{x} \in S$. ■

Proposition 9.9.20

Let $K : S \times S \rightarrow \mathbb{R}$ be a continuous kernel of order α with $0 \leq \alpha < n - 1$. Given $\lambda \in \mathbb{R}$, if $f \in L^2(S)$ and $f + \lambda T_K f \in C(S)$, then $f \in C(S)$.

Proof.

Given $0 < \epsilon < 1$, choose $\psi \in \mathcal{D}(\mathbb{R})$ such that $\psi(x) = 1$ for $x \in]-\epsilon/2, \epsilon/2[$, $\text{supp } \psi \subset]-\epsilon, \epsilon[$ and $0 \leq \psi(x) \leq 1$ for all $x \in \mathbb{R}$. Let $\phi(\mathbf{x}, \mathbf{y}) = \psi(\|\mathbf{x} - \mathbf{y}\|)$ for $\mathbf{x}, \mathbf{y} \in S$. Moreover, let $K_1 = \phi K$ and $K_2 = (1 - \phi)K$.

i) We prove that $T_{K_2} f \in C(S)$. Choose $\mathbf{x} \in S$. From Schwarz inequality, we have

$$\begin{aligned} |(T_{K_2} f)(\mathbf{x}) - (T_{K_2} f)(\mathbf{y})| &\leq \int_S |K_2(\mathbf{x}, \mathbf{z}) - K_2(\mathbf{y}, \mathbf{z})| |f(\mathbf{z})| \, dS_{\mathbf{z}} \\ &\leq \|f\|_2 \left(\int_S |K_2(\mathbf{x}, \mathbf{z}) - K_2(\mathbf{y}, \mathbf{z})|^2 \, dS_{\mathbf{z}} \right)^{1/2} \end{aligned}$$

for $\mathbf{y} \in S$. Moreover, since $K_2 = 0$ in an neighbourhood of the diagonal $\{(\mathbf{z}, \mathbf{z}) : \mathbf{z} \in S\}$, we have the K_2 is continuous on $S \times S$.

Let $h_{\mathbf{y}}(\mathbf{z}) = |K_2(\mathbf{x}, \mathbf{z}) - K_2(\mathbf{y}, \mathbf{z})|^2$ for $\mathbf{z} \in S$, and define $h : S \rightarrow \mathbb{R}$ by $h(\mathbf{z}) = 4\|K_2\|_{\infty}^2$ for all $\mathbf{z} \in S$. We have that $h \in L(S)$ because S is compact.

Since $h_{\mathbf{y}}(\mathbf{z}) \rightarrow 0$ as $\mathbf{y} \rightarrow \mathbf{x}$ for every $\mathbf{z} \in S$, and $0 \leq h_{\mathbf{y}}(\mathbf{z}) \leq h(\mathbf{z})$ for all $\mathbf{z}, \mathbf{y} \in S$, it follows from the Lebesgue Dominated Convergence Theorem that

$$\begin{aligned} |(T_{K_2} f)(\mathbf{x}) - (T_{K_2} f)(\mathbf{y})| &\leq \|f\|_2 \left(\int_S |K_2(\mathbf{x}, \mathbf{z}) - K_2(\mathbf{y}, \mathbf{z})|^2 \, dS_{\mathbf{z}} \right)^{1/2} \\ &= \|f\|_2 \left(\int_S h_{\mathbf{y}}(\mathbf{z}) \, dS_{\mathbf{z}} \right)^{1/2} \rightarrow 0 \quad \text{as } \mathbf{y} \rightarrow \mathbf{x}. \end{aligned}$$

Thus, $T_{K_2} f$ is continuous at the arbitrary point $\mathbf{x} \in S$.

ii) We prove that $f \in C(S)$. Let $g = (f + \lambda T_K f) - \lambda T_{K_2} f$. We have that $g \in C(S)$ by hypothesis and (i). Moreover, $g = f + \lambda T_{K_1} f$ by definition of K_1 and K_2 . It follows from Proposition 9.9.16 that

$$\|\lambda T_{K_1}\|_p \leq \begin{cases} \lambda C \epsilon^{n-\alpha-1} \|A\|_{\infty} & \text{if } \alpha > 0 \\ \lambda C \epsilon^{n-1} (\|A\|_{\infty} (1 + |\ln(\epsilon)|) + \|B\|_{\infty}) & \text{if } \alpha = 0 \end{cases}$$

for $p = 2$ and $p = \infty$, By taking ϵ smaller if needed, we then have that $\|\lambda T_{K_1}\|_p < 1$. It follows from the Banach Lemma that $\text{Id} + \lambda T_{K_1} : L^2(S) \rightarrow L^2(S)$ is invertible. More precisely, $(\text{Id} + \lambda T_{K_1})^{-1} = \sum_{j=0}^{\infty} (-\lambda T_{K_1})^j$. Hence $f = \sum_{j=0}^{\infty} (-\lambda T_{K_1})^j g$, where the convergence is in $L^2(S)$.

Since g is continuous on the compact set S , g is bounded. It follows from Proposition 9.9.19 that $(-\lambda T_{K_1})^j g \in C(S)$ for all j . Moreover, the series $\sum_{j=0}^{\infty} \|\lambda T_{K_1}\|_{\infty}^j$ converges because $\|\lambda T_{K_1}\|_{\infty} < 1$. Hence, given $\delta > 0$, you can choose N large enough such that

$$\left\| \sum_{j=n}^m (-\lambda T_{K_1})^j g \right\|_{\infty} \leq \sum_{j=n}^m \|\lambda T_{K_1}\|_{\infty}^j \|g\|_{\infty} < \delta$$

for $m \geq n > N$. Proving that $\left\{ \sum_{j=0}^n (-\lambda T_{K_1})^j g \right\}_{n=0}^{\infty}$ is a Cauchy sequence in $C(S)$ with respect to the uniform norm. Thus, the limit f of the series is in $C(S)$. ■

Remark 9.9.21

The results in this section are also valid if S is the closure of a bounded and open subset of $\mathbb{R}^{n-1}$. The proofs are almost identical. They are even simpler because the interior of the sets of the form $S_{\mathbf{x}, \epsilon} = \{\mathbf{y} \in S : \|\mathbf{y} - \mathbf{x}\| < \epsilon\}$ are open subsets of $\mathbb{R}^{n-1}$. So, there is no need to refer to an atlas for S and to use the relations (9.9.17) and (9.9.17). Moreover, there is no restriction associated to the Lebesgue number of an atlas. ♠

9.9.5 Double Layer Potential

As in the previous section, we set $S = \partial\Omega$. In addition to being a compact C^2 -manifold of dimension $n - 1$ in $\mathbb{R}^n$ with the topology on S induced from the topology on $\mathbb{R}^n$, S is a measurable space.

Definition 9.9.22

The **double layer potential** with moment $\phi \in C(S)$ is

$$u(\mathbf{x}) = \int_S \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) \phi(\mathbf{y}) \, dS_{\mathbf{y}} \quad , \quad \mathbf{x} \in \mathbb{R}^n \setminus S \quad ,$$

where $\frac{\partial}{\partial \nu_{\mathbf{y}}}$ denotes the directional derivative of $\mathbf{y} \mapsto N(\mathbf{x}, \mathbf{y})$ in the direction of the outward unit normal $\nu(\mathbf{y})$ to S at $\mathbf{y} \in S$ and N is the fundamental solution for the Laplace operator defined in (9.4.1) and Theorem 9.3.1.

We have that

$$\frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) = \nabla_{\mathbf{y}} N(\mathbf{x}, \mathbf{y}) \cdot \nu(\mathbf{y}) = \frac{-1}{\omega_n \|\mathbf{x} - \mathbf{y}\|^n} (\mathbf{x} - \mathbf{y}) \cdot \nu(\mathbf{y})$$

for $\mathbf{x} \neq \mathbf{y}$. Hence,

$$\Delta_{\mathbf{x}} \left(\frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) \right) = \nabla_{\mathbf{y}} (\Delta_{\mathbf{x}} N(\mathbf{x}, \mathbf{y})) \cdot \nu(\mathbf{y}) = 0$$

for $\mathbf{x} \neq \mathbf{y}$ because $N : \mathbb{R}^n \setminus \{\mathbf{0}\} \rightarrow \mathbb{R}$ is harmonic. Thus $\mathbf{x} \mapsto \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y})$ is harmonic on $\mathbb{R}^n \setminus \{\mathbf{y}\}$ for all $\mathbf{y} \in \mathbb{R}^n$. Moreover, since $(\mathbf{x}, \mathbf{y}) \mapsto D_{\mathbf{x}}^{\alpha} \left(\frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) \right)$ is continuous on $(\mathbb{R}^n \setminus S) \times S$ for all multi-index α , we may interchange the integral with respect to $\mathbf{y}$ in the definition of u with the operator $\Delta_{\mathbf{x}}$ to show that u is harmonic on $\mathbb{R}^n \setminus S$.

Since $\frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) = O(\|\mathbf{x}\|^{1-n})$ as $\|\mathbf{x}\| \rightarrow \infty$ uniformly for $\mathbf{y}$ on the compact set S ¹⁰, we have that $u(\mathbf{x}) = O(\|\mathbf{x}\|^{1-n})$ as $\|\mathbf{x}\| \rightarrow \infty$. It follows from Proposition 9.9.5 that u is harmonic at infinity.

Let

$$K(\mathbf{x}, \mathbf{y}) = \begin{cases} \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) & \text{if } \mathbf{x}, \mathbf{y} \in S \text{ and } \mathbf{x} \neq \mathbf{y} \\ \text{anything} & \text{if } \mathbf{x} = \mathbf{y} \in S \end{cases} \quad (9.9.21)$$

This function will be useful later to study the double layer potential on S . Before studying the double layer potential, we need some results about K and $\frac{\partial N}{\partial \nu_{\mathbf{y}}}$.

Lemma 9.9.23

There exists $C > 0$ such that $|(\mathbf{x} - \mathbf{y}) \cdot \nu(\mathbf{x})| \leq C \|\mathbf{x} - \mathbf{y}\|^2$ for all $\mathbf{x}, \mathbf{y} \in S$.

Proof.

i) Let $\mathcal{B} = \{(U_{\mathbf{x}_j}, \psi_{\mathbf{x}_j})\}_{j=1}^J$ be the atlas for S described in Section 9.9.3. Choose $\epsilon/2$ smaller or equal to the Lebesgue number associated to the open cover $\{V_{\mathbf{x}_j}\}_{j=1}^J$ of S . Thus, for each $\mathbf{x} \in S$, there exists an index j such that $S_{\mathbf{x}, \epsilon} = B_{\epsilon}(\mathbf{x}) \cap S \subset U_{\mathbf{x}_j}$.

Choose C_0 such that $C_0 \epsilon > 1$. Since $\nu(\mathbf{x})$ is of norm one for all $\mathbf{x}$, we get from Schwarz inequality that

$$|(\mathbf{x} - \mathbf{y}) \cdot \nu(\mathbf{x})| \leq \|\mathbf{x} - \mathbf{y}\| \leq C_0 \epsilon \|\mathbf{x} - \mathbf{y}\| \leq C_0 \|\mathbf{x} - \mathbf{y}\|^2$$

for all $\mathbf{x}, \mathbf{y} \in S$ such that $\|\mathbf{x} - \mathbf{y}\| \geq \epsilon$. Thus, we only have to consider $\|\mathbf{x} - \mathbf{y}\| < \epsilon$.

ii) Given $\mathbf{x} \in S$. We have that $S_{\mathbf{x}, \epsilon} \subset U_{\mathbf{x}_j}$ for some $j \in J$. Since we assume that $\psi_{\mathbf{x}_j}^{-1}$ is of class C^2 on the compact set $\overline{W_j}$, we have that $C_j = \max_{\mathbf{w} \in \overline{W_j}} \|D^2 \psi_j^{-1}(\mathbf{w})\| < \infty$ ¹¹.

¹⁰Let $M = \max_{\mathbf{y} \in S} \|\mathbf{y}\|$. For $\|\mathbf{x}\| \geq M$, we have that $\|\mathbf{x} - \mathbf{y}\| \geq \|\mathbf{x}\| - \|\mathbf{y}\| \geq \|\mathbf{x}\| - M$ for all $\mathbf{y} \in S$. Thus, for all $\mathbf{y} \in S$, we have that $\|\mathbf{x} - \mathbf{y}\|^{1-n} \leq (\|\mathbf{x}\| - M)^{1-n} < C \|\mathbf{x}\|^{1-n}$ for a constant C and $\|\mathbf{x}\| > M$.

¹¹For a bounded bilinear mapping $B : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^m$, $\|B\| = \max_{\|\mathbf{x}\|=\|\mathbf{y}\|=1} \|B(\mathbf{x}, \mathbf{y})\|$. Hence, $\|B(\mathbf{x}, \mathbf{y})\| \leq \|B\| \|\mathbf{x}\| \|\mathbf{y}\|$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

Let $\check{\mathbf{x}} = \psi_{\mathbf{x}_j}(\mathbf{x})$ and, given $\mathbf{y} \in S_{\mathbf{x}, \epsilon}$, let $\check{\mathbf{y}} = \psi_{\mathbf{x}_j}(\mathbf{y})$. Using the Taylor expansion of $\psi_{\mathbf{x}_j}^{-1}$ about $\check{\mathbf{x}}$, we get

$$|(\mathbf{y} - \mathbf{x}) \cdot \nu(\mathbf{x})| = |(D\psi_{\mathbf{x}_j}^{-1}(\check{\mathbf{x}})(\check{\mathbf{y}} - \check{\mathbf{x}}) + D^2\psi_{\mathbf{x}_j}^{-1}(\mathbf{w})(\check{\mathbf{y}} - \check{\mathbf{x}}, \check{\mathbf{y}} - \check{\mathbf{x}})) \cdot \nu(\mathbf{x})|$$

for some $\mathbf{w} \in W_{\mathbf{x}_j}$. However,

$$\begin{aligned} & (D\psi_{\mathbf{x}_j}^{-1}(\check{\mathbf{x}})(\check{\mathbf{y}} - \check{\mathbf{x}})) \cdot \nu(\mathbf{x}) \\ &= \left(\frac{\partial \psi_{\mathbf{x}_j}^{-1}}{\partial w_1}(\check{\mathbf{x}}) \cdot \nu(\mathbf{x}) \quad \frac{\partial \psi_{\mathbf{x}_j}^{-1}}{\partial w_2}(\check{\mathbf{x}}) \cdot \nu(\mathbf{x}) \quad \dots \quad \frac{\partial \psi_{\mathbf{x}_j}^{-1}}{\partial w_{n-1}}(\check{\mathbf{x}}) \cdot \nu(\mathbf{x}) \right) \cdot (\check{\mathbf{y}} - \check{\mathbf{x}}) = 0 \end{aligned}$$

because $\frac{\partial \psi_{\mathbf{x}_j}^{-1}}{\partial w_k}(\check{\mathbf{x}})$ is in the tangent space of S at $\mathbf{x}$ and so orthogonal to $\nu(\mathbf{x})$ for all k . Thus

$$|(\mathbf{y} - \mathbf{x}) \cdot \nu(\mathbf{x})| = |(D^2\psi_{\mathbf{x}_j}^{-1}(\mathbf{w})(\check{\mathbf{y}} - \check{\mathbf{x}}, \check{\mathbf{y}} - \check{\mathbf{x}})) \cdot \nu(\mathbf{x})| \leq C_j \|\check{\mathbf{y}} - \check{\mathbf{x}}\|^2 \leq C_j \|\mathbf{y} - \mathbf{x}\|^2$$

because $\|\nu(\mathbf{x})\| = 1$, where we have used the fact that $\|\check{\mathbf{y}} - \check{\mathbf{x}}\| = \|\psi_{\mathbf{x}_j}(\mathbf{y}) - \psi_{\mathbf{x}_j}(\mathbf{x})\| \leq \|\mathbf{y} - \mathbf{x}\|$ since $\psi_{\mathbf{x}_j}$ is the projection of $U_{\mathbf{x}_j}$ along the vector $\nu(\mathbf{x}_j)$ onto $W_{\mathbf{x}_j}$.

iii) We get the conclusion of the lemma with $C = \max_{0 \leq j \leq J} C_j$. ■

Lemma 9.9.24

$K : S \times S \rightarrow \mathbb{R}$ defined in (9.9.21) is a continuous kernel of order $n - 2$.

Proof.

We have that $K(\mathbf{x}, \mathbf{y}) = A(\mathbf{x}, \mathbf{y}) \|\mathbf{x} - \mathbf{y}\|^{2-n}$, where

$$A(\mathbf{x}, \mathbf{y}) = \frac{-1}{\omega_n \|\mathbf{x} - \mathbf{y}\|^2} (\mathbf{x} - \mathbf{y}) \cdot \nu(\mathbf{y}) .$$

From Lemma 9.9.23, there exists a constant C such that $|A(\mathbf{x}, \mathbf{y})| \leq C/\omega_n$ for all $\mathbf{x}, \mathbf{y} \in S$ with $\mathbf{x} \neq \mathbf{y}$. So we may expand A to a bounded and measurable function on $S \times S$. Note that $\{(\mathbf{x}, \mathbf{x}) : \mathbf{x} \in S\}$ is a set of measure zero. This shows that K is a kernel of order $n - 2$.

By definition, K is continuous at $\{(\mathbf{x}, \mathbf{y}) \in S \times S : \mathbf{x} \neq \mathbf{y}\}$. ■

Lemma 9.9.25

We have that

$$\int_S K(\mathbf{x}, \mathbf{y}) dS_{\mathbf{y}} = \frac{1}{2} \quad , \quad \mathbf{x} \in S \quad , \quad (9.9.22)$$

and

$$\int_S \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) dS_{\mathbf{y}} = \begin{cases} 1 & \text{if } \mathbf{x} \in \Omega \\ 0 & \text{if } \mathbf{x} \in \Omega' \end{cases} \quad (9.9.23)$$

Proof.

i) We first prove (9.9.23). Given $\mathbf{x} \in \Omega'$, we have that $w(\mathbf{y}) = N(\mathbf{x}, \mathbf{y})$ for $\mathbf{y} \in \overline{\Omega}$ is harmonic on Ω and $w \in C^2(\Omega) \cap C^1(\overline{\Omega})$. It follows from the Green's Identity (9.1.1) with $v = 1$ and $u = w$ that

$$\int_S \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) dS_{\mathbf{y}} = \int_{\Omega} \Delta_{\mathbf{y}} N(\mathbf{x}, \mathbf{y}) d\mathbf{y} = 0$$

for $\mathbf{x} \in \Omega'$.

Given $\mathbf{x} \in \Omega$, choose $\epsilon > 0$ small enough to have $\overline{B_{\epsilon}(\mathbf{x})} \subset \Omega$. Let $\Omega_{\epsilon} = \Omega \setminus \overline{B_{\epsilon}(\mathbf{x})}$. We have that $w(\mathbf{y}) = N(\mathbf{x}, \mathbf{y})$ for $\mathbf{y} \in \overline{\Omega_{\epsilon}}$ is harmonic on Ω_{ϵ} and $w \in C^2(\Omega_{\epsilon}) \cap C^1(\overline{\Omega_{\epsilon}})$. It follows from the Green's Identity (9.1.1) with $v = 1$ and $u = w$ that

$$\int_S \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) dS_{\mathbf{y}} - \int_{\partial B_{\epsilon}(\mathbf{x})} \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) dS_{\mathbf{y}} = \int_{\Omega_{\epsilon}} \Delta_{\mathbf{y}} N(\mathbf{x}, \mathbf{y}) d\mathbf{y} = 0.$$

However,

$$\int_{\partial B_{\epsilon}(\mathbf{x})} \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) dS_{\mathbf{y}} = \frac{1}{\omega_n \epsilon^{n-1}} \int_{\|\mathbf{y}\|_2 = \epsilon} dS_{\mathbf{y}} = 1$$

because $\mathbf{x} - \mathbf{y} = \epsilon \nu(\mathbf{y})$ and $\|\mathbf{x} - \mathbf{y}\| = \epsilon$. Thus, $\int_S \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) dS_{\mathbf{y}} = 1$ for $\mathbf{x} \in \Omega$.

ii) We now prove (9.9.22). Given $\mathbf{x} \in S$, let $A_{\epsilon} = S \setminus \overline{B_{\epsilon}(\mathbf{x})}$, $B_{\epsilon} = \Omega \setminus \overline{B_{\epsilon}(\mathbf{x})}$, $C_{\epsilon} = \Omega \cap \partial B_{\epsilon}(\mathbf{x})$ and $D_{\epsilon} = \{\mathbf{y} \in \partial B_{\epsilon}(\mathbf{x}) : (\mathbf{y} - \mathbf{x}) \cdot \nu(\mathbf{x}) < 0\}$ (Figure 9.10).

Since the function defined by $w(\mathbf{y}) = N(\mathbf{x}, \mathbf{y})$ for $\mathbf{y} \in \overline{B_{\epsilon}}$ is harmonic in B_{ϵ} and $w \in C^2(B_{\epsilon}) \cap C^2(\overline{B_{\epsilon}})$, we may use the Green's Identity (9.1.1) with $v = 1$ and $u = w$ to get

$$\int_{A_{\epsilon}} K(\mathbf{x}, \mathbf{y}) dS_{\mathbf{y}} - \int_{C_{\epsilon}} \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) dS_{\mathbf{y}} = \int_{B_{\epsilon}} \Delta_{\mathbf{y}} N(\mathbf{x}, \mathbf{y}) d\mathbf{y} = 0.$$

Since $\mathbf{y} \rightarrow K(\mathbf{x}, \mathbf{y})$ defines a function in $L^1(S)$ according to Proposition 9.9.17, we may use the Lebesgue Dominate Convergence theorem to conclude that

$$\int_S K(\mathbf{x}, \mathbf{y}) dS_{\mathbf{y}} = \lim_{\epsilon \rightarrow 0} \int_{A_{\epsilon}} K(\mathbf{x}, \mathbf{y}) dS_{\mathbf{y}}.$$

Thus

$$\int_S K(\mathbf{x}, \mathbf{y}) dS_{\mathbf{y}} = \lim_{\epsilon \rightarrow 0} \int_{C_{\epsilon}} \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) dS_{\mathbf{y}}. \quad (9.9.24)$$

We have that

$$\begin{aligned} \int_{C_{\epsilon}} \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) dS_{\mathbf{y}} &= \frac{1}{\omega_n \epsilon^{n-1}} \int_{C_{\epsilon}} dS_{\mathbf{y}} \\ &= \frac{1}{\omega_n \epsilon^{n-1}} \left(\int_{D_{\epsilon}} dS_{\mathbf{y}} - \int_{D_{\epsilon} \setminus C_{\epsilon}} dS_{\mathbf{y}} + \int_{C_{\epsilon} \setminus D_{\epsilon}} dS_{\mathbf{y}} \right). \end{aligned} \quad (9.9.25)$$

Moreover,

$$\frac{1}{\omega_n \epsilon^{n-1}} \int_{D_{\epsilon}} dS_{\mathbf{y}} = \frac{1}{2\omega_n \epsilon^{n-1}} \int_{\partial B_{\epsilon}(\mathbf{x})} dS_{\mathbf{y}} = \frac{1}{2}. \quad (9.9.26)$$

Let $M_\epsilon = \max_{\mathbf{y} \in S \cap B_\epsilon(\mathbf{x})} |(\mathbf{y} - \mathbf{x}) \cdot \nu(\mathbf{x})|$ and $Q_\epsilon = \{\mathbf{y} \in \partial B_\epsilon(\mathbf{x}) : |(\mathbf{y} - \mathbf{x}) \cdot \nu(\mathbf{x})| \leq M_\epsilon\}$. Since $D_\epsilon \setminus C_\epsilon \subset Q_\epsilon$, we get that the portion A_ϵ of the sphere $\partial B_\epsilon(\mathbf{x})$ covered by $D_\epsilon \setminus C_\epsilon$ is smaller than the portion $4 \arctan(M_\epsilon / \sqrt{\epsilon^2 - M_\epsilon^2})$ of the sphere $\partial B_\epsilon(\mathbf{x})$ covered by Q_ϵ . However, we have from Lemma 9.9.23 that $M_\epsilon \leq C\epsilon^2$ for some constant C . Thus, $A_\epsilon \leq 4 \arctan(M_\epsilon / \sqrt{\epsilon^2 - M_\epsilon^2}) \rightarrow 0$ as $\epsilon \rightarrow 0$. It follows that

$$\frac{1}{\omega_n \epsilon^{n-1}} \int_{D_\epsilon \setminus C_\epsilon} dS_{\mathbf{y}} = A_\epsilon \rightarrow 0 \quad \text{as } \epsilon \rightarrow 0. \quad (9.9.27)$$

Similarly,

$$\frac{1}{\omega_n \epsilon^{n-1}} \int_{C_\epsilon \setminus D_\epsilon} dS_{\mathbf{y}} \rightarrow 0 \quad \text{as } \epsilon \rightarrow 0. \quad (9.9.28)$$

If we let ϵ converge to 0 in (9.9.25), and use (9.9.26), (9.9.27) and (9.9.28), we get (9.9.22) from (9.9.24). $\blacksquare$

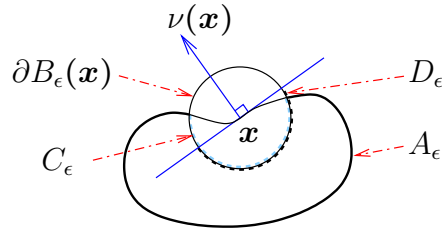


Figure 9.10: Figure for the proof of Lemma 9.9.25. The dashed blue line represents C_ϵ . The dashed black line represents D_ϵ . The tick black line represents A_ϵ .

Remark 9.9.26

Looking at the proof of the previous lemma, it is clear that (9.9.22) and (9.9.23) are still true if S is replaced by $\partial\Omega_j$ and Ω by Ω_j , where Ω_j is one of the components of Ω . For (9.9.23), the set Ω' is then replaced by the complement of $\overline{\Omega_j}$. $\spadesuit$

Lemma 9.9.27

There exists $B \in]0, \infty[$ such that

$$\int_S \left| \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) \right| dS_{\mathbf{y}} \leq B$$

for all $\mathbf{x} \in \mathbb{R}^n \setminus S$.

Proof.

Given $\mathbf{x} \in \mathbb{R}^n \setminus S$, let $d_{\mathbf{x}} = \min_{\mathbf{y} \in S} \|\mathbf{y} - \mathbf{x}\|$. Since S is compact, $d_{\mathbf{x}} > 0$.

Choose $\delta > 0$ such that $\delta < 1/(4C)$ for C given in Lemma 9.9.23 and such that

$$O_\delta = \{\mathbf{y} + t\nu(\mathbf{y}) : \mathbf{y} \in S, |t| < \delta\}$$

is a tubular neighbourhood of S . Such a δ exists according to Theorem 9.9.8.

i) If $d_{\mathbf{x}} \geq \delta$, then

$$\left| \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) \right| = \left| \frac{-1}{\omega_n \|\mathbf{x} - \mathbf{y}\|^n} (\mathbf{x} - \mathbf{y}) \cdot \nu(\mathbf{y}) \right| \leq \frac{1}{\omega_n \|\mathbf{x} - \mathbf{y}\|^{n-1}} \leq \frac{1}{\omega_n d_{\mathbf{x}}^{n-1}} \leq \frac{1}{\omega_n \delta^{n-1}} .$$

Thus,

$$\int_S \left| \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) \right| dS_{\mathbf{y}} \leq \frac{1}{\omega_n \delta^{n-1}} \int_S dS . \quad (9.9.29)$$

ii) If $d_{\mathbf{x}} < \delta$, then $\mathbf{x} \in O_{\delta}$. So, there exist a unique $\mathbf{x}_o \in S$ and $t_o \in]-\delta, \delta[$ such that $\mathbf{x} = \mathbf{x}_o + t_o \nu(\mathbf{x}_o)$. Let $S_{\mathbf{x}_o, 2\delta} = \{\mathbf{y} \in S : \|\mathbf{y} - \mathbf{x}_o\| < 2\delta\}$.

We have that

$$\int_S \left| \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) \right| dS_{\mathbf{y}} = \int_{S_{\mathbf{x}_o, 2\delta}} \left| \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) \right| dS_{\mathbf{y}} + \int_{S \setminus S_{\mathbf{x}_o, 2\delta}} \left| \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) \right| dS_{\mathbf{y}} .$$

For $\mathbf{y} \in S \setminus S_{\mathbf{x}_o, 2\delta}$, we have that

$$\|\mathbf{x} - \mathbf{y}\| \geq \|\mathbf{x}_o - \mathbf{y}\| - \|\mathbf{x} - \mathbf{x}_o\| \geq 2\delta - \delta = \delta .$$

Thus

$$\left| \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) \right| = \left| \frac{-1}{\omega_n \|\mathbf{x} - \mathbf{y}\|^n} (\mathbf{x} - \mathbf{y}) \cdot \nu(\mathbf{y}) \right| \leq \frac{1}{\omega_n \|\mathbf{x} - \mathbf{y}\|^{n-1}} \leq \frac{1}{\omega_n \delta^{n-1}}$$

and so

$$\int_{S \setminus S_{\mathbf{x}_o, 2\delta}} \left| \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) \right| dS_{\mathbf{y}} \leq \frac{1}{\omega_n \delta^{n-1}} \int_S dS . \quad (9.9.30)$$

For $\mathbf{y} \in S_{\mathbf{x}_o, 2\delta}$, we get from Lemma 9.9.23 that

$$\begin{aligned} \left| \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) \right| &= \left| \frac{-1}{\omega_n \|\mathbf{x} - \mathbf{y}\|^n} (\mathbf{x} - \mathbf{y}) \cdot \nu(\mathbf{y}) \right| \leq \frac{1}{\omega_n \|\mathbf{x} - \mathbf{y}\|^n} (|(\mathbf{x} - \mathbf{x}_o) \cdot \nu(\mathbf{y})| + |(\mathbf{x}_o - \mathbf{y}) \cdot \nu(\mathbf{y})|) \\ &\leq \frac{1}{\omega_n \|\mathbf{x} - \mathbf{y}\|^n} (\|\mathbf{x} - \mathbf{x}_o\| + C \|\mathbf{x}_o - \mathbf{y}\|)^2 \end{aligned} \quad (9.9.31)$$

because $\mathbf{x}_o, \mathbf{y} \in S$. We now find a lower bound for $\|\mathbf{x} - \mathbf{y}\|^n$ with $\mathbf{y} \in S_{\mathbf{x}_o, 2\delta}$. We have that

$$\|\mathbf{x} - \mathbf{y}\|^2 = \|(\mathbf{x} - \mathbf{x}_o) + (\mathbf{x}_o - \mathbf{y})\|^2 = \|\mathbf{x} - \mathbf{x}_o\|^2 + 2(\mathbf{x} - \mathbf{x}_o) \cdot (\mathbf{x}_o - \mathbf{y}) + \|\mathbf{x}_o - \mathbf{y}\|^2 .$$

We also have that

$$\begin{aligned} 2|(\mathbf{x} - \mathbf{x}_o) \cdot (\mathbf{x}_o - \mathbf{y})| &= 2\|\mathbf{x} - \mathbf{x}_o\| |\nu(\mathbf{x}_o) \cdot (\mathbf{x}_o - \mathbf{y})| = 2\|\mathbf{x} - \mathbf{x}_o\| |\nu(\mathbf{x}_o) \cdot (\mathbf{x}_o - \mathbf{y})| \\ &\leq 2C \|\mathbf{x} - \mathbf{x}_o\| \|\mathbf{x}_o - \mathbf{y}\|^2 , \end{aligned}$$

where again we have used Lemma 9.9.23 with $\mathbf{x}_o, \mathbf{y} \in S$. Since $\|\mathbf{x}_o - \mathbf{y}\| < 2\delta < 1/(2C)$, we get that

$$2|(\mathbf{x} - \mathbf{x}_o) \cdot (\mathbf{x}_o - \mathbf{y})| \leq \|\mathbf{x} - \mathbf{x}_o\| \|\mathbf{x}_o - \mathbf{y}\| .$$

Hence,

$$\|\mathbf{x} - \mathbf{y}\|^2 \geq \|\mathbf{x} - \mathbf{x}_o\|^2 - \|\mathbf{x} - \mathbf{x}_o\| \|\mathbf{x}_o - \mathbf{y}\| + \|\mathbf{x}_o - \mathbf{y}\|^2 \geq \frac{1}{2} (\|\mathbf{x} - \mathbf{x}_o\|^2 + \|\mathbf{x}_o - \mathbf{y}\|^2) .$$

If we substitute this expression in (9.9.31), we get

$$\begin{aligned} \left| \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) \right| &\leq \frac{2^{n/2}}{\omega_n (\|\mathbf{x} - \mathbf{x}_o\|^2 + \|\mathbf{x}_o - \mathbf{y}\|^2)^{n/2}} (\|\mathbf{x} - \mathbf{x}_o\| + C \|\mathbf{x}_o - \mathbf{y}\|)^2 \\ &\leq \frac{2^{n/2} \|\mathbf{x} - \mathbf{x}_o\|}{\omega_n (\|\mathbf{x} - \mathbf{x}_o\|^2 + \|\mathbf{x}_o - \mathbf{y}\|^2)^{n/2}} + \frac{2^{n/2} C}{\omega_n \|\mathbf{x}_o - \mathbf{y}\|^{n-2}} . \end{aligned}$$

Therefore,

$$\begin{aligned} \int_{S_{\mathbf{x}_o, 2\delta}} \left| \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) \right| dS_{\mathbf{y}} &\leq \frac{2^{n/2}}{\omega_n} \int_{S_{\mathbf{x}_o, 2\delta}} \frac{\|\mathbf{x} - \mathbf{x}_o\|}{(\|\mathbf{x} - \mathbf{x}_o\|^2 + \|\mathbf{x}_o - \mathbf{y}\|^2)^{n/2}} dS_{\mathbf{y}} \\ &\quad + \frac{2^{n/2} C}{\omega_n} \int_{S_{\mathbf{x}_o, 2\delta}} \frac{1}{\|\mathbf{x}_o - \mathbf{y}\|^{n-2}} dS_{\mathbf{y}} . \end{aligned} \quad (9.9.32)$$

If we assume that 4δ is smaller than the Lebesgue number associated to the atlas that we have introduced in Section 9.9.3, then we get from (9.9.17) with $\alpha = n - 2$ that

$$\int_{S_{\mathbf{x}_o, 2\delta}} \frac{1}{\|\mathbf{x}_o - \mathbf{y}\|^{n-2}} dS_{\mathbf{y}} \leq 2Q\omega_{n-1}\delta . \quad (9.9.33)$$

To find an upper bound for the first integral on the right side of (9.9.32), we use the fact that $S_{\mathbf{x}_o, 2\delta} \subset U_{\mathbf{x}_j}$ for one of the chart of the atlas that we have defined in Section 9.9.3. Thus

$$\begin{aligned} &\int_{S_{\mathbf{x}_o, 2\delta}} \frac{\|\mathbf{x} - \mathbf{x}_o\|}{(\|\mathbf{x} - \mathbf{x}_o\|^2 + \|\mathbf{x}_o - \mathbf{y}\|^2)^{n/2}} dS_{\mathbf{y}} \\ &= \int_{\psi_{\mathbf{x}_j}(S_{\mathbf{x}_o, 2\delta})} \frac{\|\mathbf{x} - \mathbf{x}_o\|}{(\|\mathbf{x} - \mathbf{x}_o\|^2 + \|\mathbf{x}_o - \psi_{\mathbf{x}_j}^{-1}(\mathbf{w})\|^2)^{n/2}} |\det D \psi_{\mathbf{x}_j}^{-1}(\mathbf{w})| d\mathbf{w} \\ &\leq Q \int_{B_{2\delta}(\psi_{\mathbf{x}_j}(\mathbf{x}_o))} \frac{\|\mathbf{x} - \mathbf{x}_o\|}{(\|\mathbf{x} - \mathbf{x}_o\|^2 + \|\psi_{\mathbf{x}_j}(\mathbf{x}_o) - \mathbf{w}\|^2)^{n/2}} d\mathbf{z} , \end{aligned}$$

where we have use (9.9.16), $\psi_{\mathbf{x}_j}(S_{\mathbf{x}_o, 2\delta}) \subset B_{2\delta}(\psi_{\mathbf{x}_j}(\mathbf{x}_o))$, and $\|\mathbf{x} - \mathbf{y}\| \geq \|\psi_{\mathbf{x}_j}(\mathbf{x}) - \psi_{\mathbf{x}_j}(\mathbf{y})\|$ for all $\mathbf{x}, \mathbf{y} \in U_{\mathbf{x}_j}$ to obtain the inequality. Using the change of variables $\mathbf{w} = \psi_{\mathbf{x}_j}(\mathbf{x}_o) + r\mathbf{z}$ with $0 \leq r < 2\delta$ and $\mathbf{z} \in \partial B_1(\mathbf{0}) \subset \mathbb{R}^{n-1}$, and posing $a = \|\mathbf{x} - \mathbf{x}_o\|$, we get

$$\begin{aligned} \int_{S_{\mathbf{x}_o, 2\delta}} \frac{\|\mathbf{x} - \mathbf{x}_o\|}{(\|\mathbf{x} - \mathbf{x}_o\|^2 + \|\mathbf{x}_o - \mathbf{y}\|^2)^{n/2}} dS_{\mathbf{y}} &\leq Q \int_{\partial B_1(\mathbf{0})} \int_0^{2\delta} \frac{ar^{n-2}}{(a^2 + r^2)^{n/2}} dr dS \\ &= Q\omega_{n-1} \int_0^{2\delta/a} \frac{s^{n-2}}{(1 + s^2)^{n/2}} ds \leq Q\omega_{n-1} \int_0^\infty \frac{s^{n-2}}{(1 + s^2)^{n/2}} ds , \end{aligned} \quad (9.9.34)$$

where we have used the substitution $r = as$. The last integral converges because

$$\int_0^R \frac{s^{n-2}}{(1 + s^2)^{n/2}} ds = \int_0^1 \frac{s^{n-2}}{(1 + s^2)^{n/2}} ds + \underbrace{\int_1^R \frac{s^{n-2}}{(1 + s^2)^{n/2}} ds}_{< \infty} + \underbrace{\int_1^R \frac{1}{s^2} ds}_{< 1}$$

for $R > 1$.

It follows from (9.9.30), (9.9.32), (9.9.33) and (9.9.34) that

$$\begin{aligned} \int_S \left| \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) \right| dS_{\mathbf{y}} &\leq \frac{1}{\omega_n \delta^{n-1}} \int_S dS + \frac{2^{n/2} Q \omega_{n-1}}{\omega_n} \int_0^\infty \frac{s^{n-2}}{(1+s^2)^{n/2}} ds \\ &\quad + \frac{2^{(n+2)/2} C Q \omega_{n-1} \delta}{\omega_n} \end{aligned} \quad (9.9.35)$$

iii) It follows from (9.9.29) and (9.9.35) that the Lemma is satisfied if B is equal or larger than

$$\frac{1}{\omega_n \delta^{n-1}} \int_S dS + \frac{2^{n/2} Q \omega_{n-1}}{\omega_n} \int_0^\infty \frac{s^{n-2}}{(1+s^2)^{n/2}} ds + \frac{2^{(n+2)/2} C Q \omega_{n-1} \delta}{\omega_n}.$$

Lemma 9.9.28

If $\phi \in C(S)$ is zero at some point $\mathbf{z} \in S$, then the double layer potential u is continuous at $\mathbf{z}$.

Proof.

According to Proposition 9.9.17, there exists a constant A such that $\int_S |K(\mathbf{x}, \mathbf{y})| dS_{\mathbf{y}} < A$ for all $\mathbf{x}$ since K is a continuous kernel of order $n-2$.

Given $\epsilon > 0$, choose $\eta > 0$ such that $|\phi(\mathbf{x})| < \frac{\epsilon}{3(A+B)}$ for $\mathbf{x} \in S_{\mathbf{z}, \eta} = S \cap B_\eta(\mathbf{z})$, where B is given in Lemma 9.9.27.

We have that

$$\begin{aligned} |u(\mathbf{x}) - u(\mathbf{z})| &\leq \int_{S_{\mathbf{z}, \eta}} \left| \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) - \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{z}, \mathbf{y}) \right| |\phi(\mathbf{y})| dS_{\mathbf{y}} \\ &\quad + \int_{S \setminus S_{\mathbf{z}, \eta}} \left| \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) - \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{z}, \mathbf{y}) \right| |\phi(\mathbf{y})| dS_{\mathbf{y}}, \end{aligned} \quad (9.9.36)$$

where obviously $\frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{w}, \mathbf{y}) = K(\mathbf{w}, \mathbf{y})$ when $\mathbf{w}, \mathbf{y} \in S$.

Since $(\mathbf{x}, \mathbf{y}) \rightarrow \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) - \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{z}, \mathbf{y})$ is continuous on the compact set $\overline{B_\eta(\mathbf{z})} \times (S \setminus S_{\mathbf{z}, \eta})$, it is uniformly continuous on this set. Thus, $\mathbf{y} \rightarrow \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) - \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{z}, \mathbf{y})$ converges uniformly to 0 on $S \setminus S_{\mathbf{z}, \eta}$ as $\mathbf{x} \rightarrow \mathbf{z}$. Hence, $\int_{S \setminus S_{\mathbf{z}, \eta}} \left| \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) - \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{z}, \mathbf{y}) \right| |\phi(\mathbf{y})| dS_{\mathbf{y}} \rightarrow 0$ as $\mathbf{x} \rightarrow \mathbf{z}$.

Choose $\delta > 0$ such that $\int_{S \setminus S_{\mathbf{z}, \eta}} \left| \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) - \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{z}, \mathbf{y}) \right| |\phi(\mathbf{y})| dS_{\mathbf{y}} < \frac{\epsilon}{3}$ for $\|\mathbf{x} - \mathbf{z}\| < \delta$. Since

$$\int_{S_{\mathbf{z}, \eta}} \left| \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) - \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{z}, \mathbf{y}) \right| |\phi(\mathbf{y})| dS_{\mathbf{y}}$$

$$\begin{aligned}
&\leq \int_{S_{z,\eta}} \left| \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) \right| |\phi(\mathbf{y})| dS_{\mathbf{y}} + \int_{S_{z,\eta}} \left| \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{z}, \mathbf{y}) \right| |\phi(\mathbf{y})| dS_{\mathbf{y}} \\
&\leq \frac{\epsilon}{3(A+B)} \left(\int_{S_{z,\eta}} \left| \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) \right| dS_{\mathbf{y}} + \int_{S_{z,\eta}} \left| \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{z}, \mathbf{y}) \right| dS_{\mathbf{y}} \right) \leq \frac{2\epsilon}{3},
\end{aligned}$$

we get from (9.9.36) that $|u(\mathbf{x}) - u(\mathbf{z})| < \epsilon$ for $\|\mathbf{x} - \mathbf{z}\| < \delta$. ■

Theorem 9.9.29

Consider the double layer potential given in Definition 9.9.22 with $\phi \in C(S)$. Then $u|_{\Omega}$ has a continuous extension to $\bar{\Omega}$ and $u|_{\Omega'}$ has a continuous extension to $\bar{\Omega}'$. To be more precise, let $u_t(\mathbf{x}) = u(\mathbf{x} + t\nu(\mathbf{x}))$ for $\mathbf{x} \in \partial\Omega$ and $t \in \mathbb{R}$,

$$u^{[i]}(\mathbf{x}) = \frac{1}{2}\phi(\mathbf{x}) + \int_S K(\mathbf{x}, \mathbf{y}) \phi(\mathbf{y}) dS_{\mathbf{y}} \quad (9.9.37)$$

for $\mathbf{x} \in S$, and

$$u^{[o]}(\mathbf{x}) = -\frac{1}{2}\phi(\mathbf{x}) + \int_S K(\mathbf{x}, \mathbf{y}) \phi(\mathbf{y}) dS_{\mathbf{y}} \quad (9.9.38)$$

for $\mathbf{x} \in S$. In other words, $u^{[i]} = \frac{1}{2}\phi + T_K\phi$ and $u^{[o]} = -\frac{1}{2}\phi + T_K\phi$ respectively on S . Then $u_t \rightarrow u^{[i]}$ uniformly on S as $t \rightarrow 0^-$ and $u_t \rightarrow u^{[o]}$ uniformly on S as $t \rightarrow 0^+$.

Proof.

In the discussion below, we assume that $|t|$ is small enough to have $\mathbf{x} + t\nu(\mathbf{x})$ in a tubular neighbourhood of S for all $\mathbf{x} \in S$.

i) For $\mathbf{x} \in S$ and $t < 0$, we have

$$u_t(\mathbf{x}) = \int_S \phi(\mathbf{y}) \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x} + t\nu(\mathbf{x}), \mathbf{y}) dS_{\mathbf{y}}.$$

From Lemma 9.9.25, we get that

$$\begin{aligned}
u_t(\mathbf{x}) &= \phi(\mathbf{x}) \int_S \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x} + t\nu(\mathbf{x}), \mathbf{y}) dS_{\mathbf{y}} + \int_S (\phi(\mathbf{y}) - \phi(\mathbf{x})) \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x} + t\nu(\mathbf{x}), \mathbf{y}) dS_{\mathbf{y}} \\
&= \phi(\mathbf{x}) + \int_S (\phi(\mathbf{y}) - \phi(\mathbf{x})) \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x} + t\nu(\mathbf{x}), \mathbf{y}) dS_{\mathbf{y}}.
\end{aligned}$$

Because $\mathbf{y} \mapsto \phi(\mathbf{y}) - \phi(\mathbf{x})$ is null at $\mathbf{y} = \mathbf{x}$, we have from Lemma 9.9.28 that the function

$$t \mapsto \int_{\partial\Omega} (\phi(\mathbf{y}) - \phi(\mathbf{x})) \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x} + t\nu(\mathbf{x}), \mathbf{y}) dS_{\mathbf{y}}$$

is continuous at $t = 0$ for $\mathbf{x}$ fixed. Thus,

$$\lim_{t \rightarrow 0^-} u_t(\mathbf{x}) = \phi(\mathbf{x}) + \int_S K(\mathbf{x}, \mathbf{y}) (\phi(\mathbf{y}) - \phi(\mathbf{x})) dS_{\mathbf{y}}$$

$$\begin{aligned}
&= \phi(\mathbf{x}) + \int_S K(\mathbf{x}, \mathbf{y}) \phi(\mathbf{y}) \, dS_{\mathbf{y}} - \phi(\mathbf{x}) \int_S K(\mathbf{x}, \mathbf{y}) \, dS_{\mathbf{y}} \\
&= \frac{1}{2} \phi(\mathbf{x}) + \int_S K(\mathbf{x}, \mathbf{y}) \phi(\mathbf{y}) \, dS_{\mathbf{y}} = u^{[i]}(\mathbf{x}) ,
\end{aligned}$$

where we have used Lemma 9.9.25 for the last equality.

ii) For $\mathbf{x} \in S$ and $t > 0$, we have

$$u_t(\mathbf{x}) = \int_S \phi(\mathbf{y}) \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x} + t\nu(\mathbf{x}), \mathbf{y}) \, dS_{\mathbf{y}} .$$

From Lemma 9.9.25, we get that

$$\begin{aligned}
u_t(\mathbf{x}) &= \phi(\mathbf{x}) \int_S \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x} + t\nu(\mathbf{x}), \mathbf{y}) \, dS_{\mathbf{y}} + \int_S (\phi(\mathbf{y}) - \phi(\mathbf{x})) \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x} + t\nu(\mathbf{x}), \mathbf{y}) \, dS_{\mathbf{y}} \\
&= \int_S (\phi(\mathbf{y}) - \phi(\mathbf{x})) \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x} + t\nu(\mathbf{x}), \mathbf{y}) \, dS_{\mathbf{y}} .
\end{aligned}$$

From here, the proof that

$$\lim_{t \rightarrow 0^+} u_t(\mathbf{x}) = -\frac{1}{2} \phi(\mathbf{x}) + \int_S K(\mathbf{x}, \mathbf{y}) \phi(\mathbf{y}) \, dS_{\mathbf{y}} = u^{[o]}(\mathbf{x})$$

is very similar to the proof in (i).

iii) We prove that the convergence $u_t \rightarrow u^{[i]}$ (resp. $u_t \rightarrow u^{[o]}$) is uniform on S as $t \rightarrow 0^-$ (resp. $t \rightarrow 0^+$).

Let B be the constant from Lemma 9.9.27 and let A be a constant such that $\int_S |K(\mathbf{x}, \mathbf{y})| \, dS_{\mathbf{y}} < A$ for $\mathbf{x} \in S$. Such a constant exists according to Proposition 9.9.17.

Given $\epsilon > 0$, choose $\eta > 0$ such that

$$|\phi(\mathbf{x}) - \phi(\mathbf{y})| < \frac{\epsilon}{2(A+B)} \quad (9.9.39)$$

for $\mathbf{x}, \mathbf{y} \in S$ and $\|\mathbf{x} - \mathbf{y}\| < \eta$. This is possible because ϕ is continuous on the compact set S . So ϕ is uniformly continuous on S .

We may also assume that η is small enough such that $O_\eta = \{\mathbf{x} + t\nu(\mathbf{x}) : \mathbf{x} \in S \text{ and } |t| < \eta\}$ is a tubular neighbourhood of S .

Since $(\mathbf{z}, \mathbf{y}) \mapsto \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{z}, \mathbf{y})$ is continuous (and so uniformly continuous) on the compact set $P = \{(\mathbf{z}, \mathbf{y}) \in \overline{O_{\eta/2}} \times S : \|\mathbf{z} - \mathbf{y}\| \geq \eta/2\}$, there exists $\delta < \eta/2$ such that

$$\left| \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{z}_1, \mathbf{y}) - K(\mathbf{z}_2, \mathbf{y}) \right| < \frac{\epsilon}{2} \left(\int_S |\phi(\mathbf{y})| \, dS_{\mathbf{y}} + \max_{\mathbf{x} \in S} |\phi(\mathbf{x})| \int_S dS_{\mathbf{y}} \right)^{-1} \quad (9.9.40)$$

for $(\mathbf{z}_1, \mathbf{y}), (\mathbf{z}_2, \mathbf{y}) \in P$ with $\|\mathbf{z}_1 - \mathbf{z}_2\| < \delta$ and $\mathbf{z}_2 \in S$.

Hence, for all $\mathbf{x} \in S$ and $|t| < \delta$, we have that

$$\begin{aligned}
 & |u_t(\mathbf{x}) - u^{[i]}(\mathbf{x})| \\
 &= \left| \int_S (\phi(\mathbf{y}) - \phi(\mathbf{x})) \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x} + t\nu(\mathbf{x}), \mathbf{y}) dS_{\mathbf{y}} - \int_S K(\mathbf{x}, \mathbf{y}) (\phi(\mathbf{y}) - \phi(\mathbf{x})) dS_{\mathbf{y}} \right| \\
 &\leq \int_S \left| \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x} + t\nu(\mathbf{x}), \mathbf{y}) - K(\mathbf{x}, \mathbf{y}) \right| |\phi(\mathbf{y}) - \phi(\mathbf{x})| dS_{\mathbf{y}} \\
 &\leq \int_{S \cap B_{\eta}(\mathbf{x})} \left(\left| \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x} + t\nu(\mathbf{x}), \mathbf{y}) \right| + |K(\mathbf{x}, \mathbf{y})| \right) |\phi(\mathbf{y}) - \phi(\mathbf{x})| dS_{\mathbf{y}} \\
 &\quad + \int_{S \setminus B_{\eta}(\mathbf{x})} \left| \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x} + t\nu(\mathbf{x}), \mathbf{y}) - K(\mathbf{x}, \mathbf{y}) \right| (|\phi(\mathbf{y})| + |\phi(\mathbf{x})|) dS_{\mathbf{y}} \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} < \epsilon,
 \end{aligned}$$

where we have used (9.9.39) to find an upper bound for the integral over $S \cap B_{\eta}(\mathbf{x})$, and (9.9.40) to find an upper bound for the integral over $S \setminus B_{\eta}(\mathbf{x})$. Moreover, we note that $\|\mathbf{x} + t\nu(\mathbf{x}) - \mathbf{y}\| \geq \eta/2$ for $|t| < \eta/2$ and $\|\mathbf{x} - \mathbf{y}\| \geq \eta$. So (9.9.40) can be effectively used (Figure 9.11). ■

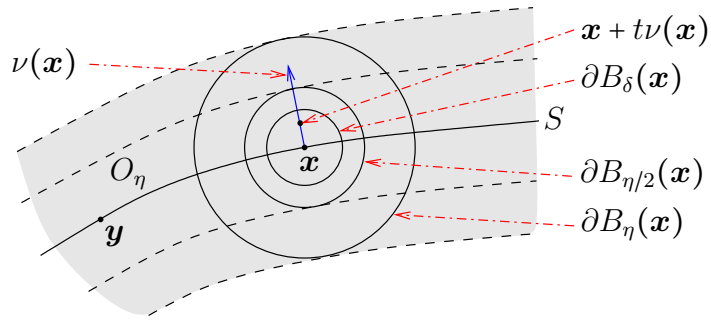


Figure 9.11: Figure for the proof of Theorem 9.9.29.

9.9.6 Single Layer Potential

As in the previous two sections, $S = \partial\Omega$.

Definition 9.9.30

The **single layer potential** with moment $\phi \in C(S)$ is

$$u(\mathbf{x}) = \int_S N(\mathbf{x}, \mathbf{y}) \phi(\mathbf{y}) dS_{\mathbf{y}} \quad , \quad \mathbf{x} \in \mathbb{R}^n \quad ,$$

where N is the fundamental solution for the Laplace operator given in (9.4.1) and Theorem 9.3.1.

Recall the following facts about N .

1. $\mathbf{y} \mapsto N(\mathbf{x}, \mathbf{y})$ is harmonic in $\mathbb{R}^n \setminus \{\mathbf{x}\}$.
2. For $n > 2$, $N(\mathbf{x}, \mathbf{y}) = O(\|\mathbf{y}\|^{2-n})$ as $\|\mathbf{y}\| \rightarrow \infty$. Hence $\mathbf{y} \mapsto N(\mathbf{x}, \mathbf{y})$ is harmonic at infinity according to Proposition 9.9.5.
3. (1) and (2) imply that the single layer potential u defined in Definition 9.9.30 is harmonic on $\mathbb{R}^n \setminus S$ and at infinity for $n > 2$.
4. The single layer potential u is well defined at $\mathbf{x} \in S$. Let $R \geq \text{diam } S = \max_{\mathbf{x}, \mathbf{y} \in S} \|\mathbf{x} - \mathbf{y}\|$, $\mathcal{B} = \{(U_{\mathbf{x}_j}, \psi_{\mathbf{x}_j})\}_{j \in J}$ be the atlas for S described in Section 9.9.3, and $\{\phi_{\mathbf{x}_j}\}_{j \in J}$ be a partition of unity for S subordinated to $\mathcal{B}$.

For $n > 2$, we have that

$$\begin{aligned}
 \int_S |N(\mathbf{x}, \mathbf{y})| |\phi(\mathbf{y})| dS_{\mathbf{y}} &\leq \frac{1}{(n-2)\omega_n} \|\phi\|_{\infty} \int_S \|\mathbf{x} - \mathbf{y}\|^{2-n} dS_{\mathbf{y}} \\
 &= \frac{1}{(n-2)\omega_n} \|\phi\|_{\infty} \sum_{j \in J} \int_{W_{\mathbf{x}_j}} \phi_{\mathbf{x}_j}(\psi_{\mathbf{x}_j}^{-1}(\mathbf{w})) \|\mathbf{x} - \psi_{\mathbf{x}_j}^{-1}(\mathbf{w})\|^{2-n} |\det \psi_{\mathbf{x}_j}^{-1}(\mathbf{w})| d\mathbf{w} \\
 &\leq \frac{Q}{(n-2)\omega_n} \|\phi\|_{\infty} \sum_{j \in J} \int_{W_{\mathbf{x}_j}} \|\psi_{\mathbf{x}_j}(\mathbf{x}) - \mathbf{w}\|^{2-n} d\mathbf{w} \\
 &\leq \frac{Q}{(n-2)\omega_n} \|\phi\|_{\infty} \sum_{j \in J} \int_{\partial B_1(\mathbf{0})} \int_0^R r^{2-n} r^{n-2} dr dS \leq \frac{Q|J|\omega_{n-1}R}{(n-2)\omega_n} \|\phi\|_{\infty} < \infty
 \end{aligned}$$

for all $\phi \in C(\partial\Omega)$ and all $\mathbf{x} \in S$, where we have used (9.9.16), $\|\mathbf{x} - \mathbf{y}\| \geq \|\psi_{\mathbf{x}_j}(\mathbf{x}) - \psi_{\mathbf{x}_j}(\mathbf{y})\|$ for all $\mathbf{x}, \mathbf{y} \in U_{\mathbf{x}_j}$ and $W_{\mathbf{x}_j} = \psi_{\mathbf{x}_j}(U_{\mathbf{x}_j}) \subset B_R(\psi_{\mathbf{x}_j}(\mathbf{x})) \subset \mathbb{R}^{n-1}$.

For $n = 2$, we have that

$$\begin{aligned}
 \int_S |N(\mathbf{x}, \mathbf{y})| |\phi(\mathbf{y})| dS_{\mathbf{y}} &\leq \frac{1}{\omega_2} \|\phi\|_{\infty} \int_S |\ln(\|\mathbf{x} - \mathbf{y}\|)| dS_{\mathbf{y}} \\
 &= \frac{1}{\omega_2} \|\phi\|_{\infty} \sum_{j \in J} \int_{W_{\mathbf{x}_j}} \phi_{\mathbf{x}_j}(\psi_{\mathbf{x}_j}^{-1}(w)) |\ln(\|\mathbf{x} - \psi_{\mathbf{x}_j}^{-1}(w)\|)| |\det \psi_{\mathbf{x}_j}^{-1}(w)| dw \\
 &\leq \frac{Q}{\omega_2} \|\phi\|_{\infty} \sum_{j \in J} \int_{W_{\mathbf{x}_j}} |\ln(|\psi_{\mathbf{x}_j}(\mathbf{x}) - w||)| dw \leq \frac{2Q}{\omega_2} \|\phi\|_{\infty} \sum_{j \in J} \int_0^R |\ln(r)| dr \\
 &\leq \frac{2Q|J|(2 + (|\ln(R)| + 1)R)}{\omega_2} \|\phi\|_{\infty} < \infty.
 \end{aligned}$$

for all $\phi \in C(\partial\Omega)$ and all $\mathbf{x} \in S$, where we have used (9.9.16), $\|\mathbf{x} - \mathbf{y}\| \geq \|\psi_{\mathbf{x}_j}(\mathbf{x}) - \psi_{\mathbf{x}_j}(\mathbf{y})\|$ for all $\mathbf{x}, \mathbf{y} \in U_{\mathbf{x}_j}$ and $W_{\mathbf{x}_j} = \psi_{\mathbf{x}_j}(U_{\mathbf{x}_j}) \subset]\psi_{\mathbf{x}_j}(\mathbf{x}) - R, \psi_{\mathbf{x}_j}(\mathbf{x}) + R[$.

Proposition 9.9.31

If $\phi \in C(S)$, then the single layer potential defined in Definition 9.9.30 is continuous on $\mathbb{R}^n$.

Proof.

We only have to show that u is continuous at an arbitrary point $\mathbf{z} \in \partial\Omega$ since u is continuous on $\mathbb{R}^n \setminus S$. Given $\mu > 0$, let $S_{\mathbf{z},\mu} = \{\mathbf{y} \in S : \|\mathbf{z} - \mathbf{y}\| < \mu\} = B_\mu(\mathbf{z}) \cap S$. We have

$$\begin{aligned} |u(\mathbf{x}) - u(\mathbf{z})| &= \left| \int_S (N(\mathbf{x}, \mathbf{y}) - N(\mathbf{z}, \mathbf{y})) \phi(\mathbf{y}) dS_{\mathbf{y}} \right| \\ &\leq \int_{S_{\mathbf{z},\mu}} |N(\mathbf{x}, \mathbf{y})| |\phi(\mathbf{y})| dS_{\mathbf{y}} + \int_{S_{\mathbf{z},\mu}} |N(\mathbf{z}, \mathbf{y})| |\phi(\mathbf{y})| dS_{\mathbf{y}} \\ &\quad + \int_{S \setminus S_{\mathbf{z},\mu}} |N(\mathbf{x}, \mathbf{y}) - N(\mathbf{z}, \mathbf{y})| |\phi(\mathbf{y})| dS_{\mathbf{y}}. \end{aligned} \quad (9.9.41)$$

Since ϕ is continuous on the compact set S , we have that $\|\phi\|_\infty = \max_{\mathbf{x} \in S} |\phi(\mathbf{x})| < \infty$.

Given $\epsilon > 0$, we first choose μ small enough such that the sum of the first two integrals is less than $2\epsilon/3$ for all $\mathbf{x} \in B_{\mu/2}(\mathbf{z}) \subset \mathbb{R}^n$ as we explain below.

i) Choose μ such that $4\mu < 1$ is smaller than the Lebesgue number associated to the atlas that we have introduced in Section 9.9.3. So $B_{2\mu}(\mathbf{z}) \subset V_{\mathbf{x}_j}$ for some j . Since $B_{2\mu}(\mathbf{z}) \supset B_{3\mu/2}(\mathbf{x}) \supset B_\mu(\mathbf{z})$ for all $\mathbf{x} \in B_{\mu/2}(\mathbf{z})$, we get that $S_{\mathbf{x},3\mu/2} = B_{3\mu/2}(\mathbf{x}) \cap S \supset B_\mu(\mathbf{z}) \cap S = S_{\mathbf{z},\mu}$ for all $\mathbf{x} \in B_{\mu/2}(\mathbf{z})$. Thus, we have for $n > 2$ that

$$\begin{aligned} \int_{S_{\mathbf{z},\mu}} |N(\mathbf{x}, \mathbf{y})| |\phi(\mathbf{y})| dS_{\mathbf{y}} &\leq \|\phi\|_\infty \int_{S_{\mathbf{x},3\mu/2}} |N(\mathbf{x}, \mathbf{y})| dS_{\mathbf{y}} = \frac{\|\phi\|_\infty}{(n-2)\omega_n} \int_{S_{\mathbf{x},3\mu/2}} \|\mathbf{x} - \mathbf{y}\|^{2-n} dS_{\mathbf{y}} \\ &= \frac{\|\phi\|_\infty}{(n-2)\omega_n} \int_{\psi_{\mathbf{x}_j}(S_{\mathbf{x},3\mu/2})} \|\mathbf{x} - \psi_{\mathbf{x}_j}^{-1}(\mathbf{w})\|^{2-n} |\det D\psi_{\mathbf{x}_j}^{-1}(\mathbf{w})| d\mathbf{w} \\ &\leq \frac{\|\phi\|_\infty Q}{(n-2)\omega_n} \int_{\psi_{\mathbf{x}_j}(S_{\mathbf{x},3\mu/2})} \|\psi_{\mathbf{x}_j}(\mathbf{x}) - \mathbf{w}\|^{2-n} d\mathbf{w} \\ &\leq \frac{\|\phi\|_\infty Q}{(n-2)\omega_n} \int_{\partial B_1(\mathbf{0})} \int_0^{3\mu/2} dr dS = \frac{3\|\phi\|_\infty \omega_{n-1} Q \mu}{2(n-2)\omega_n} \end{aligned}$$

for all $\mathbf{x} \in B_{\mu/2}(\mathbf{z})$. As is usual, for the previous computations, we have used (9.9.16) and $\|\mathbf{x} - \mathbf{y}\| \geq \|\psi_{\mathbf{x}_j}(\mathbf{x}) - \psi_{\mathbf{x}_j}(\mathbf{y})\|$ for all $\mathbf{x}, \mathbf{y} \in V_{\mathbf{x}_j}$, where we consider $\psi_{\mathbf{x}_j}$ to be the projection of $\mathbb{R}^n$ onto $T_{\mathbf{x}_j}$ as defined in Section 9.9.3.

Similarly, we find for $n = 2$ that

$$\begin{aligned} \int_{S_{\mathbf{z},\mu}} |N(\mathbf{x}, \mathbf{y})| |\phi(\mathbf{y})| dS_{\mathbf{y}} &\leq \|\phi\|_\infty \int_{S_{\mathbf{x},3\mu/2}} |N(\mathbf{x}, \mathbf{y})| dS_{\mathbf{y}} = \frac{\|\phi\|_\infty}{\omega_2} \int_{S_{\mathbf{x},3\mu/2}} |\ln(\|\mathbf{x} - \mathbf{y}\|)| dS_{\mathbf{y}} \\ &\leq \frac{\|\phi\|_\infty Q}{\omega_2} \int_0^{3\mu/2} |\ln(r)| dr = \frac{3\|\phi\|_\infty Q (|\ln(3\mu/2)| + 1)\mu}{2\omega_2} \end{aligned}$$

for all $\mathbf{x} \in B_{\mu/2}(\mathbf{z})$.

ii) We get directly from (9.9.17) and (9.9.18) that

$$\int_{S_{\mathbf{z},\mu}} |N(\mathbf{z}, \mathbf{y})| |\phi(\mathbf{y})| dS_{\mathbf{y}} \leq \frac{\|\phi\|_\infty}{(n-2)\omega_n} \int_{S_{\mathbf{z},\mu}} \|\mathbf{z} - \mathbf{y}\|^{2-n} dS_{\mathbf{y}} \leq \frac{\|\phi\|_\infty Q \omega_{n-1} \mu}{(n-2)\omega_n}$$

for $n > 2$, and

$$\int_{S_{\mathbf{z},\mu}} |N(\mathbf{z}, \mathbf{y})| |\phi(\mathbf{y})| dS_{\mathbf{y}} \leq \frac{\|\phi\|_{\infty}}{\omega_2} \int_{S_{\mathbf{z},\mu}} |\ln(\|\mathbf{z} - \mathbf{y}\|)| dS_{\mathbf{y}} \leq \frac{\|\phi\|_{\infty} Q(|\ln(\mu)| + 1) \mu}{\omega_2}$$

for $n = 2$.

iii) It follows from (i) and (ii) that we can choose μ small enough such that the sum of the first two integrals in (9.9.41) is less than $2\epsilon/3$ for all $\mathbf{x} \in B_{\mu/2}(\mathbf{z}) \subset \mathbb{R}^n$.

iv) Since $(\mathbf{x}, \mathbf{y}) \mapsto N(\mathbf{x}, \mathbf{y})$ is continuous on the compact set $\overline{B_{\mu/2}(\mathbf{z})} \times (S \setminus S_{\mathbf{z},\mu})$, it is uniformly continuous on this compact set. So, there exists $0 < \delta < \mu/2$ such that

$$\sup_{\mathbf{y} \in S \setminus S_{\mathbf{z},\mu}} |N(\mathbf{x}, \mathbf{y}) - N(\mathbf{z}, \mathbf{y})| < \epsilon \left(3\|\phi\|_{\infty} \int_S dS_{\mathbf{y}} \right)^{-1}$$

for all $\mathbf{x} \in B_{\delta}(\mathbf{z}) \subset \mathbb{R}^n$. Hence

$$\int_{S \setminus S_{\mathbf{z},\mu}} |N(\mathbf{x}, \mathbf{y}) - N(\mathbf{z}, \mathbf{y})| |\phi(\mathbf{y})| dS_{\mathbf{y}} \leq \|\phi\|_{\infty} \sup_{\mathbf{y} \in S \setminus S_{\mathbf{z},\mu}} |N(\mathbf{x}, \mathbf{y}) - N(\mathbf{z}, \mathbf{y})| \int_{S \setminus S_{\mathbf{z},\mu}} dS_{\mathbf{y}} < \frac{\epsilon}{3}$$

for all $\mathbf{x} \in B_{\delta}(\mathbf{z})$.

v) We get from (9.9.41), (iii) and (iv) that $|u(\mathbf{x}) - u(\mathbf{x}_0)| \leq \epsilon$ for all $\mathbf{x} \in B_{\delta}(\mathbf{w})$. ■

Let O_{μ} be a tubular neighbourhood of S . For all $\mathbf{z}_t \in O_{\mu}$, we define a directional derivative by the formula

$$\frac{\tilde{\partial} u}{\tilde{\partial} \nu}(\mathbf{z}_t) \equiv \frac{d}{dt} u(\mathbf{x} + t\nu(\mathbf{x})) = \nabla u(\mathbf{z}_t) \cdot \nu(\mathbf{x}) = \nabla u(\mathbf{x} + t\nu(\mathbf{x})) \cdot \nu(\mathbf{x})$$

where $\mathbf{z}_t = \mathbf{x} + t\nu(\mathbf{x})$ with $\mathbf{x} \in S$ is the unique representation of $\mathbf{z}_t$ in the tubular neighbourhood O_{μ} . The directional derivative of the single layer potential u at $\mathbf{x} \in O_{\mu} \setminus S$ is therefore

$$\frac{\tilde{\partial} u}{\tilde{\partial} \nu}(\mathbf{x}) = \int_S \phi(\mathbf{y}) \frac{\tilde{\partial} N}{\tilde{\partial} \nu_{\mathbf{x}}}(\mathbf{x}, \mathbf{y}) dS_{\mathbf{y}},$$

where $\frac{\tilde{\partial}}{\tilde{\partial} \nu_{\mathbf{x}}}$ indicates that the directional derivative is with respect to the variable $\mathbf{x}$ while $\mathbf{y}$ is constant.

Let $K^*(\mathbf{x}, \mathbf{y}) = K(\mathbf{y}, \mathbf{x})$ for $\mathbf{x}, \mathbf{y} \in S$, where K is defined in (9.9.21). We have that

$$K(\mathbf{p}, \mathbf{q}) = \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{p}, \mathbf{q}) = -\frac{1}{\omega_n \|\mathbf{p} - \mathbf{q}\|^n} (\mathbf{p} - \mathbf{q}) \cdot \nu(\mathbf{q}) = \frac{1}{\omega_n \|\mathbf{q} - \mathbf{p}\|^n} (\mathbf{q} - \mathbf{p}) \cdot \nu(\mathbf{q}) = \frac{\partial N}{\partial \nu_{\mathbf{x}}}(\mathbf{q}, \mathbf{p})$$

for $\mathbf{p}, \mathbf{q} \in S$ with $\mathbf{p} \neq \mathbf{q}$. Thus $K^*(\mathbf{p}, \mathbf{q}) = \frac{\partial N}{\partial \nu_{\mathbf{x}}}(\mathbf{p}, \mathbf{q})$ for $\mathbf{p}, \mathbf{q} \in S$ with $\mathbf{p} \neq \mathbf{q}$.

We expand $\frac{\tilde{\partial}u}{\tilde{\partial}\nu}$ to O_μ with the definition

$$\frac{\tilde{\partial}u}{\tilde{\partial}\nu}(\mathbf{x}) = \begin{cases} \int_S \phi(\mathbf{y}) \frac{\tilde{\partial}N}{\tilde{\partial}\nu_{\mathbf{x}}}(\mathbf{x}, \mathbf{y}) dS_{\mathbf{y}} & \text{if } \mathbf{x} \in O_\mu \setminus S \\ \int_S K^*(\mathbf{x}, \mathbf{y}) \phi(\mathbf{y}) dS_{\mathbf{y}} & \text{if } \mathbf{x} \in S \end{cases} \quad (9.9.42)$$

Since K is a continuous kernel of order $n-2$, we also have that K^* is a continuous kernel of order $n-2$. Hence, it follows from Propositions 9.9.16 and 9.9.19 that $\frac{\tilde{\partial}u}{\tilde{\partial}\nu}\Big|_S$ is well defined as $T_{K^*}\phi$, and $\frac{\tilde{\partial}u}{\tilde{\partial}\nu}\Big|_S \in C(S)$ if $\phi \in C(S)$.

According to Proposition 9.9.18, $T_K : L^2(S) \rightarrow L^2(S)$ and $T_{K^*} : L^2(S) \rightarrow L^2(S)$ are compact operators. Moreover, since K and K^* are real valued functions, we have

$$\begin{aligned} \langle \psi, T_K \phi \rangle_2 &= \int_S \psi(\mathbf{x}) \left(\int_S K(\mathbf{x}, \mathbf{y}) \phi(\mathbf{y}) dS_{\mathbf{y}} \right) dS_{\mathbf{x}} = \int_S \psi(\mathbf{x}) \left(\int_S K(\mathbf{x}, \mathbf{y}) \overline{\phi(\mathbf{y})} dS_{\mathbf{y}} \right) dS_{\mathbf{x}} \\ &= \int_S \left(\int_S K(\mathbf{x}, \mathbf{y}) \psi(\mathbf{x}) dS_{\mathbf{x}} \right) \overline{\phi(\mathbf{y})} dS_{\mathbf{y}} = \int_S \left(\int_S K^*(\mathbf{y}, \mathbf{x}) \psi(\mathbf{x}) dS_{\mathbf{x}} \right) \overline{\phi(\mathbf{y})} dS_{\mathbf{y}} = \langle T_{K^*} \psi, \phi \rangle_2 \end{aligned}$$

for $\phi, \psi \in L^2(S)$. Thus $T_{K^*} = T_K^*$, the adjoint operator to T_K on $L^2(S)$.

Lemma 9.9.32

For η small enough, there exists a constant C_1 such that

$$\|\nu(\mathbf{y}) - \nu(\mathbf{x})\| < C_1 \|\mathbf{y} - \mathbf{x}\|$$

for all $\mathbf{x}, \mathbf{y} \in S$ with $\|\mathbf{x} - \mathbf{y}\| < \eta$. Moreover, there exists a constant C_2 such that

$$\|(\mathbf{x} + t\nu(\mathbf{x})) - \mathbf{y}\| \geq C_2 \|\mathbf{x} - \mathbf{y}\|$$

for all $\mathbf{x}, \mathbf{y} \in S$ with $\|\mathbf{x} - \mathbf{y}\| < \eta$, and $t \in \mathbb{R}$.

Proof.

Let $\mathcal{B} = \{(U_{\mathbf{x}_j}, \psi_{\mathbf{x}_j})\}_{j \in J}$ be the atlas for S described in Section 9.9.3, and let ϵ be the Lebesgue number associated to the atlas $\mathcal{B}$.

Given $\mathbf{x} \in S$, there exists an index j such that $S_{\mathbf{x}, \epsilon/2} = B_{\epsilon/2}(\mathbf{x}) \cap S \subset U_{\mathbf{x}_j}$.

We may assume that $\nu \circ \psi_{\mathbf{x}_j}^{-1} : \overline{W_{\mathbf{x}_j}} \rightarrow \mathbb{R}^n$ is of class C^1 since we assume $\psi_{\mathbf{x}_j}^{-1} \in C^2(\overline{W_{\mathbf{x}_j}})$ ¹². Let M_j be the maximum of $\|D(\nu \circ \psi_{\mathbf{x}_j}^{-1})\|$ on the compact set $\overline{W_j}$. For all $\mathbf{y} \in S_{\mathbf{x}, \epsilon/2}$, we have that

$$\|\nu(\mathbf{y}) - \nu(\mathbf{x})\| = \|\nu(\psi_{\mathbf{x}_j}^{-1}(\check{\mathbf{y}})) - \nu(\psi_{\mathbf{x}_j}^{-1}(\check{\mathbf{x}}))\| \leq M_j \|\check{\mathbf{y}} - \check{\mathbf{x}}\| \leq M_j \|\mathbf{y} - \mathbf{x}\| ,$$

¹²If $\mathbf{z} \in S \subset \mathbb{R}^3$, we have that $\nu(\mathbf{z}) = \nu(\psi_{\mathbf{x}_j}^{-1}(\check{\mathbf{z}})) = (1/\|\mathbf{v}\|)\mathbf{v}$, where $\check{\mathbf{z}} = \psi_{\mathbf{x}_j}(\mathbf{z})$ and $\mathbf{v} = \frac{\partial \psi_{\mathbf{x}_j}^{-1}}{\partial \check{z}_1}(\check{\mathbf{z}}) \times \frac{\partial \psi_{\mathbf{x}_j}^{-1}}{\partial \check{z}_2}(\check{\mathbf{z}})$, the vector product

where $\check{y} = \psi_{x_j}(\mathbf{y})$ and $\check{x} = \psi_{x_j}(\mathbf{x})$. In the previous relation, we have used $\|\mathbf{x} - \mathbf{y}\| \geq \|\psi_{x_j}(\mathbf{x}) - \psi_{x_j}(\mathbf{y})\|$ for all $\mathbf{x}, \mathbf{y} \in U_{x_j}$ to get the last inequality.

We get the first conclusion of the lemma with $C_1 = \max_{j \in J} M_j$ and $\eta \leq \epsilon/2$.

To get the second conclusion of the lemma, we may assume that the U_{x_j} in the definition of the atlas $\mathcal{B}$ are small enough to satisfy $\nu(\mathbf{x}_1) \cdot \nu(\mathbf{x}_2) \geq 1/\sqrt{2}$ for all $\mathbf{x}_1, \mathbf{x}_2 \in U_{x_j}$. This is possible because $\nu : S \rightarrow \mathbb{R}^n$ is of class C^1 . This means that the cosine of the angle between the vectors $\nu(\mathbf{x}_1)$ and $\nu(\mathbf{x}_2)$ is between $1/\sqrt{2}$ and 1 for all $\mathbf{x}_1, \mathbf{x}_2 \in U_{x_j}$.

Let $\mathbf{y}_p$ be the orthogonal projection of $\mathbf{y}$ on the line $\{\mathbf{x} + t\nu(\mathbf{x}) : t \in \mathbb{R}\}$, and θ be the angle between the vectors $\mathbf{y}_p - \mathbf{y}$ and $\mathbf{x} - \mathbf{y}$. Then

$$\|(\mathbf{x} + t\nu(\mathbf{x})) - \mathbf{y}\| \geq \|\mathbf{y}_p - \mathbf{y}\| = \cos(\theta)\|\mathbf{y} - \mathbf{x}\| \geq \frac{1}{\sqrt{2}}\|\mathbf{y} - \mathbf{x}\| \quad (9.9.43)$$

for all $\mathbf{y} \in S_{\mathbf{x}, \epsilon/2}$ as we can observe from the sketch in Figure 9.12. We note that (9.9.43) is true for all $\mathbf{x}$. Thus, we get the second conclusion of the lemma with $C_2 = 1/\sqrt{2}$ and $\eta \leq \epsilon/2$.

By shrinking η , we may have C_2 as close to 1 as we want. This is expected because we assume that S is “smooth”. ■

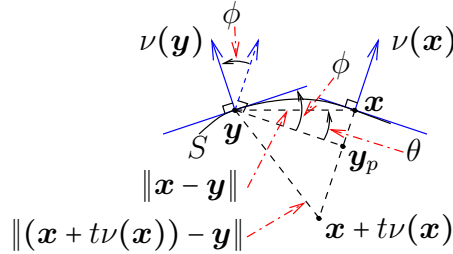


Figure 9.12: Sketch associated to the proof of Theorem 9.9.33 where $0 \leq \phi \leq \pi/4$ is the angle between $\nu(\mathbf{x})$ and $\nu(\mathbf{y})$.

Theorem 9.9.33

If $\phi \in C(S)$, then $u|_{\bar{\Omega}} \in C_{\nu(\Omega)}$ and $u|_{\bar{\Omega}'} \in C_{\nu(\Omega')}$. Moreover,

$$\frac{\partial u}{\partial \nu^-}(\mathbf{x}) = -\frac{1}{2}\phi(\mathbf{x}) + \int_{\partial\Omega} K^*(\mathbf{x}, \mathbf{y})\phi(\mathbf{y}) \, dS_{\mathbf{y}}$$

for $\mathbf{x} \in S$ and

$$\frac{\partial u}{\partial \nu^+}(\mathbf{x}) = \frac{1}{2}\phi(\mathbf{x}) + \int_{\partial\Omega} K^*(\mathbf{x}, \mathbf{y})\phi(\mathbf{y}) \, dS_{\mathbf{y}}$$

for $\mathbf{x} \in S$. In other words, $\frac{\partial u}{\partial \nu^-} = -\frac{1}{2}\phi + T_K^*\phi$ and $\frac{\partial u}{\partial \nu^+} = \frac{1}{2}\phi + T_K^*\phi$ on S .

If $w_t(\mathbf{x}) = \frac{\tilde{\partial} u}{\tilde{\partial} \nu}(\mathbf{x} + t\nu(\mathbf{x}))$ for $\mathbf{x} \in S$ and $\mathbf{x} + t\nu(\mathbf{x}) \in O_\mu$, then $w_t \rightarrow \frac{\partial u}{\partial \nu^-}$ uniformly on

S as $t \rightarrow 0^-$ and $w_t \rightarrow \frac{\partial u}{\partial \nu^+}$ uniformly on S as $t \rightarrow 0^+$.

Proof.

From Proposition 9.9.31, we have that $u \in C(\mathbb{R}^n)$. Hence, $u \in C(\overline{\Omega})$ and $u \in C(\overline{\Omega'})$. By definition of the single layer potential, we also have that $u \in C^1(\Omega)$ and $u \in C^1(\Omega')$.

Consider the double layer potential of moment ϕ defined by

$$v(\mathbf{x}) = \int_S \phi(\mathbf{y}) \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) dS_{\mathbf{y}}$$

for $\mathbf{x} \in \mathbb{R}^n \setminus S$, and the function

$$f(\mathbf{x}) = \begin{cases} v(\mathbf{x}) + \frac{\tilde{\partial} u}{\tilde{\partial} \nu_{\mathbf{x}}}(\mathbf{x}) & \text{if } \mathbf{x} \in O_{\mu} \setminus S \\ (T_K \phi)(\mathbf{x}) + (T_K^* \phi)(\mathbf{x}) & \text{if } \mathbf{x} \in S \end{cases}$$

i) We claim that f is continuous on O_{μ} . Note that O_{μ} is an open subset of $\mathbb{R}^n$ because S is a manifold of dimension $n-1$. It is clear from its definition that f is continuous on $O_{\mu} \setminus S$. To prove that f is continuous on O_{μ} , it is enough to prove that

$$\sup_{\mathbf{x} \in S} |f(\mathbf{x} + t\nu(\mathbf{x})) - f(\mathbf{x})| \rightarrow 0 \quad \text{as } t \rightarrow 0.$$

This will implies that $f|_{O_{\mu} \setminus S}$ can be continuously expanded to a continuous function on O_{μ} by $f(\mathbf{x}) = (T_K \phi)(\mathbf{x}) + (T_K^* \phi)(\mathbf{x})$ for $\mathbf{x} \in S$.

We now prove the claim. Let $\mathbf{z}_t = \mathbf{x} + t\nu(\mathbf{x})$ for $\mathbf{x} \in S$ and $S_{\mathbf{x}, \eta} = S \cap B_{\eta}(\mathbf{x})$ with $\eta < \mu$ given in Lemma 9.9.32. Recall from the proof of Lemma 9.9.32 that 2η is smaller than the Lebesgue number associated to the atlas $\mathcal{B}$ defined in Section 9.9.3. Thus, for each $\mathbf{x}$, there exists j such that $S_{\mathbf{x}, \eta} \subset U_{\mathbf{x}_j}$. we have that

$$\begin{aligned} & |f(\mathbf{z}_t) - f(\mathbf{x})| \\ &= \left| \int_S \phi(\mathbf{y}) \left(\frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{z}_t, \mathbf{y}) + \frac{\tilde{\partial} N}{\tilde{\partial} \nu_{\mathbf{x}}}(\mathbf{z}_t, \mathbf{y}) \right) dS_{\mathbf{y}} - \int_S \phi(\mathbf{y}) \left(\frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) + \frac{\partial N}{\partial \nu_{\mathbf{x}}}(\mathbf{x}, \mathbf{y}) \right) dS_{\mathbf{y}} \right| \\ &\leq I_1(\mathbf{x}, t, \eta) + I_2(\mathbf{x}, t, \eta) + I_3(\mathbf{x}, t, \eta) + I_4(\mathbf{x}, t, \eta), \end{aligned}$$

where

$$\begin{aligned} I_1(\mathbf{x}, t, \eta) &= \int_{S \setminus S_{\mathbf{x}, \eta}} |\phi(\mathbf{y})| \left| \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{z}_t, \mathbf{y}) - \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) \right| dS_{\mathbf{y}}, \\ I_2(\mathbf{x}, t, \eta) &= \int_{S \setminus S_{\mathbf{x}, \eta}} |\phi(\mathbf{y})| \left| \frac{\tilde{\partial} N}{\tilde{\partial} \nu_{\mathbf{x}}}(\mathbf{z}_t, \mathbf{y}) - \frac{\partial N}{\partial \nu_{\mathbf{x}}}(\mathbf{x}, \mathbf{y}) \right| dS_{\mathbf{y}}, \\ I_3(\mathbf{x}, t, \eta) &= \int_{S_{\mathbf{x}, \eta}} |\phi(\mathbf{y})| \left| \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{z}_t, \mathbf{y}) + \frac{\tilde{\partial} N}{\tilde{\partial} \nu_{\mathbf{x}}}(\mathbf{z}_t, \mathbf{y}) \right| dS_{\mathbf{y}} \end{aligned}$$

and

$$I_4(\mathbf{x}, t, \eta) = \int_{S_{\mathbf{x}, \eta}} |\phi(\mathbf{y})| \left| \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) + \frac{\partial N}{\partial \nu_{\mathbf{x}}}(\mathbf{x}, \mathbf{y}) \right| dS_{\mathbf{y}}.$$

i.a) We consider $I_3(\mathbf{x}, t, \eta)$. Since ϕ is continuous on the compact set S , we have that $\|\phi\|_{\infty} = \sup_{\mathbf{y} \in S} |\phi(\mathbf{y})| < \infty$. Hence,

$$I_3(\mathbf{x}, t, \eta) \leq \|\phi\|_{\infty} \int_{S_{\mathbf{x}, \eta}} \left| \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{z}_t, \mathbf{y}) + \frac{\tilde{\partial} N}{\tilde{\partial} \nu_{\mathbf{x}}}(\mathbf{z}_t, \mathbf{y}) \right| dS_{\mathbf{y}}.$$

Using Lemma 9.9.32, we have that

$$\begin{aligned} \left| \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{z}_t, \mathbf{y}) + \frac{\tilde{\partial} N}{\tilde{\partial} \nu_{\mathbf{x}}}(\mathbf{z}_t, \mathbf{y}) \right| &= \left| \frac{1}{\omega_n \|\mathbf{z}_t - \mathbf{y}\|^n} (\mathbf{z}_t - \mathbf{y}) \cdot (\nu(\mathbf{x}) - \nu(\mathbf{y})) \right| \\ &\leq \frac{1}{\omega_n \|\mathbf{z}_t - \mathbf{y}\|^n} \|\mathbf{z}_t - \mathbf{y}\| \|\nu(\mathbf{x}) - \nu(\mathbf{y})\| \leq \frac{C_1}{\omega_n \|\mathbf{z}_t - \mathbf{y}\|^{n-1}} \|\mathbf{x} - \mathbf{y}\| \leq \frac{C_1}{\omega_n C_2^{n-1} \|\mathbf{x} - \mathbf{y}\|^{n-2}} \end{aligned}$$

for $\|\mathbf{x} - \mathbf{y}\| < \eta$ and $|t| < \eta$. It follows from (9.9.17) that

$$I_3(\mathbf{x}, t, \eta) \leq \frac{C_1 \|\phi\|_{\infty}}{\omega_n C_2^{n-1}} \int_{S_{\mathbf{x}, \eta}} \|\mathbf{x} - \mathbf{y}\|^{2-n} dS_{\mathbf{y}} = \frac{C_1 \|\phi\|_{\infty} \omega_{n-1} Q}{C_2^{n-1} \omega_n} \eta$$

for $\mathbf{x} \in S$ and $\mathbf{z}_t = \mathbf{x} + t\nu(\mathbf{x})$ with $|t| < \eta$.

i.b) A similar but simpler reasoning to the one above yields

$$I_4(\mathbf{x}, t, \eta) = \int_{S_{\mathbf{x}, \eta}} |\phi(\mathbf{y})| \left| \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) + \frac{\partial N}{\partial \nu_{\mathbf{x}}}(\mathbf{x}, \mathbf{y}) \right| dS_{\mathbf{y}} \leq \frac{C_1 \|\phi\|_{\infty} \omega_{n-1} Q}{\omega_n} \eta$$

for $\mathbf{x} \in S$.

i.c) Given $\epsilon > 0$, choose η small enough such that $I_3(\mathbf{x}, t, \eta)$ and $I_4(\mathbf{x}, t, \eta)$ are smaller than $\epsilon/4$ for $\mathbf{x} \in S$ and $\mathbf{z}_t = \mathbf{x} + t\nu(\mathbf{x})$ with $|t| < \eta$.

i.d) Since $(\mathbf{x}, \mathbf{y}) \mapsto \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y})$ is continuous on the compact set $\{(\mathbf{x}, \mathbf{y}) \in \overline{O_{\nu}} \times S : \|\mathbf{x} - \mathbf{y}\| \geq \eta\}$, it is uniformly continuous on this compact set. So, there exists $0 < \delta_1 < \eta$ such that

$$\left| \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{z}_t, \mathbf{y}) - \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) \right| = \left| \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x} + t\nu(\mathbf{x}), \mathbf{y}) - \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) \right| < \frac{\epsilon}{4} \left(\int_{S \setminus S_{\mathbf{x}, \eta}} |\phi(\mathbf{y})| dS_{\mathbf{y}} \right)^{-1}$$

for all $\mathbf{y} \in S \setminus S_{\mathbf{x}, \eta}$, $|t| < \delta_1$ and $\mathbf{x} \in S$. Hence, $I_1(\mathbf{x}, t, \eta) \leq \epsilon/4$ for all $\mathbf{x} \in S$ and $|t| < \delta_1$.

i.e) Since $(t, \mathbf{x}, \mathbf{y}) \mapsto \nabla N(\mathbf{x} + t\nu(\mathbf{x}) - \mathbf{y}) \cdot \nu(\mathbf{x})$ is continuous on the compact set $[\eta/2, \eta/2] \times \{(\mathbf{x}, \mathbf{y}) \in S \times S : \|\mathbf{x} - \mathbf{y}\| \geq \eta\}$, it is uniformly continuous on this compact set. So, there exists $0 < \delta_2 < \eta/2$ such that

$$\left| \frac{\tilde{\partial} N}{\tilde{\partial} \nu_{\mathbf{x}}}(\mathbf{z}_t, \mathbf{y}) - \frac{\partial N}{\partial \nu_{\mathbf{x}}}(\mathbf{x}, \mathbf{y}) \right| = |\nabla N(\mathbf{x} + t\nu(\mathbf{x}) - \mathbf{y}) \cdot \nu(\mathbf{x}) - \nabla N(\mathbf{x} - \mathbf{y}) \cdot \nu(\mathbf{x})|$$

$$\leq \frac{\epsilon}{4} \left(\int_{S \setminus S_{\mathbf{x}, \eta}} |\phi(\mathbf{y})| \, dS_{\mathbf{y}} \right)^{-1}$$

for all $\mathbf{y} \in S \setminus S_{\mathbf{x}, \eta}$, $|t| < \delta_2$ and $\mathbf{x} \in S$. Hence, $I_2(\mathbf{x}, t, \eta) \leq \epsilon/4$ for all $\mathbf{x} \in S$ and $|t| < \delta_2$.

i.f) Hence, for $\delta < \min\{\delta_1, \delta_2\}$, we have that $|f(\mathbf{z}_t) - f(\mathbf{x})| < \epsilon$ for $\mathbf{x} \in S$ and $|t| < \delta$. This proves our claim.

ii) Since $f \in C(O_\mu)$, we have

$$\begin{aligned} (T_K \phi)(\mathbf{x}) + (T_K^* \phi)(\mathbf{x}) &= \lim_{t \rightarrow 0^-} f(\mathbf{x} + t\nu(\mathbf{x})) = \lim_{t \rightarrow 0^-} v(\mathbf{x} + t\nu(\mathbf{x})) + \lim_{t \rightarrow 0^-} \frac{\tilde{\partial} u}{\tilde{\partial} \nu_{\mathbf{x}}}(\mathbf{x} + t\nu(\mathbf{x})) \\ &= \frac{1}{2} \phi(\mathbf{x}) + (T_K \phi)(\mathbf{x}) + \frac{\partial u}{\partial \nu^-}(\mathbf{x}) \end{aligned}$$

for all $\mathbf{x} \in S$, where the last equality comes from Theorem 9.9.29 and the definition of $\frac{\partial u}{\partial \nu^-}$. Thus,

$$\frac{\partial u}{\partial \nu^-}(\mathbf{x}) = -\frac{1}{2} \phi(\mathbf{x}) + (T_K^* \phi)(\mathbf{x})$$

for all $\mathbf{x} \in S$. Similarly,

$$\begin{aligned} (T_K \phi)(\mathbf{x}) + (T_K^* \phi)(\mathbf{x}) &= \lim_{t \rightarrow 0^+} f(\mathbf{x} + t\nu(\mathbf{x})) = \lim_{t \rightarrow 0^+} v(\mathbf{x} + t\nu(\mathbf{x})) + \lim_{t \rightarrow 0^+} \frac{\tilde{\partial} u}{\tilde{\partial} \nu_{\mathbf{x}}}(\mathbf{x} + t\nu(\mathbf{x})) \\ &= -\frac{1}{2} \phi(\mathbf{x}) + (T_K \phi)(\mathbf{x}) + \frac{\partial u}{\partial \nu^+}(\mathbf{x}) \end{aligned}$$

for all $\mathbf{x} \in S$, where the last equality comes from Theorem 9.9.29 and the definition of $\frac{\partial u}{\partial \nu^+}$. Thus,

$$\frac{\partial u}{\partial \nu^+}(\mathbf{x}) = \frac{1}{2} \phi(\mathbf{x}) + (T_K^* \phi)(\mathbf{x})$$

for all $\mathbf{x} \in S$.

iii) From Theorem 9.9.29, we have that

$$\sup_{\mathbf{x} \in S} |v(\mathbf{x} + t\nu(\mathbf{x})) - v^{[i]}(\mathbf{x})| = \sup_{\mathbf{x} \in S} \left| v(\mathbf{x} + t\nu(\mathbf{x})) - \frac{1}{2} \phi(\mathbf{x}) - (T_K \phi)(\mathbf{x}) \right| \rightarrow 0 \quad \text{as } t \rightarrow 0^-.$$

From the claim in (i) above, we have that

$$\begin{aligned} \sup_{\mathbf{x} \in S} \left| \left(v(\mathbf{x} + t\nu(\mathbf{x})) + \frac{\tilde{\partial} u}{\tilde{\partial} \nu}(\mathbf{x} + t\nu(\mathbf{x})) \right) - \left((T_K \phi)(\mathbf{x}) + (T_K^* \phi)(\mathbf{x}) \right) \right| \\ = \sup_{\mathbf{x} \in S} |f(\mathbf{x} + t\nu(\mathbf{x})) - f(\mathbf{x})| \rightarrow 0 \quad \text{as } t \rightarrow 0^-. \end{aligned}$$

Thus,

$$\sup_{\mathbf{x} \in S} \left| w_t(\mathbf{x}) - \frac{\partial u}{\partial \nu^-}(\mathbf{x}) \right| = \sup_{\mathbf{x} \in S} \left| \frac{\tilde{\partial} u}{\tilde{\partial} \nu}(\mathbf{x} + t\nu(\mathbf{x})) + \frac{1}{2} \phi(\mathbf{x}) - (T_K^* \phi)(\mathbf{x}) \right| \rightarrow 0 \quad \text{as } t \rightarrow 0^-.$$

This also proves that (9.9.1) is satisfied and thus completes the proof that $u \in C_\nu(\Omega)$.

The proof that $w_t \rightarrow \frac{\partial u}{\partial \nu^+}$ uniformly on S as $t \rightarrow 0^+$ is similar. We then also get that (9.9.2) is satisfied and thus complete the proof that $u \in C_\nu(\Omega')$. ■

Corollary 9.9.34

In the previous theorem, $\phi = \frac{\partial u}{\partial \nu^+} - \frac{\partial u}{\partial \nu^-}$ on S .

The following two lemmas will be useful in the next section.

Lemma 9.9.35

If $\phi \in C(S)$ and $f = \frac{1}{2}\phi + T_K^*\phi$ on S , where K is defined in (9.9.21), then $\int_S f \, dS = \int_S \phi \, dS$.

Proof.

The result is a consequence of Fubini's theorem and Lemma 9.9.25.

$$\begin{aligned} \int_S f(\mathbf{x}) \, dS &= \frac{1}{2} \int_S \phi(\mathbf{x}) \, dS + \int_S \int_S K^*(\mathbf{x}, \mathbf{y}) \phi(\mathbf{y}) \, dS_{\mathbf{y}} \, dS_{\mathbf{x}} \\ &= \frac{1}{2} \int_S \phi(\mathbf{x}) \, dS + \int_S \left(\int_S K^*(\mathbf{x}, \mathbf{y}) \, dS_{\mathbf{x}} \right) \phi(\mathbf{y}) \, dS_{\mathbf{y}} \\ &= \frac{1}{2} \int_S \phi(\mathbf{x}) \, dS + \int_S \left(\int_S K(\mathbf{y}, \mathbf{x}) \, dS_{\mathbf{x}} \right) \phi(\mathbf{y}) \, dS_{\mathbf{y}} \\ &= \frac{1}{2} \int_S \phi(\mathbf{x}) \, dS + \frac{1}{2} \int_S \phi(\mathbf{y}) \, dS = \int_S \phi(\mathbf{x}) \, dS. \end{aligned}$$

Lemma 9.9.36

Suppose that $n = 2$ and $\phi \in C(S)$. Let u be the single layer potential with moment ϕ .

1. u is harmonic at infinity if and only if $\int_S \phi(\mathbf{x}) \, dS = 0$.
2. u harmonic at infinity implies that u vanishes at infinity.
3. If $\int_S \phi(\mathbf{x}) \, dS = 0$ and u is constant on $\bar{\Omega}$, then $\phi = 0$ on S , and therefore $u = 0$ on $\bar{\Omega}$.

Proof.

Since S is a compact set, we have that $M = \max_{\mathbf{y} \in S} \|\mathbf{y}\| < \infty$. For $\|\mathbf{x}\| > M$, we have that

$$\|\mathbf{x}\| + M \geq \|\mathbf{x}\| + \|\mathbf{y}\| \geq \|\mathbf{x} - \mathbf{y}\| \geq \|\mathbf{x}\| - \|\mathbf{y}\| \geq \|\mathbf{x}\| - M$$

and therefore

$$|\ln(\|\mathbf{x} - \mathbf{y}\|) - \ln(\|\mathbf{x}\|)| = \left| \ln\left(\frac{\|\mathbf{x} - \mathbf{y}\|}{\|\mathbf{x}\|}\right) \right| \leq \max \left\{ \left| \ln\left(1 + \frac{M}{\|\mathbf{x}\|}\right) \right|, \left| \ln\left(1 - \frac{M}{\|\mathbf{x}\|}\right) \right| \right\}$$

for all $\mathbf{y} \in S$. It follows that $\ln(\|\mathbf{x} - \mathbf{y}\|) - \ln(\|\mathbf{x}\|) \rightarrow 0$ uniformly for $\mathbf{y} \in S$ as $\|\mathbf{x}\| \rightarrow \infty$. Hence,

$$\lim_{\|\mathbf{x}\| \rightarrow \infty} \int_S (\ln(\|\mathbf{x} - \mathbf{y}\|) - \ln(\|\mathbf{x}\|)) \phi(\mathbf{y}) dS_{\mathbf{y}} = 0.$$

We get that

$$\begin{aligned} \frac{u(\mathbf{x})}{\ln(\|\mathbf{x}\|)} &= \frac{1}{\omega_2 \ln(\|\mathbf{x}\|)} \int_S (\ln(\|\mathbf{x} - \mathbf{y}\|) - \ln(\|\mathbf{x}\|)) \phi(\mathbf{y}) dS_{\mathbf{y}} + \frac{1}{\omega_2} \int_S \phi(\mathbf{y}) dS_{\mathbf{y}} \\ &\rightarrow \frac{1}{\omega_2} \int_S \phi(\mathbf{y}) dS_{\mathbf{y}} \quad \text{as } \|\mathbf{x}\| \rightarrow \infty. \end{aligned}$$

It follows from Proposition 9.9.5 that u is harmonic at infinity if and only if $\int_S \phi(\mathbf{y}) dS_{\mathbf{y}} = 0$.

Moreover, if $\int_S \phi(\mathbf{y}) dS_{\mathbf{y}} = 0$, then

$$\begin{aligned} u(\mathbf{x}) &= \frac{1}{\omega_2} \int_S (\ln(\|\mathbf{x} - \mathbf{y}\|) - \ln(\|\mathbf{x}\|)) \phi(\mathbf{y}) dS_{\mathbf{y}} + \frac{\ln(\|\mathbf{x}\|)}{\omega_2} \int_S \phi(\mathbf{y}) dS_{\mathbf{y}} \\ &= \frac{1}{\omega_2} \int_S (\ln(\|\mathbf{x} - \mathbf{y}\|) - \ln(\|\mathbf{x}\|)) \phi(\mathbf{y}) dS_{\mathbf{y}} \rightarrow 0 \quad \text{as } \|\mathbf{x}\| \rightarrow \infty. \end{aligned}$$

Thus u vanishes at infinity.

To prove the last statement of the lemma, suppose that $u(\mathbf{x}) = c$, a constant, for all $\mathbf{x} \in \overline{\Omega}$. Thus, in particular, $u(\mathbf{x}) = c$ for all $\mathbf{x} \in S$. According to (1), u is also harmonic at infinity. Thus, u is a solution of the exterior Dirichlet problem (N2) with $f(\mathbf{x}) = c$ for all $\mathbf{x} \in S$. But the solution of this problem is unique and $u(\mathbf{x}) = c$ for all $\mathbf{x} \in \Omega'$ is a solution. Therefore, $u(\mathbf{x}) = c$ for all $\mathbf{x} \in \mathbb{R}^n$. Since u is constant, it follows from Corollary 9.9.34 that $\phi(\mathbf{x}) = 0$ for all $\mathbf{x} \in S$. ■

9.9.7 Existence of Solutions

The following subspaces of $L^2(\partial\Omega)$ will play a fundamental role in the solutions of the problems (D1), (D2), (N1) and (N2) given at the beginning of the chapter.

$$\begin{aligned} V_+ &= \left\{ \phi \in L^2(\partial\Omega) : T_K \phi = \frac{1}{2} \phi \right\}, & W_+ &= \left\{ \phi \in L^2(\partial\Omega) : T_K^* \phi = \frac{1}{2} \phi \right\}, \\ V_- &= \left\{ \phi \in L^2(\partial\Omega) : T_K \phi = -\frac{1}{2} \phi \right\} & \text{and} & \quad W_- = \left\{ \phi \in L^2(\partial\Omega) : T_K^* \phi = -\frac{1}{2} \phi \right\}. \end{aligned}$$

It follows from Lemma 9.9.24 and Proposition 9.9.18 that T_K and T_K^* are compact linear operators from $L^2(\partial\Omega)$ to itself.

Proposition 9.9.37

We have

1. $\dim V_+ = \dim W_+ = m$, where m is the number of connected components of Ω .
2. $\dim V_- = \dim W_- = m'$, where m' is the number of bounded and connected components of Ω' .
3. If $\mathbf{a} \in \mathbb{C}^{m'}$, then there exists a unique $\phi_{\mathbf{a}} \in W_-$ such that the single layer potential w with moment $\phi_{\mathbf{a}}$ satisfies $w|_{\Omega'_0} = 0$ and $w|_{\Omega'_j} = a_j$ for $1 \leq j \leq m'$.
4. If $n > 2$ and $\mathbf{a} \in \mathbb{C}^m$, then there exists a unique $\phi_{\mathbf{a}} \in W_+$ such that the single layer potential u with moment $\phi_{\mathbf{a}}$ satisfies $u|_{\Omega_j} = a_j$ for $1 \leq j \leq m$.
5. If $n = 2$, there exists an $(m-1)$ -dimensional subspace X of $\mathbb{C}^m$ with the following properties.
 - (a) $\mathbb{C}^m = X \oplus \mathbb{C}\{(1, 1, 1, \dots, 1)\}$.
 - (b) If $\mathbf{a} \in X$, there exists a unique

$$\phi_{\mathbf{a}} \in W_+^0 = \left\{ \phi \in W_+ : \int_{\partial\Omega} \phi(\mathbf{x}) \, dS_{\mathbf{x}} = 0 \right\}$$

such that the single layer potential u with moment $\phi_{\mathbf{a}}$ satisfies $u|_{\Omega_j} = a_j$ for $1 \leq j \leq m$.

Proof.

Consider $\phi_i : \partial\Omega \rightarrow \mathbb{R}$ for $1 \leq i \leq m$, and $\phi'_j : \partial\Omega \rightarrow \mathbb{R}$ for $1 \leq j \leq m'$, defined by

$$\phi_i(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{x} \in \partial\Omega_i \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad \phi'_j(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{x} \in \partial\Omega'_j \\ 0 & \text{otherwise} \end{cases} \quad (9.9.44)$$

respectively.

i) It follows from Lemma 9.9.25 (see Remark 9.9.26) with $\partial\Omega$ replaced by $\partial\Omega_i$ that

$$(T_K \phi_i)(\mathbf{x}) = \int_{\partial\Omega_i} K(\mathbf{x}, \mathbf{y}) \, dS_{\mathbf{y}} = \frac{1}{2} \phi_i(\mathbf{x})$$

for all $\mathbf{x} \in \partial\Omega$. Thus $\phi_i \in V_+$ for $1 \leq i \leq m$. Since the ϕ_i are linearly independent, $\dim V_+ \geq m$.

ii) We have that

$$(T_K \phi'_j)(\mathbf{x}) = \int_{\partial\Omega'_j} K(\mathbf{x}, \mathbf{y}) \, dS_{\mathbf{y}} = - \sum_{i \in I} \int_{\partial\Omega_i} K(\mathbf{x}, \mathbf{y}) \, dS_{\mathbf{y}},$$

where $I \subset \{1, 2, \dots, m\}$ is the set of indices such that $\partial\Omega'_j = \bigcup_{i \in I} \partial\Omega_i$. The minus sign in front of the sum comes from the change of orientation between the integral over $\partial\Omega'_j$ and the

integrals over the $\partial\Omega_i$. If $\mathbf{x} \in \partial\Omega'_j$, then $\mathbf{x} \in \partial\Omega_i$ for one and only one $i_0 \in I$. In particular, $\mathbf{x}$ is in the complement of $\overline{\Omega_i}$ for all $i \neq i_0$. As in (i), it follows from Lemma 9.9.25 with $\partial\Omega$ replace by $\partial\Omega_i$ that

$$\int_{\partial\Omega_i} K(\mathbf{x}, \mathbf{y}) dS_{\mathbf{y}} = \begin{cases} 1/2 & \text{if } i = i_0 \\ 0 & \text{if } i \neq i_0 \end{cases}$$

Thus

$$(T_K \phi'_j)(\mathbf{x}) = -\frac{1}{2} \phi'_j(\mathbf{x})$$

for all $\mathbf{x} \in \partial\Omega'$. Hence, $\phi'_j \in V_-$ for $1 \leq j \leq m'$. Since the ϕ'_j are linearly independent, $\dim V_- \geq m'$.

iii) Since $-\frac{1}{2}\text{Id} + T_K$ and $\frac{1}{2}\text{Id} + T_K$ are compact operators, it follows from Theorem 4.4.24 that $\dim \mathcal{N}\left(-\frac{1}{2}\text{Id} + T_K\right) = \dim \mathcal{N}\left(-\frac{1}{2}\text{Id} + T_K^*\right)$ and $\dim \mathcal{N}\left(\frac{1}{2}\text{Id} + T_K\right) = \dim \mathcal{N}\left(\frac{1}{2}\text{Id} + T_K^*\right)$.

Since $V_+ = \mathcal{N}\left(-\frac{1}{2}\text{Id} + T_K\right)$ and $W_+ = \mathcal{N}\left(-\frac{1}{2}\text{Id} + T_K^*\right)$, we get that $\dim V_+ = \dim W_+$. Since $V_- = \mathcal{N}\left(\frac{1}{2}\text{Id} + T_K\right)$ and $W_- = \mathcal{N}\left(\frac{1}{2}\text{Id} + T_K^*\right)$, we get that $\dim V_- = \dim W_-$.

iv) If $\phi \in W_+$ and u is the single layer potential with moment ϕ , then $\frac{\partial u}{\partial \nu^-}(\mathbf{x}) = 0$ for $\mathbf{x} \in \partial\Omega$ as a consequence of Theorem 9.9.33. By uniqueness (modulo a function which is constant on each connected component of Ω) of the solution for the interior Neumann problem (N1), we have that u is constant on each connected component of Ω . We may therefore define the mapping

$$\begin{aligned} \Lambda : W_+ &\rightarrow \mathbb{C}^m \\ \phi &\mapsto \left(u|_{\Omega_1}, u|_{\Omega_2}, \dots, u|_{\Omega_m} \right) \end{aligned}$$

iv.a) Λ is one-to-one for $n > 2$. Suppose that $\Lambda u_1 = \Lambda u_2$, where u_1 and u_2 are the single layer potentials with moments ϕ_1 and ϕ_2 respectively. Then $u = u_1 - u_2$ is the single layer potential with moment $\phi = \phi_1 - \phi_2$. Moreover, $u = 0$ on each component of Ω . Thus, $u = 0$ on $\overline{\Omega}$. By uniqueness of solution for the exterior Dirichlet problem (D2) given by $\Delta v = 0$ on $\Omega' \cup \{\infty\}$ with $v|_{\partial\Omega'} = 0$, we have that $u = 0$ on Ω' because $v \equiv 0$ is a solution of this problem. Hence, $u = 0$ on $\mathbb{R}^n$. We have seen in item (2) following Definition 9.9.30 that u is harmonic at infinity. It follows from Corollary 9.9.34 that $\phi = 0$ on $\partial\Omega$. Thus $\phi_1 = \phi_2$ on $\partial\Omega$.

This shows that $\dim W_+ \leq \dim \Lambda(W_+) \leq m$. Combined with (i) and (iii), this shows that $\dim W_+ = m$. Thus proving item (1) of the proposition for $n > 2$.

We also have that $\Lambda|_{W_+}$ is onto because $\dim \Lambda(W_+) = m$. Hence, item (4) of the proposition is true because $\Lambda|_{W_+}$ is one-to-one and onto $\mathbb{C}^m$.

iv.b) If $n = 2$, Λ defined above may not be one-to-one because the single layer potential u with moment ϕ may not be harmonic at infinity. Let

$$W_+^0 = \left\{ \phi \in W_+ : \int_{\partial\Omega} \phi(\mathbf{x}) dS_{\mathbf{x}} = 0 \right\}.$$

Using Lemma 9.9.36, a reasoning very similar to the one given in (iv.a) shows that $\Lambda|_{W_+^0} : W_+^0 \rightarrow \mathbb{C}^m$ is one-to-one. In particular, this proves that $\dim W_+^0 \leq m$.

Moreover, item (3) of Lemma 9.9.36 shows that the space $Y = \{\alpha(1, 1, \dots, 1) : \alpha \in \mathbb{C}\}$ is not in the image X of $\Lambda|_{W_+^0}$. Thus $\dim W_+^0 \leq m - 1$ because $\Lambda|_{W_+^0}$ is one-to-one. Also $\dim W_+^0 \geq \dim W_+ - 1$ because W_+^0 is the kernel of the linear functional $\phi \mapsto \int_{\partial\Omega} \phi(\mathbf{x}) dS_{\mathbf{x}}$ from W_+ to $\mathbb{C}$. Therefore, $\dim W_+ \leq m$. Combined with (i) and (iii), this shows that item (1) of the proposition is true for $n = 2$.

Since $\dim W_+ = m$, it follows that $\dim W_+^0 = m - 1$. Hence, $X = \Lambda(W_+^0)$ is of dimension $m - 1$ because $\Lambda|_{W_+^0}$ is one-to-one. We then get that $\mathbb{C}^m = X \oplus \mathbb{C}\{(1, 1, 1, \dots, 1)\}$. This completes the proof of item (5) of the proposition because $\Lambda|_{W_+^0}$ is one-to-one and onto X .

v) We first claim that if $\phi \in W_-$ and u is the single layer potential with moment ϕ , then u is constant on each bounded component Ω'_j of Ω' and $u = 0$ on the unbounded component Ω'_0 . We have from Theorem 9.9.33 that $\frac{\partial u}{\partial \nu^+}(\mathbf{x}) = 0$ for $\mathbf{x} \in \partial\Omega'$. For $n > 2$, the claim is a consequence of Proposition 9.9.9. For $n = 2$, Proposition 9.9.9 still implies that u is constant on each bounded component of Ω' and also on Ω'_0 . From Lemma 9.9.35, we have that $\int_{\partial\Omega} \phi(\mathbf{x}) dS_{\mathbf{x}} = 0$. Hence, Lemma 9.9.36 implies that u vanishes at infinity. Since u is constant on Ω'_0 , $u = 0$ on Ω'_0 . This completes the proof of the claim.

We may therefore define the mapping

$$\begin{aligned} \Lambda : W_- &\rightarrow \mathbb{C}^{m'} \\ \phi &\mapsto \left(u|_{\Omega'_1}, u|_{\Omega'_2}, \dots, u|_{\Omega'_{m'}} \right) \end{aligned}$$

We now show that Λ is one-to-one. Suppose that $\Lambda u_1 = \Lambda u_2$, where u_1 and u_2 are the single layer potentials with moments ϕ_1 and ϕ_2 respectively. Then $u = u_1 - u_2$ is the single layer potential with moment $\phi = \phi_1 - \phi_2$. Moreover, $u = 0$ on each bounded component of Ω' . We also have that $u = 0$ on the unbounded components since this is true for u_1 and u_2 . Thus, $u = 0$ on $\overline{\Omega'}$. By uniqueness of the interior Dirichlet problem (D1) given by $\Delta v = 0$ on Ω with $v|_{\partial\Omega} = 0$, we have that $u = 0$ on Ω because $v \equiv 0$ is a solution of this problem. Hence, $u = 0$ on $\mathbb{R}^n$. It follows from Corollary 9.9.34 that $\phi = 0$ on $\partial\Omega'$. Thus $\phi_1 = \phi_2$ on $\partial\Omega'$.

This shows that $\dim W_- \leq \dim \Lambda(W_-) \leq m'$. Combined with (ii), this shows that $\dim W_- = m'$ and proves item (2) of the proposition.

We also have that $\Lambda|_{W_-}$ is onto because $\dim \Lambda(W_-) = m'$. Hence, item (3) of the proposition is true because $\Lambda|_{W_-}$ is one-to-one and onto $\mathbb{C}^{m'}$. ■

Proposition 9.9.38

$$L^2(\partial\Omega) = V_+^\perp \oplus W_+ = V_-^\perp \oplus W_- .$$

In the previous proposition, the orthogonality is based on the scalar product on $L^2(\partial\Omega)$ defined by

$$\langle f, g \rangle = \int_{\partial\Omega} f(\mathbf{x}) \overline{g(\mathbf{x})} dS$$

for $f, g \in L^2(\partial\Omega)$.

Proof.

i) We first prove that $L^2(\partial\Omega) = V_+^\perp \oplus W_+$. It follows from Proposition 9.9.37 that $V_+^\perp$ is a closed subspace of codimension m in $L^2(\partial\Omega)$ and W_+ is a subspace of dimension m in $L^2(\partial\Omega)$. Thus, we only have to prove that $V_+^\perp \cap W_+ = \{0\}$.

Suppose that $\phi \in V_+^\perp \cap W_+$. Since $\phi \in W_+$, we have that $\phi \in \mathcal{N}\left(T_K^* - \frac{1}{2} \text{Id}\right)$. Since $\phi \in V_+^\perp = \left(\mathcal{N}\left(T_K - \frac{1}{2} \text{Id}\right)\right)^\perp$, it follows from Theorem 4.4.41 that $\phi \in \mathcal{R}\left(T_K^* - \frac{1}{2} \text{Id}\right)$ because T_K^* is a compact operator and so $\mathcal{R}\left(T_K^* - \frac{1}{2} \text{Id}\right)$ is a closed subspace according to Theorem 4.4.24. Thus, there exists $\psi \in L^2(\partial\Omega)$ such that $\phi = T_K^* \psi - \frac{1}{2} \psi$. Moreover, since $T_K^* \phi - \frac{1}{2} \phi = 0$, it follows from Proposition 9.9.20 that $\phi \in C(\partial\Omega)$. We can use Proposition 9.9.20 a second time to conclude that $\psi \in C(\partial\Omega)$ because $T_K^* \psi - \frac{1}{2} \psi = \phi$.

Let u and v be the simple layer potentials with moments ϕ and ψ respectively. We have that $u, v \in C(\mathbb{R}^n)$ according to Proposition 9.9.31.

From Theorem 9.9.33, we have that $\frac{\partial u}{\partial \nu^-} = T_K^* \phi - \frac{1}{2} \phi = 0$ and $\frac{\partial v}{\partial \nu^-} = T_K^* \psi - \frac{1}{2} \psi = \phi = T_K^* \phi + \frac{1}{2} \phi = \frac{\partial u}{\partial \nu^+}$ on $\partial\Omega$. Hence,

$$\int_{\partial\Omega} \left(u \frac{\partial v}{\partial \nu^-} - v \frac{\partial u}{\partial \nu^-} \right) dS = \int_{\partial\Omega} u \frac{\partial u}{\partial \nu^+} dS, \quad (9.9.45)$$

where the positive direction on $\partial\Omega$ is such that the normal to $\partial\Omega$ points outside Ω , so inside Ω' .

From the Green's identity (9.1.2), we have that

$$\int_{\partial\Omega} \left(u \frac{\partial v}{\partial \nu^-} - v \frac{\partial u}{\partial \nu^-} \right) dS = \int_{\Omega} (u \Delta v - v \Delta u) d\mathbf{x} = 0 \quad (9.9.46)$$

because u and v are harmonic on Ω .

To compute the integral on the right side in (9.9.45), we choose $r > 0$ such that $\partial B_r(\mathbf{0}) \subset \Omega'_0$, the unbounded component of Ω' . We get from the Green's identities (9.1.1) that

$$\begin{aligned} \int_{\partial\Omega} u \frac{\partial u}{\partial \nu^+} dS &= - \int_{\Omega' \cap B_r(\mathbf{0})} (u \Delta u + \nabla u \cdot \nabla u) d\mathbf{x} + \int_{\partial B_r(\mathbf{0})} u \frac{\partial u}{\partial \nu} dS \\ &= - \int_{\Omega' \cap B_r(\mathbf{0})} \|\nabla u\|_2^2 d\mathbf{x} + \int_{\partial B_r(\mathbf{0})} u \frac{\partial u}{\partial \nu} dS \end{aligned} \quad (9.9.47)$$

because u is harmonic in Ω' . The normal to $\partial(\Omega' \cap B_r(\mathbf{0}))$ points inside $\Omega' \cap B_r(\mathbf{0})$.

We now show that the last integral in (9.9.47) converges to 0 as $r \rightarrow \infty$. We get from item (2) following Definition 9.9.30 that u is harmonic at infinity for $n > 2$. To prove that u is harmonic at infinity for $n = 2$ requires a little bit more work. Since $\Omega' \cap B_r(\mathbf{0}) \phi \in V_+^\perp$, we have that

$$\int_{\partial\Omega} \phi \, dS = \int_{\partial\Omega} \phi \left(\sum_{i=0}^m \phi_i \right) dS = \sum_{i=0}^m \int_{\partial\Omega} \phi \phi_i \, dS = 0,$$

where the functions $\phi_i \in V_+$ are defined at the beginning of the proof of Proposition 9.9.37. Note that $\sum_{i=0}^m \phi_i(\mathbf{x}) = 1$ for all $\mathbf{x} \in \partial\Omega$. It follows from Lemma 9.9.36 that u is harmonic at infinity.

If $n > 2$, we have from Proposition 9.9.5 and Proposition 9.9.6 that $\Omega' \cap B_r(\mathbf{0})|u(\mathbf{x})| \leq C_1 \|\mathbf{x}\|^{2-n}$ and $\left| \frac{\partial u}{\partial \nu}(\mathbf{x}) \right| \leq C_2 \|\mathbf{x}\|^{1-n}$ as $\|\mathbf{x}\| \rightarrow \infty$ for some constants C_1 and C_2 . Hence,

$$\left| \int_{\partial B_r(\mathbf{0})} u \frac{\partial u}{\partial \nu} \, dS \right| \leq \int_{\partial B_r(\mathbf{0})} |u| \left| \frac{\partial u}{\partial \nu} \right| \, dS \leq C_1 C_2 r^{3-2n} \int_{\partial B_r(\mathbf{0})} dS = C_1 C_2 \omega_n r^{2-n} \rightarrow 0$$

as $r \rightarrow \infty$. If $n = 2$, we have from Proposition 9.9.5 and Proposition 9.9.6 that $|u(\mathbf{x})| \leq C_3 |\ln(\|\mathbf{x}\|)|$ and $\left| \frac{\partial u}{\partial \nu}(\mathbf{x}) \right| \leq C_4 \|\mathbf{x}\|^{-2}$ as $\|\mathbf{x}\| \rightarrow \infty$ for some constants C_3 and C_4 . Hence,

$$\left| \int_{\partial B_r(\mathbf{0})} u \frac{\partial u}{\partial \nu} \, dS \right| \leq \int_{\partial B_r(\mathbf{0})} |u| \left| \frac{\partial u}{\partial \nu} \right| \, dS \leq C_3 C_4 |\ln(r)| r^{-2} \int_{\partial B_r(\mathbf{0})} dS = C_3 C_4 \omega_2 \frac{|\ln(r)|}{r} \rightarrow 0$$

as $r \rightarrow \infty$.

Therefore, if we let $r \rightarrow \infty$ in (9.9.47), we get from (9.9.45) and (9.9.46) that $\int_{\Omega'} \|\nabla u\|_2^2 \, d\mathbf{x} = 0$. Thus $\nabla u = 0$ on Ω' . This shows that u is constant on each component of Ω' . From this, we get that $\phi = \frac{\partial u}{\partial \nu^+} = 0$.

ii) The proof that $L^2(\partial\Omega) = V_-^\perp \oplus W_-$ is very similar to the proof given in (i). It follows from Proposition 9.9.37 that $V_-^\perp$ is a closed subspace of codimension m' in $L^2(\partial\Omega)$ and W_- is a subspace of dimension m' in $L^2(\partial\Omega)$. Thus, we only have to prove that $V_-^\perp \cap W_- = \{0\}$.

Suppose that $\phi \in V_-^\perp \cap W_-$. Since $\phi \in W_-$, we have that $\phi \in \mathcal{N}\left(T_K^* + \frac{1}{2} \text{Id}\right)$. Since $\phi \in V_-^\perp = \left(\mathcal{N}\left(T_K + \frac{1}{2} \text{Id}\right)\right)^\perp$, it follows from Theorem 4.4.41 that $\phi \in \mathcal{R}\left(T_K^* + \frac{1}{2} \text{Id}\right)$ because T_K^* is a compact operator and so $\mathcal{R}\left(T_K^* + \frac{1}{2} \text{Id}\right)$ is a closed subspace according to Theorem 4.4.24. Thus, there exists $\psi \in L^2(\partial\Omega)$ such that $\phi = T_K^* \psi + \frac{1}{2} \psi$. Moreover, since $T_K^* \phi + \frac{1}{2} \phi = 0$, it follows from Proposition 9.9.20 that $\phi \in C(\partial\Omega)$. We can use Proposition 9.9.20 a second time to conclude that $\psi \in C(\partial\Omega)$ because $T_K^* \psi + \frac{1}{2} \psi = \phi$.

Let u and v be the simple layer potentials with moments ϕ and ψ respectively. We have that $u, v \in C(\mathbb{R}^n)$ according to Proposition 9.9.31.

From Theorem 9.9.33, we have that

$$\frac{\partial u}{\partial \nu^+} = T_K^* \phi + \frac{1}{2} \phi = 0 \quad \text{and} \quad \frac{\partial v}{\partial \nu^+} = T_K^* \psi + \frac{1}{2} \psi = \phi = -T_K^* \phi + \frac{1}{2} \phi = -\frac{\partial u}{\partial \nu^-}$$

on $\partial\Omega$. Hence,

$$\int_{\partial\Omega} \left(u \frac{\partial v}{\partial \nu^+} - v \frac{\partial u}{\partial \nu^+} \right) dS = - \int_{\partial\Omega} u \frac{\partial u}{\partial \nu^-} dS, \quad (9.9.48)$$

where the positive direction on $\partial\Omega$ is such that the normal to $\partial\Omega$ points outside Ω' , so inside Ω .

From the Green's identity (9.1.1), we get that

$$\int_{\partial\Omega} u \frac{\partial u}{\partial \nu^-} dS = - \int_{\Omega} (u \Delta u + \nabla u \cdot \nabla u) d\mathbf{x} = - \int_{\Omega} \|\nabla u\|_2^2 d\mathbf{x} \quad (9.9.49)$$

because u is harmonic on Ω . The normal to $\partial\Omega$ points inside Ω .

To compute the integral on the left side in (9.9.48), we choose $r > 0$ such that $\partial B_r(\mathbf{0}) \subset \Omega'_0$. We get from the Green's identities (9.1.2) that

$$\begin{aligned} \int_{\partial\Omega} \left(u \frac{\partial v}{\partial \nu^+} - v \frac{\partial u}{\partial \nu^+} \right) dS &= \int_{\Omega' \cap B_r(\mathbf{0})} (u \Delta v + v \Delta u) d\mathbf{x} - \int_{\partial B_r(\mathbf{0})} \left(u \frac{\partial v}{\partial \nu} - v \frac{\partial u}{\partial \nu} \right) dS \\ &= - \int_{\partial B_r(\mathbf{0})} \left(u \frac{\partial v}{\partial \nu} - v \frac{\partial u}{\partial \nu} \right) dS \end{aligned} \quad (9.9.50)$$

because u and v are harmonic in Ω' . The normal to $\partial(\Omega' \cap B_r(\mathbf{0}))$ points outside $\Omega' \cap B_r(\mathbf{0})$.

We now show that the integral in (9.9.50) converges to 0 as $r \rightarrow \infty$. We get from item (2) following Definition 9.9.30 that u and v are harmonic at infinity for $n > 2$. We now prove that u and v are harmonic at infinity for $n = 2$. Since $\phi \in W_-$, we get from Lemma 9.9.35 that $\int_{\partial\Omega} \phi dS = 0$. Moreover, since $T_K^* \psi + \frac{1}{2} \psi = \phi$, we can use Lemma 9.9.35 again to conclude that $\int_{\partial\Omega} \psi dS = 0$. It follows from Lemma 9.9.36 that u and v are harmonic at infinity.

Proceeding exactly as we did in (i), we can show that $\int_{\partial B_r(\mathbf{0})} u \frac{\partial v}{\partial \nu} dS$ and $\int_{\partial B_r(\mathbf{0})} v \frac{\partial u}{\partial \nu} dS$ converge to 0 as $r \rightarrow \infty$.

Therefore, if we let $r \rightarrow \infty$ in (9.9.50), we get from (9.9.48) and (9.9.49) that $\int_{\Omega} \|\nabla u\|_2^2 d\mathbf{x} = 0$. Thus $\nabla u = 0$ on Ω . This shows that u is constant on each component of Ω . From this, we that $\phi = \frac{\partial u}{\partial \nu^-} = 0$. ■

Corollary 9.9.39

$$L^2(\partial\Omega) = \mathcal{R} \left(-\frac{1}{2} \text{Id} + T_K \right) \oplus V_+ = \mathcal{R} \left(\frac{1}{2} \text{Id} + T_K \right) \oplus V_- .$$

Proof.

Since T_K is a compact operator, it follows from Theorems 4.4.24 and 4.4.41 that $\mathcal{R}\left(-\frac{1}{2}\text{Id} + T_K\right) = \left(\mathcal{N}\left(-\frac{1}{2}\text{Id} + T_K^*\right)\right)^\perp = W_+^\perp$ and $\mathcal{R}\left(\frac{1}{2}\text{Id} + T_K\right) = \left(\mathcal{N}\left(\frac{1}{2}\text{Id} + T_K^*\right)\right)^\perp = W_-^\perp$. It is therefore enough to show that $W_+^\perp \cap V_+ = W_-^\perp \cap V_- = \{0\}$ because $W_+^\perp$ has codimension m and V_+ has dimension m , and $W_-^\perp$ has codimension m' and V_- has dimension m' .

Suppose that $\psi \in W_+^\perp \cap V_+$. It follows from Proposition 9.9.38 that $\psi = \psi_1 + \psi_2$, where $\psi_1 \in V_+^\perp$ and $\psi_2 \in W_+$. Thus

$$\langle \psi, \psi_1 \rangle_2 = \int_{\partial\Omega} \psi(\mathbf{x}) \overline{\psi_1(\mathbf{x})} dS_{\mathbf{x}} = 0$$

because $\psi \in V_+$ and

$$\langle \psi, \psi_2 \rangle_2 = \int_{\partial\Omega} \psi(\mathbf{x}) \overline{\psi_2(\mathbf{x})} dS_{\mathbf{x}} = 0$$

because $\psi \in W_+^\perp$. Hence

$$\langle \psi, \psi \rangle_2 = \int_{\partial\Omega} |\psi(\mathbf{x})|^2 dS_{\mathbf{x}} = 0$$

by linearity. Thus, $\psi = 0$ almost everywhere on $\partial\Omega$. A similar proof yields $W_-^\perp \cap V_- = \{0\}$. ■

Theorem 9.9.40

Referring to the statement of the Dirichlet and Neumann problems given at the beginning of the chapter, we have the following results.

1. **(D1)** has a unique solution for $f \in C(\partial\Omega)$.
2. **(D2)** has a unique solution for $f \in C(\partial\Omega')$.
3. **(N1)** has a solution for $f \in C(\partial\Omega)$ if and only if $\int_{\partial\Omega_i} f(\mathbf{x}) dS_{\mathbf{x}} = 0$ for $1 \leq i \leq m$.
The solution is unique modulo a function which is constant on each connected component of Ω .
4. **(N2)** has a solution for $f \in C(\partial\Omega')$ if and only if $\int_{\partial\Omega'_j} f(\mathbf{x}) dS_{\mathbf{x}} = 0$ for $1 \leq j \leq m'$ when $n > 2$, and $0 \leq j \leq m'$ when $n = 2$. The solution is unique modulo a function h such that $h(\mathbf{x}) = c$, a fixed constant, on each connected component Ω'_j for $1 \leq j \leq m'$.

Proof.

Uniqueness results has been proved in In Section 9.9.1. The need for the necessary conditions is proved in Section 9.9.2. We only need to consider the existence of the solution.

N1) Let $\phi_i \in V_+$ for $1 \leq i \leq m$ be the functions defined in (9.9.44). We have shown in the proof of Proposition 9.9.37 that $\{\phi_i\}_{i=1}^m$ is a basis of V_+ . Hence,

$$V_+^\perp = \left\{ g \in L^2(\partial\Omega) : \int_{\partial\Omega} g(\mathbf{x}) \phi_i(\mathbf{x}) dS = \int_{\partial\Omega_i} g(\mathbf{x}) dS = 0 \text{ for } 1 \leq i \leq m \right\}.$$

From Theorems 4.4.24 and 4.4.41, we have that $g \in \mathcal{R}\left(-\frac{1}{2}\text{Id} + T_K^*\right)$ if and only if $g \in \left(\mathcal{N}\left(-\frac{1}{2}\text{Id} + T_K\right)\right)^\perp = V_+^\perp$.

By hypothesis $f \in V_+^\perp$. Therefore, there exists $\phi \in L^2(\partial\Omega)$ such that

$$f = -\frac{1}{2}\phi + T_K^*\phi .$$

Since $f \in C(\partial\Omega)$, we have that $\phi \in C(\partial\Omega)$ by Proposition 9.9.20. Finally, from Theorem 9.9.33, the single layer potential with moment ϕ defined by

$$u(\mathbf{x}) = \int_{\partial\Omega} N(\mathbf{x}, \mathbf{y})\phi(\mathbf{y}) \, dS_{\mathbf{y}} \quad , \quad \mathbf{x} \in \overline{\Omega} ,$$

is the solution of (N1).

N2) Let $\phi'_j \in V_-$ for $1 \leq j \leq m'$ be the functions defined in (9.9.44). We have shown in the proof of Proposition 9.9.37 that $\{\phi'_j\}_{j=1}^{m'}$ is a basis of V_- . Hence,

$$V_-^\perp = \left\{ g \in L^2(\partial\Omega) : \int_{\partial\Omega'} g(\mathbf{x})\phi'_j(\mathbf{x}) \, dS_{\mathbf{x}} = \int_{\partial\Omega'_j} g(\mathbf{x}) \, dS_{\mathbf{x}} = 0 \text{ for } 1 \leq j \leq m' \right\} .$$

From Theorems 4.4.24 and 4.4.41, we have that $g \in \mathcal{R}\left(\frac{1}{2}\text{Id} + T_K^*\right)$ if and only if $g \in \left(\mathcal{N}\left(\frac{1}{2}\text{Id} + T_K\right)\right)^\perp = V_-^\perp$.

By hypothesis $f \in V_-^\perp$. Therefore, there exists $\phi \in L^2(\partial\Omega)$ such that

$$f = \frac{1}{2}\phi + T_K^*\phi .$$

Since $f \in C(\partial\Omega')$, we have that $\phi \in C(\partial\Omega')$ by Proposition 9.9.20. Finally, from Theorem 9.9.33, the single layer potential with moment ϕ defined by

$$u(\mathbf{x}) = \int_{\partial\Omega} N(\mathbf{x}, \mathbf{y})\phi(\mathbf{y}) \, dS_{\mathbf{y}} \quad , \quad \mathbf{x} \in \overline{\Omega'} ,$$

is the solution of (N2). If $n > 2$, we showed in item (2) following Definition 9.9.30 that u is harmonica at infinity. If $n = 2$, the extra condition $\int_{\partial\Omega'_0} f(\mathbf{x}) \, dS_{\mathbf{x}} = 0$ implies that u is harmonic at infinity (and vanishes at infinity) according to Lemmas 9.9.35 and 9.9.36.

D1) From Corollary 9.9.39, we may write

$$f = \frac{1}{2}\phi + T_K\phi + \sum_{j=1}^{m'} a_j \phi'_j$$

for some $\phi \in L^2(\partial\Omega)$ and $a_j \in \mathbb{C}$. Since $f \in C(\partial\Omega)$ and $\phi'_j \in C(\partial\Omega)$ for all j , we have that $\phi + 2T_K\phi = 2f - 2\sum_{j=1}^{m'} a_j\phi'_j \in C(\partial\Omega)$. It follows from Proposition 9.9.20 that $\phi \in C(\partial\Omega)$. From Theorem 9.9.29, the double layer potential with moment ϕ ,

$$v(\mathbf{x}) = \int_{\partial\Omega} \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) \phi(\mathbf{y}) \, dS_{\mathbf{y}} \quad , \quad \mathbf{x} \in \Omega \quad ,$$

can be extended to $\overline{\Omega}$ to give a continuous solution for the interior Dirichlet problem (D1) with $\frac{1}{2}\phi + T_K\phi$ instead of f as boundary condition. By item (3) of Proposition 9.9.37, there exists a unique $\psi \in W_-$ such that the single layer potential with moment ψ ,

$$w(\mathbf{x}) = \int_{\partial\Omega} N(\mathbf{x}, \mathbf{y}) \psi(\mathbf{y}) \, dS_{\mathbf{y}} \quad , \quad \mathbf{x} \in \mathbb{R}^n \quad ,$$

satisfies $w|_{\Omega'_j} = a_j$ for $1 \leq j \leq m'$, and $w|_{\Omega'_0} = 0$. Since $w \in C(\mathbb{R}^n)$ by Proposition 9.9.31, we have that $w|_{\partial\Omega} = \sum_{j=1}^{m'} a_j\phi'_j$. Hence,

$$u(\mathbf{x}) = v(\mathbf{x}) + w(\mathbf{x}) \quad , \quad \mathbf{x} \in \overline{\Omega} \quad ,$$

is a solution of (D1).

D2) From Proposition 9.9.39, we may write

$$f = -\frac{1}{2}\phi + T_K\phi + \sum_{i=1}^m a_i\phi_i$$

for some $\phi \in L^2(\partial\Omega)$ and $a_i \in \mathbb{C}$. Since $f \in C(\partial\Omega)$ and $\phi_i \in C(\partial\Omega)$ for all j , we have that $\phi - 2T_K\phi = -2f + 2\sum_{j=1}^m a_j\phi_j \in C(\partial\Omega)$. It follows from Proposition 9.9.20 that $\phi \in C(\partial\Omega)$. From Theorem 9.9.29, the double layer potential with moment ϕ

$$v(\mathbf{x}) = \int_{\partial\Omega} \frac{\partial N}{\partial \nu_{\mathbf{y}}}(\mathbf{x}, \mathbf{y}) \phi(\mathbf{y}) \, dS_{\mathbf{y}} \quad , \quad \mathbf{x} \in \Omega' \quad ,$$

can be extended to $\overline{\Omega'}$ to give a continuous solution of the exterior Dirichlet problem (D2) with $-\frac{1}{2}\phi + T_K\phi$ instead of f as boundary condition.

The rest of the proof is split in two cases: $n > 2$ and $n = 2$.

n > 2) By item (4) of Proposition 9.9.37, there exists a unique $\psi \in W_+$ such that the single layer potential with moment ψ ,

$$w(\mathbf{x}) = \int_{\partial\Omega} N(\mathbf{x}, \mathbf{y}) \psi(\mathbf{y}) \, dS_{\mathbf{y}} \quad , \quad \mathbf{x} \in \mathbb{R}^n \quad ,$$

satisfies $w|_{\Omega_i} = a_i$ for $1 \leq i \leq m$. Since $w \in C(\mathbb{R}^n)$ by Proposition 9.9.31, $w|_{\partial\Omega'} = \sum_{i=1}^{\infty} a_i \phi_i$. Therefore,

$$u(\mathbf{x}) = v(\mathbf{x}) + w(\mathbf{x}) \quad , \quad \mathbf{x} \in \overline{\Omega'} ,$$

is a solution of (D2).

n = 2) Since $\sum_{i=1}^m \phi_i(\mathbf{x}) = 1$ for $\mathbf{x} \in \partial\Omega$, we may write

$$\sum_{i=1}^m a_i \phi_i(\mathbf{x}) = \sum_{i=1}^m b_i \phi_i(\mathbf{x}) + c ,$$

where c is a constant and $(b_1, b_2, \dots, b_m)$ is an element of the set X defined in Proposition 9.9.37. Hence, by item (5) of Proposition 9.9.37, there exists $\psi \in W_+^0$ such that the single layer potential with moment ψ ,

$$w(\mathbf{x}) = \int_{\partial\Omega} N(\mathbf{x}, \mathbf{y}) \psi(\mathbf{y}) \, dS_{\mathbf{y}} \quad , \quad \mathbf{x} \in \mathbb{R}^n ,$$

satisfies $w|_{\Omega_i} = b_i$ for $1 \leq i \leq m$. Since $\int_{\partial\Omega} \psi(\mathbf{y}) \, dS_{\mathbf{y}} = 0$ because $\psi \in W_+^0$, we also have from Lemma 9.9.36 that w is harmonic at infinity. Since $w \in C(\mathbb{R}^n)$ by Proposition 9.9.31, $w|_{\partial\Omega'} = \sum_{i=1}^m b_i \phi_i$. Therefore,

$$u(\mathbf{x}) = v(\mathbf{x}) + w(\mathbf{x}) + c \quad , \quad \mathbf{x} \in \overline{\Omega'}$$

is a solution of (D2). ■

9.10 Separation of Variables

The method of separation of variables is a simple method to solve the Laplace equation however it applies only to nice domains like a rectangle or a sphere. As for the wave equation, our presentation of the method of separation of variables for the Laplace equation is going to be formal.

9.10.1 Cartesian Coordinates

Example 9.10.1

Find the solution of the Laplace equation

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \quad , \quad 0 < x < a \text{ and } 0 < y < b , \quad (9.10.1)$$

with the boundary conditions $u(0, y) = 0$, $u(a, y) = 0$, $u(x, 0) = 0$ and $u(x, b) = f(x)$ for $0 < x < a$

Separation of Variables: If we substitute $u(x, y) = F(x)G(y)$ in (9.10.1), we get

$$F''(x)G(y) + F(x)G''(y) = 0 \quad , \quad 0 < x < a \text{ and } 0 < y < b \quad ,$$

After dividing both sides by $F(x)G(y)$, we get

$$\frac{F''(x)}{F(x)} + \frac{G''(y)}{G(y)} = 0 \quad , \quad 0 < x < a \text{ and } 0 < y < b \quad ,$$

Thus,

$$\frac{F''(x)}{F(x)} = -\frac{G''(y)}{G(y)} \quad , \quad 0 < x < a \text{ and } 0 < y < b \quad .$$

Since the right hand side is independent of x and the left hand side is independent of y , we get

$$\frac{F''(x)}{F(x)} = -\frac{G''(y)}{G(y)} = -k \quad , \quad 0 < x < a \text{ and } 0 < y < b \quad ,$$

for some constant k . We end up with two ordinary differential equations.

$$F''(x) + kF(x) = 0 \quad \text{and} \quad G''(y) - kG(y) = 0 \quad . \quad (9.10.2)$$

The first ordinary differential equation in (9.10.2) satisfies two boundary conditions. From $u(0, y) = 0$, we get $F(0)G(y) = 0$. Since we assume that G is non-null, we get $F(0) = 0$. Similarly, from $u(a, y) = 0$, we get $F(a)G(y) = 0$. Again, since we assume that G is non-null, we get $F(a) = 0$. The boundary conditions for the first ordinary differential equation in (9.10.2) are $F(0) = F(a) = 0$.

Simple Functions: We consider the boundary value problem

$$F''(x) + kF(x) = 0 \quad \text{with} \quad F(0) = F(a) = 0 \quad . \quad (9.10.3)$$

The form of the general solution of the ordinary differential equation above is determined by the roots of the characteristic equation $\lambda^2 + k = 0$.

If $k < 0$, the roots of the characteristic equation are $\pm\sqrt{-k}$. Since the roots are real, the solution of the ordinary differential equation is of the form $F(x) = Ae^{\sqrt{-k}x} + Be^{-\sqrt{-k}x}$. Moreover, $A + B = 0$ from $F(0) = 0$ and $Ae^{a\sqrt{-k}} + Be^{-a\sqrt{-k}} = 0$ from $F(a) = 0$. The only solution of the system

$$\begin{pmatrix} 1 & 1 \\ e^{a\sqrt{-k}} & e^{-a\sqrt{-k}} \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

is $A = B = 0$ because

$$\det \begin{pmatrix} 1 & 1 \\ e^{a\sqrt{-k}} & e^{-a\sqrt{-k}} \end{pmatrix} = e^{a\sqrt{-k}} - e^{-a\sqrt{-k}} = e^{-a\sqrt{-k}} (e^{2a\sqrt{-k}} - 1) \neq 0$$

for $k \neq 0$. Therefore, the null solution is the only solution of the boundary value problem (9.10.3) for $k < 0$.

If $k = 0$, the solution of the ordinary differential equation is $F(x) = Ax + B$. Moreover, $F(0) = 0$ implies that $B = 0$ and thus $F(a) = 0$ implies that $Aa + B = 0$. Hence, $A = B = 0$ and again the null solution is the only solution of the boundary value problem (9.10.3) for $k = 0$.

If $k > 0$, the roots of the characteristic equation are $\pm i\sqrt{k}$. Since the roots are complex, the solution of the ordinary differential equation is of the form

$$F(x) = A \cos(\sqrt{k}x) + B \sin(\sqrt{k}x) .$$

Moreover, $F(0) = 0$ implies that $A = 0$ and thus $F(a) = 0$ implies that $B \sin(a\sqrt{k}) = 0$. If we take $B = 0$, we get the null solution. We must therefore have $\sin(a\sqrt{k}) = 0$. This implies that $k = k_n \equiv (n\pi/a)^2$ for any integer $n > 0$. The boundary value problem (9.10.3) has non-null solutions only for $k = k_n = (n\pi/a)^2$ with any integer $n > 0$, and the solutions associated to k_n are of the form

$$F(x) = F_n(x) \equiv B_n \sin\left(\frac{n\pi x}{a}\right) .$$

With $k = k_n = (n\pi/a)^2$, the second ordinary differential equation in (9.10.3) is therefore

$$G''(y) - \left(\frac{n\pi}{a}\right)^2 G(y) = 0$$

for any integer $n > 0$. For each n , the general solution of this second order ordinary differential equation is

$$G(t) = G_n(t) \equiv C_n e^{n\pi y/a} + D_n e^{-n\pi y/a} .$$

We have found the simple functions

$$u_n(x, t) \equiv F_n(x)G_n(t) = (a_n e^{n\pi y/a} + b_n e^{-n\pi y/a}) \sin\left(\frac{n\pi x}{a}\right)$$

for $n > 0$. The constants a_n and b_n come from the product of the constants B_n and C_n , and B_n and D_n respectively.

We seek a solution of the form

$$u(x, y) = \sum_{n=1}^{\infty} u_n(x, y) = \sum_{n=1}^{\infty} (a_n e^{n\pi y/a} + b_n e^{-n\pi y/a}) \sin\left(\frac{n\pi x}{a}\right) .$$

From $u(x, 0) = 0$, we get

$$0 = \sum_{n=1}^{\infty} (a_n + b_n) \sin\left(\frac{n\pi x}{a}\right) .$$

Thus, $a_n = -b_n$ for all n . The solution can be written

$$u(x, t) = \sum_{n=1}^{\infty} c_n \sinh\left(\frac{n\pi y}{a}\right) \sin\left(\frac{n\pi x}{a}\right) , \quad (9.10.4)$$

where $c_n = 2a_n$. From $u(x, b) = f(x)$, we get

$$f(x) = \sum_{n=1}^{\infty} c_n \sinh\left(\frac{n\pi b}{a}\right) \sin\left(\frac{n\pi x}{a}\right)$$

for $0 < x < a$. This is the Fourier sine series of f on $[0, a]$. Thus,

$$c_n \sinh\left(\frac{n\pi b}{a}\right) = \frac{2}{a} \int_0^a f(x) \sin\left(\frac{n\pi x}{a}\right) dx$$

for $n > 0$. The solution of the Laplace equation is given by (9.10.4) with the coefficient

$$c_n = \frac{2}{a \sinh(n\pi b/a)} \int_0^a f(x) \sin\left(\frac{n\pi x}{a}\right) dx$$

for $n > 0$. ♣

In general, the solutions of the Laplace equation $\Delta u = 0$ on a rectangular domain are linear combinations of simple functions of the following form. If $u = 0$ on the vertical sides (x is constant),

$$u(x, y) = (A \cos(kx) + B \sin(kx))(C e^{ky} + D e^{-ky}) .$$

If $u = 0$ on the horizontal sides (y is constant),

$$u(x, y) = (A e^{kx} + B e^{-kx})(C \cos(ky) + D \sin(ky)) .$$

Example 9.10.2

Solve the potential equation

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \quad , \quad 0 < x < 3 \text{ and } 0 < y < 7 ,$$

with the boundary conditions $u(x, 0) = u(x, 7) = 0$ for $0 < x < 3$, $u(0, y) = 3 \sin(4\pi y/7)$ for $0 < y < 7$, and $u(3, y) = y(7 - y)$ for $0 < y < 7$.

The simple solutions are of the form

$$u(x, y) = (A e^{kx} + B e^{-kx})(C \cos(ky) + D \sin(ky)) .$$

The boundary condition $u(x, 0) = 0$ gives $C(A e^{kx} + B e^{-kx}) = 0$ for $0 < x < 3$. Thus $C = 0$ because the expression between parentheses cannot be null for all x if we want to get a non-null solution. The boundary condition $u(x, 7) = 0$ then gives $D(A e^{kx} + B e^{-kx}) \sin(7k) = 0$. Since D and the expression between parentheses cannot be null for all x if we want to get a non-null solution, we get $\sin(7k) = 0$ and therefore $k = k_n \equiv n\pi/7$ for $n > 0$.

We end up with simple functions of the form

$$u_n(x, y) \equiv (a_n e^{n\pi x/7} + b_n e^{-n\pi x/7}) \sin(n\pi y/7)$$

for $n > 0$, where a_n and b_n come from the product of D and A , and D and B respectively for each value of n .

We seek a solution of the form

$$u(x, t) = \sum_{n=1}^{\infty} u_n(x, t) = \sum_{n=1}^{\infty} (a_n e^{n\pi x/7} + b_n e^{-n\pi x/7}) \sin\left(\frac{n\pi y}{7}\right).$$

From $u(0, y) = 3 \sin(4\pi y/7)$, we get

$$\sum_{n=1}^{\infty} (a_n + b_n) \sin\left(\frac{n\pi y}{7}\right) = 3 \sin\left(\frac{4\pi y}{7}\right).$$

Thus

$$a_n + b_n = \begin{cases} 3 & \text{for } n = 4 \\ 0 & \text{otherwise} \end{cases}$$

From $u(3, y) = y(7 - y)$, we get

$$\sum_{n=1}^{\infty} (a_n e^{3n\pi/7} + b_n e^{-3n\pi/7}) \sin\left(\frac{n\pi y}{7}\right) = y(7 - y).$$

This is the Fourier sine series of $y(7 - y)$ on the interval $[0, 7]$. Thus,

$$a_n e^{3n\pi/7} + b_n e^{-3n\pi/7} = \frac{2}{7} \int_0^7 y(7 - y) \sin\left(\frac{n\pi y}{7}\right) dy.$$

To evaluate the integral, we use the method of integration by parts. The integrand is $f(y)g'(y)$, where $f(y) = y(7 - y)$ and $g'(y) = \sin(n\pi y/7)$. Hence, $f'(y) = 7 - 2y$ and $g(y) = -7/(n\pi) \cos(n\pi y/7)$. We get

$$\begin{aligned} a_n e^{3n\pi/7} + b_n e^{-3n\pi/7} &= \frac{-7}{n\pi} y(7 - y) \cos\left(\frac{n\pi y}{7}\right) \Big|_0^7 + \frac{7}{n\pi} \int_0^7 (7 - 2y) \cos\left(\frac{n\pi y}{7}\right) dy \\ &= \frac{7}{n\pi} \int_0^7 (7 - 2y) \cos\left(\frac{n\pi y}{7}\right) dy \end{aligned}$$

Again, we can compute this integral by the method of integration by parts. The integrand is $f(y)g'(y)$, where $f(y) = 7 - 2y$ and $g'(y) = \cos(n\pi y/7)$. Hence, $f'(y) = -2$ and $g(y) = 7/(n\pi) \sin(n\pi y/7)$. We get

$$\begin{aligned} a_n e^{3n\pi/7} + b_n e^{-3n\pi/7} &= \frac{7^2}{n^2\pi^2} (7 - 2y) \sin\left(\frac{n\pi y}{7}\right) \Big|_0^7 + \frac{2 \cdot 7^2}{n^2\pi^2} \int_0^7 \sin\left(\frac{n\pi y}{7}\right) dy \\ &= \frac{-2 \cdot 7^3}{n^3\pi^3} \cos\left(\frac{n\pi y}{7}\right) \Big|_0^7 = \frac{-2 \cdot 7^3}{n^3\pi^3} ((-1)^n - 1) \end{aligned}$$

for $n > 0$.

For $n = 4$, we have the linear system

$$a_4 + b_4 = 3$$

$$a_4 e^{12\pi/7} + b_4 e^{-12\pi/7} = 0$$

The solution of this system is $a_4 = \frac{-3e^{-12\pi/7}}{e^{12\pi/7} - e^{-12\pi/7}} = \frac{-3e^{-12\pi/7}}{2 \sinh(12\pi/7)}$ and $b_4 = \frac{3e^{12\pi/7}}{e^{12\pi/7} - e^{-12\pi/7}} = \frac{3e^{12\pi/7}}{2 \sinh(12\pi/7)}$.

For $n \neq 4$, we get the linear system

$$\begin{aligned} a_n + b_n &= 0 \\ a_n e^{3n\pi/7} + b_n e^{-3n\pi/7} &= \frac{-2 \cdot 7^3}{n^3 \pi^3} ((-1)^n - 1) \end{aligned}$$

The solution of this system is

$$a_n = -b_n = \frac{-2 \cdot 7^3}{n^3 \pi^3} ((-1)^n - 1) \frac{1}{e^{3n\pi/7} - e^{-3n\pi/7}} = \frac{2 \cdot 7^3 (1 - (-1)^n)}{n^3 \pi^3 \sinh(3n\pi/7)}.$$

The solution of the potential problem is

$$\begin{aligned} u(x, y) &= \frac{-3}{2 \sinh(12\pi/7)} (e^{-12\pi/7+4\pi x/7} - e^{12\pi/7-4\pi x/7}) \sin\left(\frac{4\pi y}{7}\right) \\ &\quad + \sum_{\substack{n=1 \\ n \neq 4}}^{\infty} \frac{2 \cdot 7^3 (1 - (-1)^n)}{n^3 \pi^3 \sinh(3n\pi/7)} (e^{n\pi x/7} - e^{-n\pi x/7}) \sin\left(\frac{n\pi y}{7}\right) \\ &= \frac{-3}{\sinh(12\pi/7)} \sinh\left(\frac{4\pi(x-3)}{7}\right) \sin\left(\frac{4\pi y}{7}\right) + \sum_{\substack{n=1 \\ n \neq 4}}^{\infty} \frac{7^3 (1 - (-1)^n)}{n^3 \pi^3 \sinh(3n\pi/7)} \sinh\left(\frac{n\pi x}{7}\right) \sin\left(\frac{n\pi y}{7}\right). \end{aligned}$$

♣

Example 9.10.3

Find the solution of the following Dirichlet problem.

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \quad , \quad 0 < y < L \text{ and } -L < x < L ,$$

with the boundary conditions $u(x, 0) = u(x, L) = 0$ for $-L < x < L$, $u(-L, y) = 50 \sin(\pi y/(2L))$ and $u(L, y) = 50 \sin(\pi y/(2L))$ for $0 < y < L$.

The simple functions obtained by separation of variables are of the form

$$u(x, y) = (Ae^{kx} + Be^{-kx}) (C \cos(ky) + D \sin(ky)) .$$

From $u(x, 0) = 0$, we get $C (Ae^{kx} + Be^{-kx}) = 0$ for $-L < x < L$. Since we are looking for non-null solutions, we must have $C = 0$ because $Ae^{kx} + Be^{-kx}$ cannot be equal to 0 for all x . From $u(x, L) = 0$, we then get $D \sin(kL) (Ae^{kx} + Be^{-kx}) = 0$ for $-L < x < L$. Hence, $D \neq 0$ and $\sin(kL) = 0$ since we are looking for non-null solutions. This implies that $k = k_n \equiv n\pi/L$ for any integer $n > 0$.

We have found that the simple functions are of the form

$$u_n(x, y) \equiv \sin\left(\frac{n\pi y}{L}\right) (a_n e^{n\pi x/L} + b_n e^{-n\pi x/L})$$

for $n > 0$, where a_n and b_n come from the product of D and A , and D and B respectively for each value of n . The solution is of the form

$$u(x, y) = \sum_{n=1}^{\infty} u_n(x, y) = \sum_{n=1}^{\infty} \sin\left(\frac{n\pi y}{L}\right) (a_n e^{n\pi x/L} + b_n e^{-n\pi x/L}) .$$

From $u(-L, y) = 50 \sin(\pi y/(2L))$, we get

$$\sum_{n=1}^{\infty} \sin\left(\frac{n\pi y}{L}\right) (a_n e^{-n\pi} + b_n e^{n\pi}) = 50 \sin\left(\frac{\pi y}{2L}\right) .$$

This is the Fourier sine series of the odd function $50 \sin(\pi y/(2L))$. Therefore,

$$a_n e^{-n\pi} + b_n e^{n\pi} = \frac{100}{L} \int_0^L \sin\left(\frac{n\pi y}{L}\right) \sin\left(\frac{\pi y}{2L}\right) dy .$$

If we use the trigonometric identity $\sin(\theta_1) \sin(\theta_2) = (\cos(\theta_1 - \theta_2) - \cos(\theta_1 + \theta_2))/2$, we get

$$\begin{aligned} a_n e^{-n\pi} + b_n e^{n\pi} &= \frac{50}{L} \int_0^L \left(\cos\left(\frac{n\pi y}{L} - \frac{\pi y}{2L}\right) - \cos\left(\frac{n\pi y}{L} + \frac{\pi y}{2L}\right) \right) dy \\ &= \frac{50}{L} \int_0^L \left(\cos\left(\frac{(2n-1)\pi y}{2L}\right) - \cos\left(\frac{(2n+1)\pi y}{2L}\right) \right) dy \\ &= \frac{100}{(2n-1)\pi} \sin\left(\frac{(2n-1)\pi y}{2L}\right) \Big|_0^L - \frac{100}{(2n+1)\pi} \sin\left(\frac{(2n+1)\pi y}{2L}\right) \Big|_0^L \\ &= \frac{100}{(2n-1)\pi} \sin\left(\frac{(2n-1)\pi}{2}\right) - \frac{100}{(2n+1)\pi} \sin\left(\frac{(2n+1)\pi}{2}\right) \\ &= \left(\frac{100}{(2n-1)\pi} + \frac{100}{(2n+1)\pi} \right) \sin\left(\frac{(2n-1)\pi}{2}\right) \end{aligned} \quad (9.10.5)$$

since

$$\sin\left(\frac{(2n+1)\pi}{2}\right) = \sin\left(\frac{(2n-1)\pi}{2} + \pi\right) = -\sin\left(\frac{(2n-1)\pi}{2}\right) ,$$

From $u(L, y) = 50 \sin(\pi y/(2L))$, we get

$$\sum_{n=1}^{\infty} \sin\left(\frac{n\pi y}{L}\right) (a_n e^{n\pi} + b_n e^{-n\pi}) = 50 \sin\left(\frac{\pi y}{2L}\right) .$$

As for the boundary condition at $x = -L$, this is the Fourier sine series of this odd function $50 \sin(\pi y/(2L))$. Therefore,

$$a_n e^{n\pi} + b_n e^{-n\pi} = \frac{100}{L} \int_0^L \sin\left(\frac{n\pi y}{L}\right) \sin\left(\frac{\pi y}{2L}\right) dy$$

$$= \left(\frac{100}{(2n-1)\pi} + \frac{100}{(2n+1)\pi} \right) \sin \left(\frac{(2n-1)\pi}{2} \right) \quad (9.10.6)$$

From (9.10.5) and (9.10.6), we get the linear system

$$\begin{pmatrix} e^{n\pi} & e^{-n\pi} \\ e^{-n\pi} & e^{n\pi} \end{pmatrix} \begin{pmatrix} a_n \\ b_n \end{pmatrix} = \begin{pmatrix} \alpha_n \\ \alpha_n \end{pmatrix},$$

where $\alpha_n = \left(\frac{100}{(2n-1)\pi} + \frac{100}{(2n+1)\pi} \right) \sin \left(\frac{(2n-1)\pi}{2} \right)$. The solution of this linear system is

$$\begin{aligned} a_n = b_n &= \frac{e^{n\pi} - e^{-n\pi}}{e^{2n\pi} - e^{-2n\pi}} \alpha_n = \frac{e^{n\pi} - e^{-n\pi}}{e^{2n\pi} - e^{-2n\pi}} \left(\frac{100}{(2n-1)\pi} + \frac{100}{(2n+1)\pi} \right) \sin \left(\frac{(2n-1)\pi}{2} \right) \\ &= (-1)^{n+1} \frac{\sinh(n\pi)}{\sinh(2n\pi)} \left(\frac{100}{(2n-1)\pi} + \frac{100}{(2n+1)\pi} \right), \end{aligned}$$

where we have used $\sin \left(\frac{(2n-1)\pi}{2} \right) = \sin \left(n\pi - \frac{\pi}{2} \right) = (-1)^{n+1}$.

Finally, the solution of the Dirichlet problem above is

$$\begin{aligned} u(x, y) &= \sum_{n=1}^{\infty} u_n(x, y) \\ &= \sum_{n=1}^{\infty} (-1)^{n+1} \frac{\sinh(n\pi)}{\sinh(2n\pi)} \left(\frac{100}{(2n-1)\pi} + \frac{100}{(2n+1)\pi} \right) \sin \left(\frac{n\pi y}{L} \right) (e^{n\pi x/L} + e^{-n\pi x/L}) \\ &= 2 \sum_{n=1}^{\infty} (-1)^{n+1} \frac{\sinh(n\pi)}{\sinh(2n\pi)} \left(\frac{200}{(2n-1)\pi} + \frac{200}{(2n+1)\pi} \right) \sin \left(\frac{n\pi y}{L} \right) \cosh \left(\frac{n\pi x}{L} \right). \end{aligned}$$

♣

Example 9.10.4

Find the solution of the following Dirichlet problem.

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \quad , \quad 0 < x < a \text{ and } 0 < y < b ,$$

with the boundary conditions $u(x, 0) = 0$ and $u(x, b) = 0.5 \sin(7\pi x/a)$ for $0 < x < a$, and $u(0, y) = 0.7 \sin(2\pi y/b)$ and $u(a, y) = y(b-y)$ for $0 < y < b$.

We split the problem into two smaller problems (Remark 9.10.5 below).

Problem 1: Solve the Laplace equation

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \quad , \quad 0 < x < a \text{ and } 0 < y < b ,$$

with the boundary conditions $u(x, 0) = 0$ and $u(x, b) = 0.5 \sin(7\pi x/a)$ for $0 < x < a$, and $u(0, y) = u(a, y) = 0$ for $0 < y < b$.

Problem 2: Solve the Laplace equation

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \quad , \quad 0 < x < a \text{ and } 0 < y < b \quad ,$$

with the boundary conditions $u(x, 0) = u(x, b) = 0$ for $0 < x < a$, and $u(0, y) = 0.7 \sin(2\pi y/b)$ and $u(a, y) = y(b - y)$ for $0 < y < b$.

Problem 1: The simple functions obtained by separation of variables are of the form

$$u(x, y) = (A \cos(kx) + B \sin(kx)) (Ce^{ky} + De^{-ky}) \quad .$$

From $u(0, y) = 0$, we get $A(Ce^{ky} + De^{-ky}) = 0$ for $0 < y < b$. Thus $A = 0$ because we are looking for non-null solutions. From $u(a, y) = 0$, we then get $B \sin(ka) (Ce^{ky} + De^{-ky}) = 0$ for $0 < y < b$. Hence, $B \neq 0$ and $\sin(ka) = 0$ because we are looking for non-null solutions. therefore, $k = k_n \equiv n\pi/a$ for any integer $n > 0$.

We have found that the simple functions are of the form

$$u_n(x, y) \equiv \sin\left(\frac{n\pi x}{a}\right) (a_n e^{n\pi y/a} + b_n e^{-n\pi y/a})$$

for $n > 0$. The solution is of the form

$$u(x, y) = \sum_{n=1}^{\infty} u_n(x, y) = \sum_{n=1}^{\infty} \sin\left(\frac{n\pi x}{a}\right) (a_n e^{n\pi y/a} + b_n e^{-n\pi y/a}) \quad .$$

From $u(x, 0) = 0$, we get

$$\sum_{n=1}^{\infty} \sin\left(\frac{n\pi x}{a}\right) (a_n + b_n) = 0 \quad .$$

Thus

$$a_n + b_n = 0 \tag{9.10.7}$$

for $n > 0$. From $u(x, b) = 0.5 \sin(7\pi x/a)$, we get

$$\sum_{n=1}^{\infty} \sin\left(\frac{n\pi x}{a}\right) (a_n e^{n\pi b/a} + b_n e^{-n\pi b/a}) = 0.5 \sin\left(\frac{7\pi x}{a}\right) \quad .$$

Thus

$$a_n e^{n\pi b/a} + b_n e^{-n\pi b/a} = \begin{cases} 0.5 & \text{if } n = 7 \\ 0 & \text{if } n \neq 7 \end{cases} \tag{9.10.8}$$

For $n = 7$, we get from (9.10.7) and (9.10.8) the linear system

$$\begin{aligned} a_7 + b_7 &= 0 \\ a_7 e^{7\pi b/a} + b_7 e^{-7\pi b/a} &= 0.5 \end{aligned}$$

The solution of this system is $a_7 = -b_7 = \frac{1}{2(e^{7\pi b/a} - e^{-7\pi b/a})} = \frac{1}{4 \sinh(7\pi b/a)}$.

For $n \neq 7$, we get from (9.10.7) and (9.10.8) the linear system

$$\begin{aligned} a_n + b_n &= 0 \\ a_n e^{n\pi b/a} + b_n e^{-n\pi b/a} &= 0 \end{aligned}$$

The solution of this system is $a_n = b_n = 0$ because

$$\det \begin{pmatrix} 1 & 1 \\ e^{n\pi b/a} & e^{-n\pi b/a} \end{pmatrix} = e^{-n\pi b/a} - e^{n\pi b/a} = e^{-n\pi b/a} (1 - e^{2n\pi b/a}) \neq 0$$

for $n \neq 0$. The solution of the first problem is

$$u(x, y) = \sin\left(\frac{7\pi x}{a}\right) \left(\frac{e^{7\pi y/a} - e^{-7\pi y/a}}{4 \sinh(7\pi b/a)} \right) = \frac{1}{2 \sinh(7\pi b/a)} \sin\left(\frac{7\pi x}{a}\right) \sinh\left(\frac{7\pi y}{a}\right) . \quad (9.10.9)$$

Problem 2: The simple functions obtained by separation of variables are of the form

$$u(x, y) = (Ae^{kx} + Be^{-kx})(C \cos(ky) + D \sin(ky)) .$$

From $u(x, 0) = 0$, we get $C(Ae^{kx} + Be^{-kx}) = 0$ for $0 < x < a$. Thus $C = 0$ because we are looking for non-null solutions. From $u(x, b) = 0$, we then get $D \sin(kb)(Ae^{kx} + Be^{-kx}) = 0$ for $0 < x < a$. Hence, $D \neq 0$ and $\sin(kb) = 0$ because we are looking for non-null solutions. Therefore, $k = k_n \equiv n\pi/b$ for any integer $n > 0$.

We have found that the simple functions are of the form

$$u_n(x, y) \equiv \sin\left(\frac{n\pi y}{b}\right) (a_n e^{n\pi x/b} + b_n e^{-n\pi x/b})$$

for $n > 0$. The solution is of the form

$$u(x, y) = \sum_{n=1}^{\infty} u_n(x, y) = \sum_{n=1}^{\infty} \sin\left(\frac{n\pi y}{b}\right) (A_n e^{n\pi x/b} + B_n e^{-n\pi x/b}) .$$

From $u(0, y) = 0.7 \sin(2\pi y/b)$, we get

$$\sum_{n=1}^{\infty} \sin\left(\frac{n\pi y}{b}\right) (a_n + b_n) = 0.7 \sin\left(\frac{2\pi y}{b}\right) .$$

Thus

$$a_n + b_n = \begin{cases} 0.7 & \text{if } n = 2 \\ 0 & \text{if } n \neq 2 \end{cases} \quad (9.10.10)$$

From $u(a, y) = y(b - y)$, we get

$$\sum_{n=1}^{\infty} \sin\left(\frac{n\pi y}{b}\right) (A_n e^{n\pi a/b} + B_n e^{-n\pi a/b}) = y(b - y) .$$

This is the Fourier sine series of $y(b - y)$, therefore

$$a_n e^{n\pi a/b} + b_n e^{-n\pi a/b} = \frac{2}{b} \int_0^b y(b - y) \sin\left(\frac{n\pi y}{b}\right) dy .$$

We can compute this integral by the method of integration by parts. The integrand is $f(y)g'(y)$, where $f(y) = y(b-y)$ and $g'(y) = \sin(n\pi y/b)$. Thus, $f'(y) = b-2y$ and $g(y) = -b/(n\pi) \cos(n\pi y/b)$. We get

$$\begin{aligned} a_n e^{n\pi a/b} + b_n e^{-n\pi a/b} &= -\frac{2}{b} \left(\frac{b}{n\pi} y(b-y) \cos\left(\frac{n\pi y}{b}\right) \right) \Big|_0^b + \frac{2}{n\pi} \int_0^b (b-2y) \cos\left(\frac{n\pi y}{b}\right) dy \\ &= \frac{2}{n\pi} \int_0^b (b-2y) \cos\left(\frac{n\pi y}{b}\right) dy. \end{aligned}$$

Again, we can compute this integral by the method of integration by parts. The integrand is $f(y)g'(y)$, where $f(y) = b-2y$ and $g'(y) = \cos(n\pi y/b)$. Thus, $f'(y) = -2$ and $g(y) = b/(n\pi) \sin(n\pi y/b)$. We get

$$\begin{aligned} a_n e^{n\pi a/b} + b_n e^{-n\pi a/b} &= \frac{2}{n\pi} \left(\frac{b}{n\pi} (b-2y) \sin\left(\frac{n\pi y}{b}\right) \right) \Big|_0^b + \frac{4b}{n^2\pi^2} \int_0^b \sin\left(\frac{n\pi y}{b}\right) dy \\ &= \frac{4b}{n^2\pi^2} \int_0^b \sin\left(\frac{n\pi y}{b}\right) dy = -\frac{4b^2}{n^3\pi^3} \cos\left(\frac{n\pi y}{b}\right) \Big|_0^b = \frac{4b^2}{n^3\pi^3} (1 - (-1)^n) \end{aligned} \quad (9.10.11)$$

for $n > 0$.

For $n = 2$, we get from (9.10.10) and (9.10.11) the linear system

$$\begin{aligned} a_2 + b_2 &= 0.7 \\ a_2 e^{2\pi a/b} + b_2 e^{-2\pi a/b} &= 0 \end{aligned}$$

The solution of this system is $a_2 = \frac{-0.7 e^{-2\pi a/b}}{2 \sinh(2\pi a/b)}$ and $b_2 = \frac{0.7 e^{2\pi a/b}}{2 \sinh(2\pi a/b)}$.

For $n \neq 2$, we get from (9.10.7) and (9.10.8) the linear system

$$\begin{aligned} a_n + b_n &= 0 \\ a_n e^{n\pi a/b} + b_n e^{-n\pi a/b} &= \frac{4b^2}{n^3\pi^3} (1 - (-1)^n) \end{aligned}$$

The solution of this system is

$$a_n = -b_n = \left(\frac{1}{e^{n\pi a/b} - e^{-n\pi a/b}} \right) \frac{4b^2}{n^3\pi^3} (1 - (-1)^n) = \frac{2b^2 (1 - (-1)^n)}{n^3\pi^3 \sinh(n\pi a/b)}.$$

The solution of the second problem is

$$\begin{aligned} u(x, y) &= \left(\frac{-0.7 e^{-2\pi a/b}}{2 \sinh(2\pi a/b)} e^{2\pi x/b} + \frac{0.7 e^{2\pi a/b}}{2 \sinh(2\pi a/b)} e^{-2\pi x/b} \right) \sin\left(\frac{2\pi y}{b}\right) \\ &\quad + \sum_{\substack{n=1 \\ n \neq 2}}^{\infty} \frac{2b^2 (1 - (-1)^n)}{n^3\pi^3 \sinh(n\pi a/b)} (e^{n\pi x/b} - e^{-n\pi x/b}) \sin\left(\frac{n\pi y}{b}\right) \\ &= \frac{-0.7}{\sinh(2\pi a/b)} \sinh\left(\frac{2\pi(x-a)}{b}\right) \sin\left(\frac{2\pi y}{b}\right) \end{aligned}$$

$$+ \sum_{\substack{n=1 \\ n \neq 2}}^{\infty} \frac{4b^2 (1 - (-1)^n)}{n^3 \pi^3 \sinh(n\pi a/b)} \sinh\left(\frac{n\pi x}{b}\right) \sin\left(\frac{n\pi y}{b}\right).$$

Since

$$1 - (-1)^n = \begin{cases} 2 & \text{for } n \text{ odd} \\ 0 & \text{for } n \text{ even} \end{cases}$$

we get the solution

$$\begin{aligned} u(x, y) = & \frac{-0.7}{\sinh(2\pi a/b)} \sin\left(\frac{2\pi y}{b}\right) \sinh\left(\frac{2\pi(x-a)}{b}\right) \\ & + \sum_{k=0}^{\infty} \frac{8b^2}{(2k+1)^3 \pi^3 \sinh((2k+1)\pi a/b)} \sin\left(\frac{(2k+1)\pi y}{b}\right) \sinh\left(\frac{(2k+1)\pi x}{b}\right). \end{aligned} \quad (9.10.12)$$

Complete Solution: The sum of (9.10.9) and (9.10.12) gives the solution of the full Dirichlet problem. ♣

Remark 9.10.5

Consider the Dirichlet problem

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \quad , \quad 0 < x < a \text{ and } 0 < y < b, \quad (9.10.13)$$

with the boundary conditions $u(x, 0) = f_1(x)$ and $u(x, b) = f_2(x)$ for $0 < x < a$, and $u(0, y) = g_1(y)$ and $u(a, y) = g_2(y)$ for $0 < y < b$.

Suppose that the compatibility conditions $f_1(0) = g_1(0)$, $f_1(a) = g_2(0)$, $f_2(0) = g_1(b)$ and $f_2(a) = g_2(b)$ are satisfied. Namely, the boundary conditions match at the corners. These compatibility conditions are necessary to get a classical solution.

If we split the problem in two as we did in Example 9.10.4, we get:

Problem 1: Solve

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \quad , \quad 0 < x < a \text{ and } 0 < y < b,$$

with the boundary conditions $u(x, 0) = f_1(x)$ and $u(x, b) = f_2(x)$ for $0 < x < a$, and $u(0, y) = 0$ and $u(a, y) = 0$ for $0 < y < b$.

Problem 2: Solve

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \quad , \quad 0 < x < a \text{ and } 0 < y < b,$$

with the boundary conditions $u(x, 0) = 0$ and $u(x, b) = 0$ for $0 < x < a$, and $u(0, y) = g_1(y)$ and $u(a, y) = g_2(y)$ for $0 < y < b$.

However, the boundary conditions may not match at the corners for the two reduced problems above. If $f_1(0) = g_1(0) \neq 0$, then the compatibility condition is broken at the $(0, 0)$

corner of the domain for both reduced problems. In Example 9.10.4, there was no such problem because $f_1(0) = f_1(a) = f_2(0) = f_2(a) = 0$ and $g_1(0) = g_1(b) = g_2(0) = g_2(b) = 0$.

To be able to split the problem and expect a classical solution, we must make sure that $f_1(0) = f_1(a) = f_2(0) = f_2(a) = 0$ and $g_1(0) = g_1(b) = g_2(0) = g_2(b) = 0$. To do so, we first choose a polynomial $p : \mathbb{R}^2 \rightarrow \mathbb{R}$ of degree two such that $\Delta p = 0$ with $p(0,0) = f_1(0) = g_1(0)$, $p(a,0) = f_1(a) = g_2(0)$, $p(0,b) = f_2(0) = g_1(b)$ and $p(a,b) = f_2(a) = g_2(b)$. We consider instead of (9.10.13) the Dirichlet problem

$$\Delta w = \frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} = 0 \quad , \quad 0 < x < a \text{ and } 0 < y < b \quad ,$$

with the boundary conditions $w(x,0) = f_1(x) - p(x,0)$ and $w(x,b) = f_2(x) - p(x,b)$ for $0 < x < a$, and $w(0,y) = g_1(y) - p(0,y)$ and $w(a,y) = g_2(y) - p(a,y)$ for $0 < y < b$.

For this problem, the boundary conditions match and are null at the corners. We may split this Dirichlet problem to find its solution w . The final solution of (9.10.13) is $u = w + p$.

Because $\Delta p = 0$, the polynomial p has to be of the form $p(x,y) = a_{2,2}(x^2 - y^2) + a_{1,1}xy + a_{1,0}x + a_{0,1}y + a_{0,0}$. To satisfy the conditions $p(0,0) = f_1(0) = g_1(0)$, $p(a,0) = f_1(a) = g_2(0)$, $p(0,b) = f_2(0) = g_1(b)$ and $p(a,b) = f_2(a) = g_2(b)$, we find that we may take $a_{2,2} = 0$ to get $a_{0,0} = f_1(0)$, $a_{1,0} = (f_1(a) - f_1(0))/a$, $a_{0,1} = (f_2(0) - f_1(0))/b$ and $a_{1,1} = (f_2(a) - f_1(a) - f_2(0) + f_1(0))/(ab)$. ♠

Example 9.10.6

Find the solution of the following Neumann problem.

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \quad , \quad 0 < y < \pi \text{ and } 0 < x < \pi \quad ,$$

with the boundary conditions $\frac{\partial u}{\partial x}(0,y) = \frac{\partial u}{\partial x}(\pi,y) = 0$ for $0 < y < \pi$, and $\frac{\partial u}{\partial y}(x,0) = 0$ and $\frac{\partial u}{\partial y}(x,\pi) = x - \frac{\pi}{2}$ for $0 < x < \pi$.

Note that the consistency condition for the Neumann problem, Proposition 9.9.10, is satisfied; namely,

$$\int_{\partial R} \frac{\partial u}{\partial \nu} ds = \int_0^\pi \frac{\partial u}{\partial y}(x,\pi) dx = \int_0^\pi \left(x - \frac{\pi}{2}\right) dx = \frac{1}{2} \left(x - \frac{\pi}{2}\right)^2 \Big|_0^\pi = 0 \quad , \quad (9.10.14)$$

where $R = \{(x,y) : 0 \leq y \leq \pi \text{ and } 0 \leq x \leq \pi\}$ and $\nu(x,y)$ is the outward unit normal to the surface ∂R at the point $(x,y) \in \partial R$ (except at the corners of the domain).

Separation of Variables: If we substitute $u(x,y) = F(x)G(y)$ in $\Delta u = 0$, we get

$$\frac{d^2 F}{dx^2}(x) G(y) + F(x) \frac{d^2 G}{dy^2}(y) = 0 \quad , \quad 0 < y < \pi \text{ and } 0 < x < \pi \quad .$$

Thus, after dividing both sides by $F(x)G(y)$, we get

$$\frac{1}{F(x)} \frac{d^2 F}{dx^2}(x) = -\frac{1}{G(y)} \frac{d^2 G}{dy^2}(y) \quad , \quad 0 < y < \pi \text{ and } 0 < x < \pi \quad .$$

Since the right hand side is independent of x and the left hand side is independent of y , we get

$$\frac{1}{F(x)} \frac{d^2 F}{dx^2}(x) = -\frac{1}{G(y)} \frac{d^2 G}{dy^2}(y) = k \quad , \quad 0 < y < \pi \text{ and } 0 < x < \pi .$$

for some constant k . We end up with two ordinary differential equation.

$$\frac{d^2 F}{dx^2}(x) - kF(x) = 0 \quad \text{and} \quad \frac{d^2 G}{dy^2}(y) + kG(y) = 0 . \quad (9.10.15)$$

Simple Functions: The first ordinary differential equation in (9.10.15) satisfies two boundary conditions. From $\frac{\partial u}{\partial x}(0, y) = 0$, we get $\frac{dF}{dx}(0)G(y) = 0$. Since we are looking for non-null solutions, G must be non-null and we get $\frac{dF}{dx}(0) = 0$. Similarly, from $\frac{\partial u}{\partial x}(\pi, y) = 0$, we get $\frac{dF}{dx}(\pi)G(y) = 0$. Again, since we are looking for non-null solutions, G must be non-null and we get $\frac{dF}{dx}(\pi) = 0$. The boundary conditions for the first ordinary differential equation in (9.10.15) are $\frac{dF}{dx}(0) = \frac{dF}{dx}(\pi) = 0$.

We consider the boundary value problem

$$\frac{d^2 F}{dx^2}(x) - kF(x) = 0 \quad , \quad 0 < x < \pi , \quad (9.10.16)$$

with the boundary conditions $\frac{dF}{dx}(0) = \frac{dF}{dx}(\pi) = 0$.

If $k > 0$, the roots of the characteristic equation are $\pm\sqrt{k}$. Since the roots are real, the general solution of the ordinary differential equation in (9.10.16) is of the form

$$F(x) = Ae^{\sqrt{k}x} + Be^{-\sqrt{k}x} .$$

However, $\frac{dF}{dx}(0) = 0$ implies that $A - B = 0$ and $\frac{dF}{dx}(\pi) = 0$ implies $Ae^{\pi\sqrt{k}} - Be^{-\pi\sqrt{k}} = 0$. The only solution of the system

$$\begin{pmatrix} 1 & -1 \\ e^{\pi\sqrt{k}} & -e^{-\pi\sqrt{k}} \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

is $A = B = 0$ because

$$\det \begin{pmatrix} 1 & -1 \\ e^{\pi\sqrt{k}} & -e^{-\pi\sqrt{k}} \end{pmatrix} = -e^{-\pi\sqrt{k}} + e^{\pi\sqrt{k}} = -e^{-\pi\sqrt{k}} (1 - e^{2\pi\sqrt{k}}) \neq 0$$

for $k \neq 0$. Therefore, the null solution is the only solution of the boundary value problem for $k > 0$.

If $k = 0$, the general solution of the ordinary differential equation in (9.10.16) is $F(x) = B_0x + A_0$. However, both $\frac{dF}{dx}(0) = 0$ and $\frac{dF}{dx}(\pi) = 0$ imply that $B_0 = 0$. Hence, A_0 is free. We get the solution $F_0(x) \equiv A_0$ for the boundary value problem.

If $k < 0$, the roots of the characteristic equation are $\pm i\sqrt{-k}$. Since the roots are complex, the general solution of the ordinary differential equation in (9.10.16) is of the form

$$F(x) = A \cos(\sqrt{-k}x) + B \sin(\sqrt{-k}x) .$$

However, we get $B = 0$ from $\frac{dF}{dx}(0) = 0$. It then follows from $\frac{dF}{dx}(\pi) = 0$ that $-A\sqrt{-k} \sin(\pi\sqrt{-k}) = 0$. If we take $A = 0$, we get the null solution. We must therefore have $\sin(\pi\sqrt{-k}) = 0$ with $k > 0$. This implies that $k = k_n \equiv -n^2$ for $n > 0$. The boundary value problem (9.10.16) has non-null solutions only for $k = k_n = -n^2 < 0$ with any integer $n > 0$, and the solutions associated are of the form $F(x) = F_n(x) \equiv A_n \cos(nx)$ for $n > 0$.

We consider the differential equation

$$\frac{d^2G}{dy^2}(y) + kG(y) = 0 .$$

For $k = 0$, we get the general solution $G_0(y) \equiv C_0y + D_0$ and, for $k = -n^2$ with $n > 0$, we get the general solution $G_n(y) \equiv C_n e^{ny} + D_n e^{-ny}$.

The simple functions obtained by separation of variables are $u_0(x, y) \equiv a_0y + b_0$ and $u_n(x, y) \equiv \cos(nx)(a_n e^{ny} + b_n e^{-ny})$ for all integers $n > 0$, where $a_n = C_n A_n$ and $b_n = D_n A_n$.

The solution is of the form

$$u(x, y) = \sum_{n=0}^{\infty} u_n(x, y) = a_0y + b_0 + \sum_{n=1}^{\infty} \cos(nx)(a_n e^{ny} + b_n e^{-ny}) .$$

Since

$$\frac{\partial u}{\partial y}(x, y) = a_0 + \sum_{n=1}^{\infty} \cos(nx)(na_n e^{ny} - nb_n e^{-ny}) ,$$

we get from $\frac{\partial u}{\partial y}(x, 0) = 0$ that

$$a_0 + \sum_{n=1}^{\infty} \cos(nx)(na_n - nb_n) = 0 .$$

Thus

$$a_0 = 0 \quad \text{and} \quad na_n - nb_n = 0 \tag{9.10.17}$$

for all integers $n > 0$. We also get from $\frac{\partial u}{\partial y}(x, \pi) = x - \frac{\pi}{2}$ that

$$a_0 + \sum_{n=1}^{\infty} \cos(nx)(na_n e^{n\pi} - nb_n e^{-n\pi}) = x - \frac{\pi}{2} .$$

This is the Fourier cosine series of $x - \pi/2$. Since

$$a_0 = \frac{1}{\pi} \int_0^{\pi} \left(x - \frac{\pi}{2}\right) dx = 0 ,$$

the constant term of the Fourier cosine series of $x - \pi/2$ is null. This is exactly what we have in (9.10.17). This is a concrete justification for the consistency condition (9.10.14). The other terms of the Fourier cosine series are given by

$$na_ne^{n\pi} - nb_ne^{-n\pi} = \frac{2}{\pi} \int_0^\pi \left(x - \frac{\pi}{2}\right) \cos(nx) \, dx$$

for $n > 0$. We use integration by parts to compute this last integral. The integrand is of the form $f(x)g'(x)$ where $f(x) = x - \pi/2$ and $g'(x) = \cos(nx)$. Thus $f'(x) = 1$ and $g(x) = (1/n)\sin(nx)$. We end up with

$$\begin{aligned} na_ne^{n\pi} - nb_ne^{-n\pi} &= \frac{2}{\pi} \left(\frac{1}{n} \left(x - \frac{\pi}{2}\right) \sin(nx) \Big|_0^\pi - \frac{1}{n} \int_0^\pi \sin(nx) \, dx \right) \\ &= \frac{2}{n^2\pi} \cos(nx) \Big|_0^\pi = \frac{2}{n^2\pi} ((-1)^n - 1) \end{aligned} \quad (9.10.18)$$

for $n > 0$.

For $n > 0$, we get the following system of linear equations from (9.10.17) and (9.10.18).

$$\begin{pmatrix} n & -n \\ ne^{n\pi} & -ne^{-n\pi} \end{pmatrix} \begin{pmatrix} a_n \\ b_n \end{pmatrix} = \begin{pmatrix} 0 \\ 2((-1)^n - 1)/(n^2\pi) \end{pmatrix}.$$

The solution of this system is

$$a_n = b_n = \frac{2((-1)^n - 1)}{n^3\pi(e^{n\pi} - e^{-n\pi})} = \frac{(-1)^n - 1}{n^3\pi \sinh(n\pi)} = \begin{cases} \frac{-2}{n^3\pi \sinh(n\pi)} & \text{for } n \text{ odd} \\ 0 & \text{for } n \text{ even} \end{cases}$$

The solution of the Neumann problem is

$$u(x, y) = b_0 - \sum_{k=1}^{\infty} \frac{4}{(2k-1)^3\pi \sinh((2k-1)\pi)} \cos((2k-1)x) \cosh((2k-1)y).$$

Since the boundary conditions involve only derivatives, it is not surprising that there are no constraints on b_0 . Namely, b_0 is a free parameter. ♣

9.10.2 Cylindrical and Spherical Coordinates

We have seen the polar, cylindrical and spherical changes of coordinates in Section 8.7.2. In particular, we have seen that the Laplacian in spherical coordinates is

$$\Delta u = \frac{\partial^2 u}{\partial r^2} + \frac{2}{r} \frac{\partial u}{\partial r} + \frac{1}{r^2} \frac{\partial^2 u}{\partial \phi^2} + \frac{\cot(\phi)}{r^2} \frac{\partial u}{\partial \phi} + \frac{1}{r^2 \sin^2(\phi)} \frac{\partial^2 u}{\partial \theta^2}.$$

Example 9.10.7

Solve the potential equation

$$\Delta u = 0 \quad , \quad \|(x, y, z)\|_2 < R$$

with the condition that u is bounded at the origin and u satisfies the boundary condition $u(R, \theta, \phi) = f(\phi)$ for $0 \leq \phi \leq \pi$.

Since the boundary condition is independent of θ , we may assume that the solution u is also independent of θ . Thus, the potential equation becomes

$$\Delta u = \frac{\partial^2 u}{\partial r^2} + \frac{2}{r} \frac{\partial u}{\partial r} + \frac{1}{r^2} \frac{\partial^2 u}{\partial \phi^2} + \frac{\cot(\phi)}{r^2} \frac{\partial u}{\partial \phi} = 0 \quad , \quad \|(x, y, z)\|_2 < R . \quad (9.10.19)$$

Separation of Variables: If we substitute $u(r, \phi) = F(r)G(\phi)$ in (9.10.19) and multiply by r^2 , we get

$$r^2 \frac{d^2 F}{dr^2}(r)G(\phi) + 2r \frac{dF}{dr}(r)G(\phi) + F(r) \frac{d^2 G}{d\phi^2}(\phi) + \cot(\phi)F(r) \frac{dG}{d\phi}(\phi) = 0 .$$

After dividing by $F(r)G(\phi)$, we get

$$\frac{1}{F(r)} \frac{d}{dr} \left(r^2 \frac{dF}{dr}(r) \right) = - \frac{1}{\sin(\phi)G(\phi)} \frac{d}{d\phi} \left(\sin(\phi) \frac{dG}{d\phi}(\phi) \right) .$$

Since the left hand side is independent of ϕ and the right hand side is independent of r , we get

$$\frac{1}{F(r)} \frac{d}{dr} \left(r^2 \frac{dF}{dr}(r) \right) = - \frac{1}{\sin(\phi)G(\phi)} \frac{d}{d\phi} \left(\sin(\phi) \frac{dG}{d\phi}(\phi) \right) = k ,$$

where k is constant. This yields two ordinary differential equations.

$$r^2 \frac{d^2 F}{dr^2}(r) + 2r \frac{dF}{dr}(r) - kF(r) = 0 \quad (9.10.20)$$

and

$$\frac{1}{\sin(\phi)} \frac{d}{d\phi} \left(\sin(\phi) \frac{dG}{d\phi}(\phi) \right) + kG(\phi) = 0 . \quad (9.10.21)$$

Simple Functions: We first solve (9.10.21). Let $w = \cos(\phi)$. We have $\frac{dH}{d\phi}(\phi) = \frac{dH}{dw}(w) \frac{dw}{d\phi}(\phi) = -\sin(\phi) \frac{dH}{dw}(w)$ and so $\frac{1}{\sin(\phi)} \frac{dH}{d\phi}(\phi) = -\frac{dH}{dw}(w)$ for any differentiable function H . As usual, to simplify the notation, we keep the same name H as a function of ϕ and H as a function of w . We then have $H(\phi) = H(w)$ for $w = \cos(\phi)$. Hence, for $H = G$, we get

$$\sin(\phi) \frac{dG}{d\phi}(\phi) = -\sin^2(\phi) \frac{dG}{dw}(w) = (\cos^2(\phi) - 1) \frac{dG}{dw}(w) = (w^2 - 1) \frac{dG}{dw}(w) .$$

With the change of variable $w = \cos(\phi)$, the equation (9.10.21) becomes

$$-\frac{d}{dw} \left((w^2 - 1) \frac{dG}{dw}(w) \right) + kG(w) = 0 .$$

If we expand this equation, we get the Legendre equation

$$(1-w^2)\frac{d^2G}{dw^2}(w) - 2w\frac{dG}{dw}(w) + kG(w) = 0$$

for $-1 \leq w \leq 1$ because $0 \leq \phi \leq \pi$. As we saw in Section 5.3 the Legendre equation has bounded solutions on $] -1, 1[$ only when $k = n(n+1)$. These solutions are scalar multiple of the Legendre polynomial $P_n(w)$. Thus, we take $G(\phi) = G_n(\phi) \equiv P_n(\cos(\phi))$ for integers $n \geq 0$.

We now turn our attention to (9.10.20) with $k = n(n+1)$. This is an Euler-Cauchy ordinary differential equation. The method to solve this ordinary differential equation is well known. If we substitute $F(r) = r^\alpha$ in (9.10.20), we get the characteristic equation $\alpha(\alpha-1) + 2\alpha + k = 0$. If α_1 and α_2 are distinct roots of this equation, then the general solution of (9.10.20) is given by $F(r) = Ar^{\alpha_1} + Br^{\alpha_2}$. If there is only one root α of algebraic multiplicity two, then the general solution of (9.10.20) is given by $F(r) = Ar^\alpha + Br^\alpha \ln(r)$.

If we substitute $k = n(n+1)$ in (9.10.20), we get the characteristic equation

$$\alpha(\alpha+1) - n(n+1) = (\alpha-n)(\alpha+n+1) = 0 .$$

Its roots are $\alpha_1 = n$ and $\alpha_2 = -n-1$. The general solution of (9.10.20) with $k = n(n+1)$ is therefore

$$F(r) = F_n(r) \equiv A_n r^n + B_n r^{-n-1} \quad (9.10.22)$$

for $n \geq 0$. Since we are looking for bounded solutions at the origin, we must have $B_n = 0$ for all n .

The simple functions are

$$u_n(r, \phi) \equiv F_n(r)G_n(\phi) = A_n r^n P_n(\cos(\phi))$$

for $n \geq 0$.

Boundary Condition: We seek a solution of the form

$$u(r, \phi) = \sum_{n=0}^{\infty} u_n(r, \phi) = \sum_{n=0}^{\infty} A_n r^n P_n(\cos(\phi)) \quad , \quad 0 < r < R \text{ and } 0 \leq \phi \leq \pi .$$

From the Boundary condition $u(R, \phi) = f(\phi)$, we get

$$f(\phi) = \sum_{n=0}^{\infty} A_n R^n P_n(\cos(\phi)) .$$

The change of variable $w = \cos(\phi)$ (or $\phi = \arccos(w)$) shows that this is the Fourier-Legendre series of $f(\arccos(w))$ with respect to the polynomials $P_n(w)$ for $n \geq 0$. Hence

$$A_n R^n = \frac{2n+1}{2} \int_{-1}^1 f(\arccos(w)) P_n(w) dw = \frac{2n+1}{2} \int_0^\pi f(\phi) P_n(\cos(\phi)) \sin(\phi) d\phi ,$$

Example 9.10.8

Solve the potential equation

$$\Delta u = 0 \quad , \quad \|(x, y, z)\|_2 > R$$

with the condition that u vanished at infinity and u satisfies the boundary condition $u(R, \theta, \phi) = f(\phi)$ for $0 \leq \phi \leq \pi$.

The solution of this problem starts exactly like in Example 9.10.7. The solution takes a different direction at (9.10.22). Since we are now looking for a solution that goes to 0 as r converges to infinity, we must take $A_n = 0$ for $n \geq 0$.

The simple functions are

$$u_n(r, \phi) \equiv F_n(r)G_n(\phi) = B_n r^{-n-1} P_n(\phi)$$

for $n \geq 0$.

Boundary Condition: We seek a solution of the form

$$u(r, \phi) = \sum_{n=0}^{\infty} u_n(r, \phi) = \sum_{n=0}^{\infty} B_n r^{-n-1} P_n(\phi) \quad , \quad r > R \text{ and } 0 \leq \phi \leq \pi .$$

From the Boundary condition $u(R, \phi) = f(\phi)$, we get

$$f(\phi) = \sum_{n=0}^{\infty} B_n R^{-n-1} P_n(\cos(\phi)) .$$

This is the Fourier-Legendre series of $f(\arccos(w))$ with respect to the polynomials $P_n(w)$ for $n \geq 0$. Therefore

$$B_n R^{-n-1} = \frac{2n+1}{2} \int_{-1}^1 f(\arccos(w)) P_n(w) dw = \frac{2n+1}{2} \int_0^\pi f(\phi) P_n(\cos(\phi)) \sin(\phi) d\phi ,$$

where we have used the change of variable $w = \cos(\phi)$ for $0 \leq \phi \leq \pi$. ♣

Example 9.10.9

Solve the potential equation

$$\Delta u = 0 \quad , \quad 0 < R_1 < \|(x, y, z)\|_2 < R_2$$

with the condition that u is bounded and the boundary conditions $u(R_1, \theta, \phi) = f(\phi)$ and $u(R_2, \theta, \phi) = g(\phi)$ for $0 \leq \phi \leq \pi$.

The solution of this problem starts exactly like in Example 9.10.7. The solution takes a different direction at (9.10.22). Since the domain is the bounded annulus $\{(x, y, z) : R_1 < \|(x, y, z)\|_2 < R_2\}$, there is no reason for the A_n and B_n to be zero.

The simple functions are

$$u_n(r, \phi) \equiv F_n(r)G_n(\phi) = (A_n r^n + B_n r^{-n-1}) P_n(\phi)$$

for $n \geq 0$.

Boundary Condition: We seek a solution of the form

$$u(r, \phi) = \sum_{n=0}^{\infty} u_n(r, \phi) = \sum_{n=0}^{\infty} (A_n r^n + B_n r^{-n-1}) P_n(\phi) \quad , \quad R_1 < r < R_2 \text{ and } 0 \leq \phi \leq \pi .$$

From the Boundary condition $u(R_1, \phi) = f(\phi)$, we get

$$f(\phi) = \sum_{n=0}^{\infty} (A_n R_1^n + B_n R_1^{-n-1}) P_n(\cos(\phi)) .$$

This is the Fourier-Legendre series of $f(\arccos(w))$ with respect to the polynomials $P_n(w)$ for $n \leq 0$. Therefore,

$$A_n R_1^n + B_n R_1^{-n-1} = \frac{2n+1}{2} \int_0^\pi f(\phi) P_n(\cos(\phi)) \sin(\phi) d\phi .$$

From the Boundary condition $u(R_2, \phi) = g(\phi)$, we get

$$g(\phi) = \sum_{n=0}^{\infty} (A_n R_2^n + B_n R_2^{-n-1}) P_n(\cos(\phi)) .$$

This is the Fourier-Legendre series of $g(\arccos(w))$ with respect to the polynomials $P_n(w)$ for $n \geq 0$. Therefore,

$$A_n R_2^n + B_n R_2^{-n-1} = \frac{2n+1}{2} \int_0^\pi g(\phi) P_n(\cos(\phi)) \sin(\phi) d\phi .$$

We have an infinite numbers of linear system of equations

$$\begin{pmatrix} R_1^n & R_1^{-n-1} \\ R_2^n & R_2^{-n-1} \end{pmatrix} \begin{pmatrix} A_n \\ B_n \end{pmatrix} = \begin{pmatrix} \frac{2n+1}{2} \int_0^\pi f(\phi) P_n(\cos(\phi)) \sin(\phi) d\phi \\ \frac{2n+1}{2} \int_0^\pi g(\phi) P_n(\cos(\phi)) \sin(\phi) d\phi \end{pmatrix}$$

for $n \geq 0$ to solve. ♣

By separation of variables, we find that the solutions of the potential equation on a spherical domain are linear combinations of simple functions of the following form. If the problem is independent of θ ,

$$u_n(r, \phi) = \left(A_n r^n + \frac{B_n}{r^{n+1}} \right) P_n(\cos(\phi)) \quad (9.10.23)$$

for $n \geq 0$. If the problem is independent of θ and ϕ , $u(r) = C/r + D$.

Example 9.10.10

Consider two concentric spheres S_1 and S_2 of radius $R_1 = 3$ and $R_2 = 4$ respectively. The potential on the first sphere is given by $u(R_1, \phi) = \cos(\phi)$ and on the second sphere by $u(R_2, \phi) = 3 \cos^2(\phi) + \cos(\phi) - 1$. Find the potential $u(r, \phi)$:

- a) inside the sphere S_1 .
- b) outside the sphere S_2
- c) between both spheres.

For (a), since the simple functions (9.10.23) must be bounded at the origine, we must have $B_n = 0$ for all n . The solution is therefore of the form

$$u(r, \phi) = \sum_{n=0}^{\infty} A_n r^n P_n(\cos(\phi)) \quad , \quad 0 < r < 3 \text{ and } 0 \leq \phi \leq \pi .$$

Since $P_1(\cos(\phi)) = \cos(\phi)$, we get

$$P_1(\cos(\phi)) = u(3, \phi) = \sum_{n=0}^{\infty} A_n 3^n P_n(\cos(\phi)) .$$

Thus

$$A_n = \begin{cases} 1/3 & \text{if } n = 1 \\ 0 & \text{if } n \neq 1 \end{cases}$$

The potential inside the sphere S_1 is given by

$$u(r, \phi) = \frac{1}{3} r P_1(\cos(\phi)) = \frac{r \cos(\phi)}{3} \quad , \quad 0 < r < 3 \text{ and } 0 \leq \phi \leq \pi .$$

For (b), since the simple functions (9.10.23) must converge to 0 as r goes to infinity, we must have $A_n = 0$ for all n . The solution is therefore of the form

$$u(r, \phi) = \sum_{n=0}^{\infty} \frac{B_n}{r^{n+1}} P_n(\cos(\phi)) \quad , \quad r > 4 \text{ and } 0 \leq \phi \leq \pi .$$

Since

$$u(4, \phi) = 3 \cos^2(\phi) + \cos(\phi) - 1 = 2P_2(\cos(\phi)) + P_1(\cos(\phi)) ,$$

we get

$$2P_2(\cos(\phi)) + P_1(\cos(\phi)) = u(4, \phi) = \sum_{n=0}^{\infty} \frac{B_n}{4^{n+1}} P_n(\cos(\phi)) .$$

Thus

$$B_n = \begin{cases} 16 & \text{if } n = 1 \\ 128 & \text{if } n = 2 \\ 0 & \text{if } n \neq 1, 2 \end{cases}$$

The potential outside the sphere S_2 is given by

$$u(r, \phi) = \frac{16}{r^2} P_1(\cos(\phi)) + \frac{128}{r^3} P_2(\cos(\phi)) \quad , \quad r > 4 \text{ and } 0 \leq \phi \leq \pi .$$

Finally, for (c), there is no restriction on the simple functions (9.10.23), The solution is therefore of the form

$$u(r, \phi) = \sum_{n=0}^{\infty} \left(A_n r^n + \frac{B_n}{r^{n+1}} \right) P_n(\cos(\phi)) \quad , \quad 3 < r < 4 \text{ and } 0 \leq \phi \leq \pi .$$

From

$$P_1(\cos(\phi)) = u(3, \phi) = \sum_{n=0}^{\infty} \left(A_n 3^n + \frac{B_n}{3^{n+1}} \right) P_n(\cos(\phi)) ,$$

we get

$$A_n 3^n + \frac{B_n}{3^{n+1}} = \begin{cases} 1 & \text{if } n = 1 \\ 0 & \text{if } n \neq 1 \end{cases}$$

From

$$2P_2(\cos(\phi)) + P_1(\cos(\phi)) = u(4, \phi) = \sum_{n=0}^{\infty} \left(A_n 4^n + \frac{B_n}{4^{n+1}} \right) P_n(\cos(\phi)) ,$$

we get

$$A_n 4^n + \frac{B_n}{4^{n+1}} = \begin{cases} 1 & \text{if } n = 1 \\ 2 & \text{if } n = 2 \\ 0 & \text{if } n \neq 1, 2 \end{cases}$$

For $n = 1$, the coefficients A_1 and B_1 are the solutions of

$$\begin{pmatrix} 3 & 1/9 \\ 4 & 1/16 \end{pmatrix} \begin{pmatrix} A_1 \\ B_1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} .$$

We find $A_1 = 7/37$ and $B_1 = 144/37$. For $n = 2$, the coefficients A_2 and B_2 are the solutions of

$$\begin{pmatrix} 9 & 1/27 \\ 16 & 1/64 \end{pmatrix} \begin{pmatrix} A_2 \\ B_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \end{pmatrix} .$$

We find $A_2 = -2(27(9/64 - 16/27))^{-1} = 128/781$ and

$B_2 = 18(9/64 - 16/27)^{-1} = -31104/781$. For all the other values of n , $A_n = B_n = 0$. The potential between the spheres S_1 and S_2 is given by

$$u(r, \phi) = \left(\frac{7r}{37} + \frac{144}{37r^2} \right) P_1(\cos(\phi)) + \left(\frac{128r^2}{781} + \frac{-31104}{781r^3} \right) P_2(\cos(\phi))$$

for $3 < r < 4$ and $0 \leq \phi \leq \pi$. ♣

We have seen in Section 8.7.2 that the Laplacian in cylindrical coordinates is

$$\Delta u = \frac{\partial^2 u}{\partial r^2} + \frac{1}{r} \frac{\partial u}{\partial r} + \frac{1}{r^2} \frac{\partial^2 u}{\partial \theta^2} + \frac{\partial^2 u}{\partial z^2} .$$

By separation of variables, we find that the solutions of the potential equation in a cylindrical domain (this includes the planar circular domain if the solution is independent of the z -coordinate) are linear combinations of simple functions of the following form. If the problem is independent of z only $u_0(r, \theta) = A_0 + B_0 \ln(r)$ and $u_n(r, \theta) = (A_n r^n + B_n r^{-n}) \cos(n\theta) + (C_n r^n + D_n r^{-n}) \sin(n\theta)$ for $n > 0$. If the problem is independent of θ and z , $u(r) = C \ln r + D$.

9.11 Eigenvalue Problems for the Laplacian Operator

The concept of eigenvalues and eigenvectors is first introduced in Linear Algebra. Namely, if A is an $n \times n$ matrix, then λ is an eigenvalue of A if there exists a non-null vector $\mathbf{v} \in \mathbb{R}^n$ such that $A\mathbf{v} = \lambda\mathbf{v}$. The vector $\mathbf{v}$ is called an eigenvector associated to λ .

The same concept may be applied to the linear differential operator Δ . Suppose that Ω is a open set in $\mathbb{R}^n$ with a sufficiently smooth boundary. Let

$$\mathcal{X} = \{u \in C^2(\Omega) \cap C^1(\overline{\Omega}) : u = 0 \text{ on } \partial\Omega\} .$$

We say that $\lambda \in \mathbb{R}$ is an **eigenvalue** for the Laplacian operator $-\Delta$ on $\mathcal{X}$ if there exists a non-null function $u \in \mathcal{X}$ solving the Laplace equation $-\Delta u = \lambda u$ on Ω . We then say that u is an **eigenvector** or **eigenfunction** associated to λ .

We provide a rigorous treatment of this subject in Section 12.5.1. We consider the operator $-\Delta$ instead of Δ to obtain a nice “variational formulation” of the Laplace equation as we will see in Section 12.5.1.

For the rest of this section about the method of separation of variables, we will continue with a more formal presentation.

9.11.1 Eigenvalue Problem in a Rectangle

In this section, we consider the eigenvalue problem for the mapping $-\Delta$ on $\mathcal{X}$, where

$$\Omega = \{(x, y) : 0 < x < a, 0 < y < b\} .$$

Namely, we seek λ and $u \in C^2(\Omega) \cap C(\overline{\Omega})$ such that

$$-\Delta u = -\frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2} = \lambda u \quad , \quad (x, y) \in \Omega , \quad (9.11.1)$$

with the boundary conditions $u(x, 0) = u(x, b) = 0$ for $0 < x < a$ and $u(0, y) = u(a, y) = 0$ for $0 < y < b$.

Separation of Variables: If we substitute $u(x, y) = F(x)G(y)$ in (9.11.1), we get

$$-\frac{d^2 F}{dx^2}(x) G(y) - F(x) \frac{d^2 G}{dy^2}(y) = \lambda F(x) G(y) \quad , \quad (x, y) \in \Omega .$$

After dividing both sides by $F(x)G(y)$, we get

$$-\frac{1}{G(y)} \frac{d^2 G}{dy^2}(y) = \lambda + \frac{1}{F(x)} \frac{d^2 F}{dx^2}(x) \quad , \quad (x, y) \in \Omega .$$

Since the right hand side is independent of y and the left hand side is independent of x , we get

$$-\frac{1}{G(y)} \frac{d^2 G}{dy^2}(y) = \lambda + \frac{1}{F(x)} \frac{d^2 F}{dx^2}(x) = k \quad , \quad (x, y) \in \Omega ,$$

for some constant k . We end up with two ordinary differential equation.

$$\frac{d^2 F}{dx^2}(x) + (\lambda - k)F(x) = 0 \quad \text{and} \quad \frac{d^2 G}{dy^2}(y) + kG(y) = 0. \quad (9.11.2)$$

The first ordinary differential equation in (9.11.2) satisfies two boundary conditions. From $u(0, y) = 0$, we get $F(0)G(y) = 0$. Since we are looking for non-null solutions, we must have that G is non-null and $F(0) = 0$. Similarly, from $u(a, y) = 0$, we get $F(a)G(y) = 0$. Again, since we are looking for non-null solutions, we must have that G is non-null and we get $F(a) = 0$. The boundary conditions for the first ordinary differential equation in (9.11.2) are $F(0) = F(a) = 0$.

Similarly, the second ordinary differential equation in (9.11.2) satisfies the boundary conditions $G(0) = G(b) = 0$.

Eigenfunctions: We first consider the boundary value problem

$$\frac{d^2 G}{dy^2}(y) + kG(y) = 0 \quad \text{with} \quad G(0) = G(b) = 0. \quad (9.11.3)$$

The form of the general solution of the second order ordinary differential equation above is determined by the roots of the characteristic equation $\rho^2 + k = 0$.

For $k > 0$, we get $G(y) = C \cos(\sqrt{k}y) + D \sin(\sqrt{k}y)$. The boundary condition $G(0) = 0$ forces C to be 0. The boundary condition $G(b) = 0$ implies that $k = k_n \equiv (n\pi/b)^2$ for all integers $n > 0$ if we want to get a non-null solution. We obtain

$$G(y) = G_n(y) \equiv D_n \sin\left(\frac{n\pi y}{b}\right)$$

for any integer $n > 0$. The usual analysis based on the sign of k yields no other non-null solution.

We now consider the boundary value problems

$$\frac{d^2 F}{dx^2}(x) + (\lambda - k_n)F(x) = 0 \quad \text{with} \quad F(0) = F(\pi) = 0 \quad (9.11.4)$$

for $n > 0$. The form of the general solution of the second order ordinary differential equation above is determined by the roots of the characteristic equation $\rho^2 + (\lambda - k_n) = 0$.

For $\lambda - k_n > 0$, we get $F(x) = A_n \cos(\sqrt{\lambda - k_n}x) + B_n \sin(\sqrt{\lambda - k_n}x)$. The boundary condition $F(0) = 0$ forces A to be 0. The boundary condition $F(a) = 0$ implies that $\lambda - k_n = (m\pi/a)^2$ for all integers $m > 0$ if we want to get a non-null solution. We obtain

$$F(x) = F_m(x) \equiv B_m \sin\left(\frac{m\pi x}{a}\right)$$

for $m > 0$. The usual analysis based on the sign of $\lambda - k_n$ yields no other non-null solution.

For each m and n , we have found the non-null solution $u_{n,m}(x, y) \equiv F_m(x)G_n(y)$ associated to $\lambda = \lambda_{m,n} \equiv (m\pi/a)^2 + (n\pi/b)^2$ with $m, n > 0$. The $\lambda_{m,n}$ are the eigenvalues and, for each m and n , an eigenfunction associated to $\lambda_{m,n}$ is

$$u_{m,n}(x, y) = F_m(x)G_n(y) = \sin\left(\frac{m\pi x}{a}\right)\sin\left(\frac{n\pi y}{b}\right).$$

Remark 9.11.1

The eigenspace associated to an eigenvalue $\lambda_{m,n}$ can sometime be of dimension greater than 1. For instance, if $a = b$ then $\lambda_{1,2} = \lambda_{2,1}$ and the eigenfunctions $u_{1,2}$ and $u_{2,1}$ are two linearly independent functions in the eigenspace associated to $\lambda_{1,2} = \lambda_{2,1}$. ♠

9.11.2 Eigenvalue Problem in a Disk

In this section, we consider the eigenvalue problem for the mapping $-\Delta$ on $\mathcal{X}$, where

$$\Omega = \{(x, y) : \|(x, y)\| < R\}.$$

We seek λ and $u \in C^2(\Omega) \cap C(\overline{\Omega})$ such that

$$-\Delta u = -\frac{\partial^2 u}{\partial r^2} - \frac{1}{r} \frac{\partial u}{\partial r} - \frac{1}{r^2} \frac{d^2 u}{d\theta^2} = \lambda u \quad (9.11.5)$$

on Ω with the boundary conditions $u(R, \theta) = 0$ for $0 \leq \theta < 2\pi$.

Separation of Variables: If we substitute $u(r, \theta) = F(r)G(\theta)$ in (9.11.5), we get

$$-\frac{d^2 F}{dr^2}(r)G(\theta) - \frac{1}{r} \frac{\partial F}{\partial r}(r)G(\theta) - \frac{1}{r^2} F(r) \frac{d^2 G}{d\theta^2}(\theta) = \lambda F(r)G(\theta).$$

After multiplying both sides by $\frac{r^2}{F(r)G(\theta)}$, we get

$$-\frac{1}{G(\theta)} \frac{d^2 G}{d\theta^2}(\theta) = \frac{r^2}{F(r)} \frac{d^2 F}{dr^2}(r) + \frac{r}{F(r)} \frac{\partial F}{\partial r}(r) + \lambda r^2.$$

Since the right hand side is independent of θ and the left hand side is independent of r , we get

$$-\frac{1}{G(\theta)} \frac{d^2 G}{d\theta^2}(\theta) = \frac{r^2}{F(r)} \frac{d^2 F}{dr^2}(r) + \frac{r}{F(r)} \frac{\partial F}{\partial r}(r) + \lambda r^2 = k$$

for some constant k . We end up with two ordinary differential equation.

$$r^2 \frac{d^2 F}{dr^2}(r) + r \frac{\partial F}{\partial r}(r) + (\lambda r^2 - k) F(r) = 0 \quad \text{and} \quad \frac{d^2 G}{d\theta^2}(\theta) + k G(\theta) = 0. \quad (9.11.6)$$

The first ordinary differential equation in (9.11.6) satisfies two boundary conditions. From $u(R, \theta) = 0$, we get $F(R)G(\theta) = 0$. Since we are looking for non-null solutions, we must have

that G is non-null and $F(R) = 0$. There is another condition which is implicit. The function u , and thus F , must be bounded. As we will see, this condition is that F must be bounded at the origin.

The second ordinary differential equation has also an implicit boundary condition. The solution u must be periodic of period 2π in θ ; namely, $u(r, \theta) = u(r, \theta + 2\pi)$ for all θ . This gives $F(r)G(\theta) = F(r)G(\theta + 2\pi)$ for all r and θ . Since we are looking for non-null solutions, we must have that F is non-null and $G(\theta) = G(\theta + 2\pi)$ for all θ .

Eigenfunctions: We first consider the boundary value problem

$$\frac{d^2 G}{d\theta^2}(\theta) + k G(\theta) = 0 \quad \text{with} \quad G(\theta) = G(\theta + 2\pi). \quad (9.11.7)$$

If $k < 0$, the general solution of (9.11.7) is of the form $G(\theta) = Ae^{\sqrt{-k}\theta} + Be^{-\sqrt{-k}\theta}$. This cannot satisfy the condition $G(\theta) = G(\theta + 2\pi)$ for all θ unless $A = B = 0$.

If $k = 0$, the general solution of (9.11.7) is of the form $G(\theta) = B_0\theta + A_0$. We need $B_0 = 0$ to satisfy the condition $G(\theta) = G(\theta + 2\pi)$. Associated to $k = k_0 \equiv 0$, we get the solution $G(\theta) = G_0(\theta) \equiv A_0$ for all θ , where A_0 is a constant,

If $k > 0$, the general solution of (9.11.7) is of the form $G(\theta) = A \cos(\sqrt{k}\theta) + B \sin(\sqrt{k}\theta)$. To obtain non-null solutions satisfying the condition $G(\theta) = G(\theta + 2\pi)$ for all θ , we must have that $k = k_n = n^2$ for all integers $n > 0$. We get the non-null solution

$$G(\theta) = G_n(\theta) \equiv A_n \cos(n\theta) + B_n \sin(n\theta)$$

for any integer $n > 0$. Note that $n = 0$ in the formula above yields the constant solution $G = A_0$ found for $k = k_0 = 0$.

We now consider the boundary value problems

$$r^2 \frac{d^2 F}{dr^2}(r) + r \frac{\partial F}{\partial r}(r) + (\lambda r^2 - n^2) F(r) = 0 \quad (9.11.8)$$

with $F(R) = 0$ and F bounded for any integer $n \geq 0$.

As we will see in Section 12.5.1, the eigenvalues of $-\Delta$ are positive. We can therefore define the change of variable $s = \sqrt{\lambda}r$. Since

$$\frac{dH}{dr} = \frac{dH}{ds} \frac{ds}{dr} = \sqrt{\lambda} \frac{dH}{ds} \quad \text{and} \quad \frac{d^2 H}{dr^2} = \frac{d}{dr} \left(\frac{dH}{dr} \right) = \sqrt{\lambda} \frac{d}{ds} \left(\sqrt{\lambda} \frac{dH}{ds} \right) = \lambda \frac{d^2 H}{ds^2}$$

for any twice differentiable function H , the differential equation in (9.11.8) becomes

$$s^2 \frac{d^2 F}{ds^2}(s) + s \frac{\partial F}{\partial s}(s) + (s^2 - n^2) F(s) = 0$$

for any integer $n \geq 0$, where, as usual, we are using the same name for F as a function of s and F as a function of r . For each n , this is the Bessel equation of order n whose general solution is of the form $F(s) = C_n J_n(s) + D_n Y_n(s)$. Thus, $F(r) = C_n J_n(\sqrt{\lambda}r) + D_n Y_n(\sqrt{\lambda}r)$. The

condition that u be bounded (at the origin) implies that $D_n = 0$. To get a non-null solution, we must have that $C_m \neq 0$. Hence, the condition $F(R) = 0$ implies that $J_n(\sqrt{\lambda}R) = 0$. Thus, $\lambda = \lambda_{m,n} \equiv (\alpha_{m,n}/R)^2$, where $\alpha_{m,n}$ is the m^{th} zero of the Bessel function J_n . We get the non-null solutions

$$F(r) = F_{m,n}(r) \equiv C_{m,n} J_n\left(\frac{\alpha_{m,n} r}{R}\right)$$

for any integers $n, m \geq 0$. For $n \geq 1$, we have $\alpha_{0,n} = 0$ so we may ignore $m = 0$ for $n > 0$ because $J_n(\alpha_{0,n} r/R) = J_n(0) = 0$ for all r .

We have found non-null solutions only for $\lambda = \lambda_{m,n} = (\alpha_{m,n}/R)^2$ with $m > 0$ for $n > 0$ and $m \geq 0$ for $n = 0$. The solution associated to $\lambda_{m,n}$ is

$$u_{m,n}(x, y) \equiv F_{m,n}(r) G_n(\theta) = (A_{m,n} \cos(n\theta) + B_{m,n} \sin(n\theta)) J_n\left(\frac{\alpha_{m,n} r}{R}\right),$$

where $A_{m,n} = A_n C_{m,n}$ and $B_{m,n} = B_n C_{m,n}$ are constants.

Hence, the eigenvalues are $\lambda_{m,n} = (\alpha_{m,n}/R)^2$ with $m > 0$ for $n > 0$ and $m \geq 0$ for $n = 0$. The eigenspace associated to the eigenvalue $\lambda_{m,n}$ is generated by

$$\cos(n\theta) J_n\left(\frac{\alpha_{m,n} r}{R}\right) \quad \text{and} \quad \sin(n\theta) J_n\left(\frac{\alpha_{m,n} r}{R}\right).$$

Note that for the smallest eigenvalue $\lambda_{0,0} = (\alpha_{0,0}/R)^2$, the eigenfunctions are multiple of $J_0(\alpha_{0,0} r/R)$. Thus, the eigenspace associated to $\lambda_{0,0}$ is of dimension one as predicted by the theory.

9.11.3 Eigenvalue Problem on a Sphere

In this section, we consider the eigenvalue problem for the mapping $-\Delta$ on $\mathcal{X}$, where the region

$$\Omega = \{(x, y, z) : \|(x, y, z)\| = 1\}.$$

We seek λ and $u \in C^2(\Omega)$ such that

$$-\Delta u = -\frac{\partial^2 u}{\partial \phi^2} - \cot(\phi) \frac{\partial u}{\partial \phi} - \frac{1}{\sin^2(\phi)} \frac{\partial^2 u}{\partial \theta^2} = \lambda u \quad (9.11.9)$$

on Ω . There is no boundary conditions but u is periodic of period 2π in θ and u is bounded on the sphere, in particular at the poles.

Separation of Variables: If we substitute $u(\theta, \phi) = G(\theta)H(\phi)$ in (9.11.9), we get

$$-G(\theta) \frac{d^2 H}{d\phi^2}(\phi) - \cot(\phi) G(\theta) \frac{dH}{d\phi}(\phi) - \frac{1}{\sin^2(\phi)} \frac{d^2 G}{d\theta^2} H(\phi) = \lambda G(\theta) H(\phi). \quad (9.11.10)$$

After multiplying both sides by $\frac{\sin^2(\phi)}{G(\theta)H(\phi)}$, we get

$$\frac{1}{H(\phi)} \left(\sin^2(\phi) \frac{d^2 H}{d\phi^2}(\phi) + \cos(\phi) \sin(\phi) \frac{dH}{d\phi}(\phi) \right) + \lambda \sin^2(\phi) = -\frac{1}{G(\theta)} \frac{d^2 G}{d\theta^2}(\theta).$$

Since the right hand side is independent of ϕ and the left hand side is independent of θ , we get

$$\frac{1}{H(\phi)} \left(\sin^2(\phi) \frac{d^2 H}{d\phi^2}(\phi) + \cos(\phi) \sin(\phi) \frac{dH}{d\phi}(\phi) \right) + \lambda \sin^2(\phi) = -\frac{1}{G(\theta)} \frac{d^2 G}{d\theta^2}(\theta) = k$$

for some constant k . We get two ordinary differential equation.

$$\frac{d^2 G}{d\theta^2}(\theta) + kG(\theta) = 0 \quad (9.11.11)$$

and

$$\sin^2(\phi) \frac{d^2 H}{d\phi^2}(\phi) + \cos(\phi) \sin(\phi) \frac{dH}{d\phi}(\phi) + (\lambda \sin^2(\phi) - k) H(\phi) = 0 \quad (9.11.12)$$

Eigenfunctions: We first solve (9.11.11). Since u is periodic of period 2π in θ , the function G must also be periodic of period 2π . As for the eigenvalue problem on a disk, we find that $k = k_m \equiv m^2$ and

$$G(\theta) = G_m(\theta) \equiv A_m \cos(m\theta) + B_m \sin(m\theta)$$

for all integers $m \geq 0$. Since $G_0(\theta) = A_0$ for all θ , the term $B_m \sin(m\theta)$ can be ignored when $m = 0$.

We now solve (9.11.12) with $k = m^2$ for $m \geq 0$. Namely, we consider

$$\sin(\phi) \frac{d}{d\phi} \left(\sin(\phi) \frac{dH}{d\phi}(\phi) \right) + (\lambda \sin^2(\phi) - m^2) H(\phi) = 0 \quad (9.11.13)$$

for $m \geq 0$.

As we did for the potential equation in spherical coordinates, we use the change of variable $w = \cos(\phi)$ for $0 < \phi < \pi$. We have that $\frac{dV}{d\phi}(\phi) = \frac{dV}{dw}(w) \frac{dw}{d\phi}(\phi) = -\sin(\phi) \frac{dV}{dw}(w)$ and $\sin(\phi) \frac{dV}{d\phi}(\phi) = -\sin^2(\phi) \frac{dV}{dw}(w) = (\cos^2(\phi) - 1) \frac{dV}{dw}(w) = (w^2 - 1) \frac{dV}{dw}(w)$ for any differentiable functions V . As usual, to simplify the notation, we use the same name to designate V as a function of ϕ and V as a function of w . We then have that $V(\phi) = V(w)$ for $w = \cos(\phi)$.

Hence, with the change of variable $w = \cos(\phi)$ for $0 < \phi < \pi$, the equation (9.11.13) becomes

$$(w^2 - 1) \frac{d}{dw} \left((w^2 - 1) \frac{dH}{dw}(w) \right) + (\lambda (w^2 - 1) - m^2) H(w) = 0 \quad (9.11.14)$$

for $m \geq 0$. This is the associated Legendre equation of order m and has regular singular points at $w = -1$ and $w = 1$. For each integer $m > 0$, (9.11.14) possesses non-null bounded solutions when $\lambda = \lambda_n \equiv n(n+1)$ for any integer $n \geq m$. For fixed values of m and n , the non-null bounded solutions are multiple of the associated Legendre polynomial

$$P_n^m(w) = (1 - w^2)^{m/2} \frac{d^m P_n}{dw^m}(w) .$$

We note that $\frac{d^m P_n}{dw^m}(w) = 0$ for all w when $m > n$ because P_n is the Legendre polynomial of degree n ,

Hence, for each integer $m \geq 0$, (9.11.12) has non-null bounded solutions when $\lambda = \lambda_n = n(n+1)$ with n an integer such that $n \geq m$. For fixed values of m and n , the non-null bounded solutions are multiple of $H(\phi) = H_{m,n}(\phi) \equiv P_n^m(\cos(\phi))$.

We have found non-null solutions only for $\lambda = \lambda_n = n(n+1)$. The solutions associated to λ_n are of the form

$$u_{m,n}(x, y) \equiv \sum_{m=0}^n G_m(\theta) H_{m,n}(\phi) = \sum_{m=0}^n A_m \cos(m\theta) P_n^m(\cos(\phi)) + \sum_{m=1}^n B_m \sin(m\theta) P_n^m(\cos(\phi)),$$

where the A_n and B_n are constants.

Hence, the eigenvalues are $\lambda = \lambda_n = n(n+1)$ for any integer $n \geq 0$. The eigenspace associated to the eigenvalue λ_n is generated by the eigenfunctions $\cos(m\theta) P_n^m(\cos(\phi))$ for $0 \leq m \leq n$ and $\sin(m\theta) P_n^m(\cos(\phi))$ for $1 \leq m \leq n$. These eigenfunctions are known as **Surface Harmonics** of order n . The dimension of the eigenspace associated to λ_n is of dimension $2n+1$.

9.11.4 Eigenvalue Problem in a Ball

In this section, we consider the eigenvalue problem for the mapping $-\Delta$ on $\mathcal{X}$, where

$$\Omega = \{(x, y, z) : \|(x, y, z)\| < R\}.$$

We seek λ and $u \in C^2(\Omega) \cap C(\overline{\Omega})$ such that

$$-\Delta u = -\frac{\partial^2 u}{\partial r^2} - \frac{2}{r} \frac{\partial u}{\partial r} - \frac{1}{r^2} \frac{\partial^2 u}{\partial \phi^2} - \frac{\cot(\phi)}{r^2} \frac{\partial u}{\partial \phi} - \frac{1}{r^2 \sin^2(\phi)} \frac{\partial^2 u}{\partial \theta^2} = \lambda u \quad (9.11.15)$$

on Ω with the boundary conditions $u(R, \theta, \phi) = 0$ for $0 \leq \theta < 2\pi$ and $0 < \phi < \pi$.

Separation of Variables: If we substitute $u(r, \theta, \phi) = F(r)G(\theta)H(\phi)$ in (9.11.15), we get

$$\begin{aligned} & -\frac{d^2 F}{dr^2}(r) G(\theta) H(\phi) - \frac{2}{r} \frac{dF}{dr}(r) G(\theta) H(\phi) - \frac{1}{r^2} F(r) G(\theta) \frac{d^2 H}{d\phi^2}(\phi) \\ & - \frac{\cot(\phi)}{r^2} F(r) G(\theta) \frac{dH}{d\phi}(\phi) - \frac{1}{r^2 \sin^2(\phi)} F(r) \frac{d^2 G}{d\theta^2} H(\phi) = \lambda F(r) G(\theta) H(\phi). \end{aligned}$$

After multiplying both sides by $\frac{r^2}{F(r)G(\theta)H(\phi)}$, we get

$$\begin{aligned} & -\frac{1}{F(r)} \left(r^2 \frac{d^2 F}{dr^2}(r) + 2r \frac{dF}{dr}(r) \right) - \lambda r^2 \\ & = \frac{1}{H(\phi)} \left(\frac{d^2 H}{d\phi^2}(\phi) + \cot(\phi) \frac{dH}{d\phi}(\phi) \right) + \frac{1}{\sin^2(\phi) G(\theta)} \frac{d^2 G}{d\theta^2}(\theta). \end{aligned}$$

Since the right hand side is independent of r and the left hand side is independent of θ and ϕ , we get

$$\begin{aligned} & -\frac{1}{F(r)} \left(r^2 \frac{d^2 F}{dr^2}(r) + 2r \frac{dF}{dr}(r) \right) - \lambda r^2 \\ & = \frac{1}{H(\phi)} \left(\frac{d^2 H}{d\phi^2}(\phi) + \cot(\phi) \frac{dH}{d\phi}(\phi) \right) + \frac{1}{\sin^2(\phi)G(\theta)} \frac{d^2 G}{d\theta^2}(\theta) = p \end{aligned}$$

for some constant p . We get the ordinary differential equation

$$r^2 \frac{d^2 F}{dr^2}(r) + 2r \frac{dF}{dr}(r) + (p + \lambda r^2)F(r) = 0 \quad (9.11.16)$$

and the equation

$$\frac{1}{H(\phi)} \left(\frac{d^2 H}{d\phi^2}(\phi) + \cot(\phi) \frac{dH}{d\phi}(\phi) \right) + \frac{1}{\sin^2(\phi)G(\theta)} \frac{d^2 G}{d\theta^2}(\theta) = p$$

This equation is (9.11.10) with λ replaced by $-p$. In the previous section, we have found the conditions that must be satisfied by p to get non-null and bounded solutions. More precisely, we have found that non-null bounded solutions exist only for $p = p_n \equiv -n(n+1)$ for $n \geq 0$. For each n , the bounded solutions are linear combinations of terms of the form $G(\theta)H(\phi) = \cos(m\theta)P_n^m(\cos(\phi))$ for $0 \leq m \leq n$ and $G(\theta)H(\phi) = \sin(m\theta)P_n^m(\cos(\phi))$ for $0 \leq m \leq n$.

Eigenfunctions: We only have to solve (9.11.16) with $p = p_n = -n(n+1)$ for $n \geq 0$. We get from $u(R, \theta, \phi) = 0$ that $F(R)G(\theta)H(\phi) = 0$ for all θ and ϕ . Since G and H are non-null functions, we must have that $F(R) = 0$. Moreover, F must be bounded at the origin since it is required for u .

Let $F(r) = r^{-1/2} \rho(r)$. We have that $\frac{dF}{dr} = -\frac{1}{2} r^{-3/2} \rho(r) + r^{-1/2} \frac{d\rho}{dr}(r)$ and $\frac{d^2 F}{dr^2} = \frac{3}{4} r^{-5/2} \rho(r) - r^{-3/2} \frac{d\rho}{dr}(r) + r^{-1/2} \frac{d^2 \rho}{dr^2}$. So, If we substitute $F(r) = r^{-1/2} \rho(r)$ in (9.11.16) and simplify, we get

$$r^2 \frac{d^2 \rho}{dr^2} + r \frac{d\rho}{dr} + \left(\lambda r^2 - \left(n + \frac{1}{2} \right)^2 \right) \rho = 0 .$$

With $z = \lambda^{1/2} r$, we have $\frac{dV}{dr} = \frac{dV}{dz} \frac{dz}{dr} = \lambda^{1/2} \frac{dV}{dz}$ and $\frac{d^2 V}{dr^2} = \frac{d}{dr} \left(\frac{dV}{dr} \right) = \lambda^{1/2} \frac{d}{dz} \left(\lambda^{1/2} \frac{dV}{dz} \right) = \lambda \frac{d^2 V}{dz^2}$ for any twice differentiable function V . Hence, with the substitution $z = \lambda^{1/2} r$, the previous ordinary differential equation becomes

$$z^2 \frac{d^2 \rho}{dz^2} + z \frac{d\rho}{dz} + \left(z^2 - \left(n + \frac{1}{2} \right)^2 \right) \rho = 0 .$$

This is a Bessel Equation of order $n + 1/2$. Its general solution is $\rho(z) = AJ_{n+1/2}(z) + BJ_{-n-1/2}(z)$. The general solution of (9.11.16) with $p = p_n = -n(n+1)$ is therefore

$$F(r) = Ar^{-1/2} J_{n+1/2}(\sqrt{\lambda} r) + Br^{-1/2} J_{-n-1/2}(\sqrt{\lambda} r) .$$

Since F must be bounded at the origin, we have $B = 0$. It follows from the condition $F(R) = 0$ that $0 = F(R) = AR^{-1/2}J_{n+1/2}(\sqrt{\lambda}R)$. Thus, $\lambda = \lambda_{j,n} \equiv (\alpha_{j,n+1/2}/R)^2$ for $j \geq 1$, where $\alpha_{j,n+1/2}$ is the j^{th} root of the Bessel function $J_{n+1/2}$.

We have shown that (9.11.15) has bounded solutions only when $\lambda = \lambda_{j,n} = (\alpha_{j,n+1/2}/R)^2$ and the non-null bounded solutions associated to $\lambda_{j,n}$ are linear combinations of $r^{-1/2}J_{n+1/2}(\alpha_{j,n+1/2}r/R)\cos(m\theta)P_n^m(\cos(\phi))$ for $0 \leq m \leq n$, and $r^{-1/2}J_{n+1/2}(\alpha_{j,n+1/2}r/R)\sin(m\theta)P_n^m(\cos(\phi))$ for $1 \leq m \leq n$.

Hence, the eigenvalues of (9.11.15) are $\lambda_{j,n} = (\alpha_{j,n+1/2}/R)^2$ for any integers $n \geq 0$ and $j > 0$. The eigenspace associated to the eigenvalue $\lambda_{j,n}$ is generated by the eigenfunctions $r^{-1/2}J_{n+1/2}(\alpha_{j,n+1/2}r/R)\cos(m\theta)P_n^m(\cos(\phi))$ for $0 \leq m \leq n$, and $r^{-1/2}J_{n+1/2}(\alpha_{j,n+1/2}r/R)\sin(m\theta)P_n^m(\cos(\phi))$ for $1 \leq m \leq n$. These eigenfunctions are known as **Spherical Harmonics** of order n . The dimension of the eigenspace associated to $\lambda_{j,n}$ is of dimension $2n + 1$.

9.12 Exercises

Suggested exercises:

- In [16]: all the numbers in Section 4.2.
- In [22]: all the numbers in Sections 4.1; numbers 1 to 9 in Section 4.2; numbers 2, 5 to 7 in Section 4.4.
- In [24]: numbers 7.2 to 7.5, 7.8, 7.10 to 7.12, 7.19, 7.20 and 7.22 in Section 7.9; numbers 9.8 to 9.10 in Section 9.13.
- In [38]: numbers 2 to 12 in Section 6.1; all the numbers in Section 6.2; numbers 1 to 3 in Section 6.3; numbers 1, 2, 5 to 11, and 13 in Section 6.4; numbers 5 to 9, 12 to 18 in Section 7.4; numbers 14 to 19 in Section 10.5; numbers 6 to 8 in Sections 10.6.

Chapter 10

Heat Equation

Let Ω be an open subset of $\mathbb{R}^n$. The **heat equation** is the partial differential equation

$$L(\mathbf{x}, t, D)u = \frac{\partial u}{\partial t} - c^2 \Delta u = 0$$

on Ω . The constant c^2 represent the heat conductivity in the region Ω . In this chapter Δ stands for $\Delta_{\mathbf{x}}$ unless otherwise specified.

The problem is to find u that satisfies the heat equation, some boundary conditions, and an initial condition. The most frequent boundary conditions are

- $u(\mathbf{x}, t) = f(\mathbf{x})$ for $\mathbf{x} \in \partial\Omega$ and $t \geq 0$. This means that the temperature is maintained to $f(\mathbf{x})$ at $\mathbf{x} \in \partial\Omega$ for all time.
- $\frac{\partial u}{\partial \nu}(\mathbf{x}, t) = g(\mathbf{x})$ for $\mathbf{x} \in \partial\Omega$ and $t \geq 0$, where $\nu(\mathbf{x})$ is the outward unit normal to the surface $\partial\Omega$ at $\mathbf{x} \in \partial\Omega$. This means that the flow of heat through $\mathbf{x} \in \partial\Omega$ is $g(\mathbf{x})$ for all time. The most frequent situation is when $g = 0$ on $\partial\Omega$; namely, the boundary of Ω is insulated and no heat flow in or out of Ω .
- $u(\mathbf{x}, t) = f(\mathbf{x})$ on some parts of $\partial\Omega$ and $\frac{\partial u}{\partial \nu}(\mathbf{x}, t) = g(\mathbf{x})$ on the other parts of $\partial\Omega$ for all $t \geq 0$.
- $\frac{\partial u}{\partial \nu}(\mathbf{x}, t) + \alpha(u(\mathbf{x}, t) - u_o(\mathbf{x}, t)) = 0$ for $\mathbf{x} \in \partial\Omega$ and $t \geq 0$, where α is a constant. This means that the heat flow across $\partial\Omega$ at all time is proportional to the difference between the temperature u inside Ω and the temperature u_o outside Ω . This is **Newton's Law** of heating.

The initial condition is normally $u(\mathbf{x}, 0) = v(\mathbf{x})$ for $\mathbf{x} \in \Omega$ and represent the heat distribution in Ω initially.

10.1 Cauchy Problem

10.1.1 One Dimensional Heat (Diffusion) Equation

We consider

$$\frac{\partial u}{\partial t} - c^2 \frac{\partial^2 u}{\partial x^2} = 0 \quad , \quad x \in \mathbb{R} \text{ and } t > 0 \quad , \quad (10.1.1)$$

$$u(x, 0) = f(x) \quad , \quad x \in \mathbb{R} \quad . \quad (10.1.2)$$

We assume that f is continuous and integrable on $\mathbb{R}$. This implies in particular that $\lim_{x \rightarrow \pm\infty} f(x) = 0$.

The solution of the heat equation satisfies the following properties which will guide us in the search of this solution.

1. Spatial translation of a solution u of (10.1.1) is still a solution of (10.1.1); namely, $v(x, t) = u(x - y, t)$ satisfies (10.1.1).
2. Any partial derivative of a solution u of (10.1.1) is still a solution of (10.1.1).
3. Linear combination of solutions of (10.1.1) is a solution of (10.1.1).
4. If K is a solution of (10.1.1), then

$$v(x, t) = \int_{-\infty}^{\infty} K(x - y, t) f(y) dy$$

is also a solution of (10.1.1) if proper convergence conditions are assumed.

5. If $a > 0$, then $v(x, t) = u(\sqrt{a}x, at)$ is a solution of (10.1.1) if u is a solution of (10.1.1). The solution of (10.1.1) is invariant under the transformation $(x, t) \mapsto (\sqrt{a}x, at)$.

To solve the heat equation, we first find the solution of the following problem.

$$\frac{\partial Q}{\partial t} - c^2 \frac{\partial^2 Q}{\partial x^2} = 0 \quad , \quad x \in \mathbb{R} \text{ and } t > 0 \quad , \quad (10.1.3)$$

$$Q(x, 0) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x < 0 \end{cases} \quad (10.1.4)$$

The solution that we will find has a singularity at the origin.

Because the solution of (10.1.3) is invariant under the transformation $(x, t) \mapsto (\sqrt{a}x, at)$, we seek a solution of the form $Q(x, t) = q\left(\frac{x}{2c\sqrt{t}}\right)$. We introduce the factor $\frac{1}{2c}$ for aesthetic reasons as it will become apparent later. Hence, (10.1.3) becomes

$$q''(s) + 2sq'(s) = 0 \quad , \quad (10.1.5)$$

where $s = x/(2c\sqrt{t})$. The general solution of (10.1.5) is

$$q(s) = C_1 \int_0^s e^{-r^2} dr + C_2$$

for some constants C_1 and C_2 . Hence, we get

$$Q(x, t) = C_1 \int_0^{x/(2c\sqrt{t})} e^{-r^2} dr + C_2 \quad (10.1.6)$$

To determine C_1 and C_2 , we use the initial condition (10.1.4). We require that

$$1 = \lim_{t \rightarrow 0^+} Q(x, t) = C_1 \int_0^\infty e^{-r^2} dr + C_2 = C_1 \frac{\sqrt{\pi}}{2} + C_2$$

for $x > 0$ and

$$0 = \lim_{t \rightarrow 0^+} Q(x, t) = C_1 \int_0^{-\infty} e^{-r^2} dr + C_2 = -C_1 \frac{\sqrt{\pi}}{2} + C_2$$

for $x < 0$. We get $C_1 = 1/\sqrt{\pi}$ and $C_2 = 1/2$. Thus

$$Q(x, t) = \begin{cases} \frac{1}{2} + \frac{1}{\sqrt{\pi}} \int_0^{x/(2c\sqrt{t})} e^{-r^2} dr & \text{if } x \in \mathbb{R}, t > 0 \\ 1 & \text{if } x > 0, t = 0 \\ 1/2 & \text{if } x = 0 \\ 0 & \text{if } x < 0, t = 0 \end{cases}$$

Let

$$K(x, t) = \frac{\partial Q}{\partial x}(x, t) = \frac{1}{2c\sqrt{\pi t}} e^{-x^2/(4c^2 t)}$$

for $x \in \mathbb{R}$ and $t > 0$. We show that

$$u(x, t) = \begin{cases} \int_{-\infty}^\infty K(x - y, t) f(y) dy & \text{if } x \in \mathbb{R}, t > 0 \\ f(x) & \text{if } x \in \mathbb{R}, t = 0 \end{cases} \quad (10.1.7)$$

is the solution of (10.1.1) and (10.1.2). The second and fourth properties for the solution of (10.1.1) mentioned above implies that u satisfies (10.1.1). For the fourth properties, we note that for every $T > 0$ there exists a constant C such that $|K(x - y, t)f(y)| \leq C|f(y)|$, $\left| \frac{\partial K}{\partial t}(x - y, t)f(y) \right| \leq C|f(y)|$, $\left| \frac{\partial}{\partial x} K(x - y, t)f(y) \right| \leq C|f(y)|$ and $\left| \frac{\partial^2}{\partial x^2} K(x - y, t)f(y) \right| \leq C|f(y)|$ for $x \in \mathbb{R}$ and $t \geq T > 0$. Since $f \in L^1(\mathbb{R})$, we may interchange the integral in the definition of u with the derivatives of u with respect to t and x (see [7]).

To prove that u is a continuous function for $x \in \mathbb{R}$ and $t \geq 0$, we only have to prove that u is continuous at $(x, 0)$ for $x \in \mathbb{R}$ because u is continuous for $\mathbb{R} \times]0, \infty[$ follows from $(x, t) \mapsto K(x, t)f(x)$ being continuous on $\mathbb{R} \times]0, \infty[$. We first note that $\int_{-\infty}^\infty K(y, t) dy = 1$ for all $t > 0$. Let $M = \sup_{x \in \mathbb{R}} |f(x)|$. We have that $M < \infty$ because f is a continuous function in

$L^1(\mathbb{R})$. Given $x \in \mathbb{R}$ and $\epsilon > 0$, choose $\delta > 0$ such that $|f(y) - f(x)| < \epsilon/2$ for $|y - x| < \delta$. Then, choose δ smaller if needed such that

$$\int_{|x|>\delta} K(x, t) dx = \frac{1}{c\sqrt{\pi t}} \int_{\delta}^{\infty} e^{-x^2/(4c^2 t)} dx = \frac{2}{\sqrt{\pi}} \int_{\delta/(2c\sqrt{t})}^{\infty} e^{-s^2} ds < \frac{\epsilon}{4M}$$

for $0 < t < \delta$. Hence,

$$\begin{aligned} |u(z, t) - u(x, 0)| &= \left| \int_{-\infty}^{\infty} K(z - y, t)(f(y) - f(x)) dy \right| \\ &\leq \left| \int_{|z-y|<\delta} K(z - y, t)(f(y) - f(x)) dy \right| + \left| \int_{|z-y|\geq\delta} K(z - y, t)(f(y) - f(x)) dy \right| \\ &\leq \int_{|z-y|<\delta} |K(z - y, t)| |f(y) - f(x)| dy + \int_{|z-y|\geq\delta} |K(z - y, t)| |f(y) - f(x)| dy \\ &\leq \underbrace{\frac{\epsilon}{2} \int_{|z-y|<\delta} |K(z - y, t)| dy}_{\leq 1} + \underbrace{2M \int_{|z-y|\geq\delta} |K(z - y, t)| dy}_{\epsilon/(4M)} \leq \epsilon \end{aligned}$$

for $|z - x| < \delta$ and $0 < t < \delta$.

Remark 10.1.1

If f is differentiable and $f' \in L^1(\mathbb{R})$, there is an easy to prove that $\lim_{t \rightarrow 0^+} u(x, t) = f(x)$. Note that this is not enough to prove the continuity of u at $(x, 0)$ as a function from $\mathbb{R} \times [0, \infty[$ to $\mathbb{R}$ because we approach $(x, 0)$ in only one direction.

Using integration by parts, we have for $t > 0$ that

$$\begin{aligned} u(x, t) &= \int_{-\infty}^{\infty} \frac{\partial Q}{\partial x}(x - y, t) f(y) dy = - \int_{-\infty}^{\infty} \left(\frac{\partial}{\partial y} Q(x - y, t) \right) f(y) dy \\ &= \int_{-\infty}^{\infty} Q(x - y, t) f'(y) dy - Q(x - y, t) f(y) \Big|_{-\infty}^{\infty} \end{aligned}$$

Since $\lim_{x \rightarrow \pm\infty} f(x) = 0$, we get

$$u(x, t) = \int_{-\infty}^{\infty} Q(x - y, t) f'(y) dy .$$

Hence,

$$\lim_{t \rightarrow 0^+} u(x, t) = \int_{-\infty}^{\infty} Q(x - y, 0) f'(y) dy = \int_{-\infty}^x f'(y) dy = f(y) \Big|_{-\infty}^x = f(x) ,$$

where we have again used $\lim_{x \rightarrow \pm\infty} f(x) = 0$. We should note that $|Q(x - y, t) f'(y)| < |f'(y)|$ for all x and $t \geq 0$. Since $f' \in L^1(\mathbb{R})$, we may use the Lebesgue Dominated Convergence Theorem to justify the first equality in the previous expression. ♠

Explicitly, (10.1.7) is

$$u(x, t) = \frac{1}{2c\sqrt{\pi t}} \int_{-\infty}^{\infty} e^{-(x-y)^2/(4c^2 t)} f(y) dy$$

for $t > 0$.

The function $(x, t) \mapsto K(x, t)$ is known under several names: **source function**, **Green's function**, **fundamental solution**, etc. As t converges toward 0, the function $x \mapsto K(x, t)$ behaves like the Dirac delta function at the origin. From the solution of the heat equation, we observe that if we have a hot spot near the origin at $t = 0$, then the heat will spread evenly as t increases.

The heat equation can be solved “in some sense” even when the initial condition is not a continuous function. This will be the subject of a section later on. First, we consider the non-homogeneous wave equation.

10.1.2 Non-homogeneous Heat Equation, Diffusion with a Source

We consider

$$\frac{\partial u}{\partial t} - c^2 \frac{\partial^2 u}{\partial x^2} = g(x, t) \quad , \quad x \in \mathbb{R} \text{ and } t > 0 \quad , \quad (10.1.8)$$

$$u(x, 0) = f(x) \quad , \quad x \in \mathbb{R} \quad . \quad (10.1.9)$$

We assume that f is continuous and integrable on $\mathbb{R}$. In particular, this implies that $\lim_{x \rightarrow \pm\infty} f(x) = 0$. Moreover, we assume that g is also continuously differentiable (at least with respect to x) on $\mathbb{R} \times]0, \infty[$ and integrable on $\mathbb{R} \times [0, T]$ for all $T > 0$. Again, this implies that $\lim_{x \rightarrow \pm\infty} g(x, t) = 0$ for all t . With these assumptions, all the following integrals involving g and $\frac{\partial g}{\partial x}$ converge and we can interchange limits and integrals using the Lebesgue Dominated Convergence Theorem.

We show that the solution of (10.1.8) and (10.1.9) is

$$u(x, t) = \int_{-\infty}^{\infty} K(x - y, t) f(y) dy + \int_0^t \int_{-\infty}^{\infty} K(x - y, t - s) g(y, s) dy ds$$

for $t > 0$. The first integral is the solution of the homogeneous heat equation (10.1.1) and (10.1.2), and the second integral is a solution of the non-homogeneous heat equation

$$\frac{\partial u}{\partial t} - c^2 \frac{\partial^2 u}{\partial x^2} = g(x, t) \quad , \quad x \in \mathbb{R} \text{ and } t > 0 \quad , \quad (10.1.10)$$

$$u(x, 0) = 0 \quad , \quad x \in \mathbb{R} \quad . \quad (10.1.11)$$

We only have to prove that

$$v(x, t) = \int_0^t \int_{-\infty}^{\infty} K(x - y, t - s) g(y, s) dy ds$$

satisfies (10.1.10) because (10.1.11) is obviously satisfied by v .

Since

$$\frac{\partial K}{\partial t}(x - y, t - s) = c^2 \frac{\partial^2 K}{\partial x^2}(x - y, t - s) \quad ,$$

we have

$$\begin{aligned}\frac{\partial v}{\partial t} &= \int_0^t \int_{-\infty}^{\infty} \frac{\partial K}{\partial t}(x-y, t-s)g(y, s) \, dy \, ds + \lim_{s \rightarrow t^-} \int_{-\infty}^{\infty} K(x-y, t-s)g(y, s) \, dy \\ &= \int_0^t \int_{-\infty}^{\infty} c^2 \frac{\partial^2 K}{\partial x^2}(x-y, t-s)g(y, s) \, dy \, ds + g(x, t) = c^2 \frac{\partial^2 v}{\partial x^2}(x, t) + g(x, t)\end{aligned}$$

for $t > 0$ and $x \in \mathbb{R}$ because

$$\begin{aligned}\lim_{s \rightarrow t^-} \int_{-\infty}^{\infty} K(x-y, t-s)g(y, s) \, dy &= \lim_{\epsilon \rightarrow 0^+} \int_{-\infty}^{\infty} K(x-y, \epsilon)g(y, t-\epsilon) \, dy \\ &= - \lim_{\epsilon \rightarrow 0^+} \int_{-\infty}^{\infty} \left(\frac{\partial}{\partial y} Q(x-y, \epsilon) \right) g(y, t-\epsilon) \, dy \\ &= \lim_{\epsilon \rightarrow 0^+} \left(\int_{-\infty}^{\infty} Q(x-y, \epsilon) \frac{\partial g}{\partial y}(y, t-\epsilon) \, dy - Q(x-y, \epsilon)g(y, t-\epsilon) \Big|_{y=-\infty}^{\infty} \right) \\ &= \int_{-\infty}^{\infty} Q(x-y, 0) \frac{\partial g}{\partial y}(y, t) \, dy = \int_{-\infty}^x \frac{\partial g}{\partial y}(y, t) \, dy = g(y, t) \Big|_{y=-\infty}^x = g(x, t) ,\end{aligned}$$

where we have used several times that $\lim_{x \rightarrow \pm\infty} g(x, t) = 0$ for all $t > 0$.

10.1.3 Higher Dimensional Heat Equation

Consider the Cauchy problem

$$\frac{\partial u}{\partial t}(\mathbf{x}, t) - \Delta u(\mathbf{x}, t) = 0 \quad , \quad \mathbf{x} \in \mathbb{R}^n \text{ and } t > 0 , \quad (10.1.12)$$

$$u(\mathbf{x}, 0) = f(\mathbf{x}) \quad , \quad \mathbf{x} \in \mathbb{R}^n , \quad (10.1.13)$$

where $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is a function to be specified later.

Remark 10.1.2

The hyperplane $t = 0$ is a characteristic surface because the outward unit normal to this hyperplane is $\nu(\mathbf{x}, 0) = (\mathbf{0}, -1)$ for all $\mathbf{x} \in \mathbb{R}^n$, where $\mathbf{0}$ is the origin in $\mathbb{R}^n$, and the principal symbol for the heat equation is

$$Q((\mathbf{x}, t), (\mathbf{y}, s)) = \sum_{j=1}^n y_j^2 .$$

Hence, $Q((\mathbf{x}, 0), \nu(\mathbf{x}, 0)) = 0$ for all $\mathbf{x} \in \mathbb{R}^n$. ♠

Let us assume that $f \in \mathcal{S}(\mathbb{R}^n)$. Moreover, in addition of u being of class C^2 on $\mathbb{R}^n \times]0, \infty[$ and continuous on $\mathbb{R}^n \times [0, \infty[$, let us assume that $\mathbf{x} \mapsto u(\mathbf{x}, t)$ is and $\mathbf{x} \mapsto \frac{\partial u}{\partial t}(\mathbf{x}, t)$ are absolutely integrable on $\mathbb{R}^n$ for all $t \geq 0$. We may then use Fourier transform with respect to $\mathbf{x}$ to transform (10.1.12) and (10.1.13) into

$$\frac{\partial \hat{u}}{\partial t}(\mathbf{y}, t) + \|\mathbf{y}\|^2 \hat{u}(\mathbf{y}, t) = 0 \quad , \quad \mathbf{y} \in \mathbb{R}^n \text{ and } t \geq 0 ,$$

$$\hat{u}(\mathbf{y}, 0) = \hat{f}(\mathbf{y}) \quad , \quad \mathbf{y} \in \mathbb{R}^n .$$

The solution of this initial value problem is

$$\hat{u}(\mathbf{y}, t) = \hat{f}(\mathbf{y}) e^{-\|\mathbf{y}\|^2 t}$$

for $\mathbf{y} \in \mathbb{R}^n$ and $t \geq 0$. Using the Inverse Fourier transform, we get from Proposition 6.6.3 that

$$u(\mathbf{x}, t) = (f * K_t)(\mathbf{x}) \tag{10.1.14}$$

for $\mathbf{x} \in \mathbb{R}^n$ and $t \geq 0$, where the convolution is with respect to $\mathbf{x}$ only and K_t is the inverse Fourier transform of $\mathbf{y} \mapsto (2\pi)^{-n/2} e^{-\|\mathbf{y}\|^2 t}$ ¹. If we use (6.6.3) in Example 6.6.2 with $a = 1/(4t)$, we find that

$$K_t(\mathbf{x}) = (4\pi t)^{-n/2} e^{-\|\mathbf{x}\|^2/(4t)}$$

for $\mathbf{x} \in \mathbb{R}^n$ and $t > 0$. Proceeding as we did for the one dimensional heat equation, we can prove that

$$u(\mathbf{x}, t) = (f * K_t)(\mathbf{x}) = (4\pi t)^{-n/2} \int_{\mathbb{R}^n} e^{-\|\mathbf{x}-\mathbf{y}\|^2/(4t)} f(\mathbf{y}) d\mathbf{y}$$

is a classical solution of the heat equation given by (10.1.12) and (10.1.13).

This discussion motivates the introduction of the function

$$K(\mathbf{x}, t) = \begin{cases} (4\pi t)^{-n/2} e^{-\|\mathbf{x}\|^2/(4t)} & \text{if } t > 0 \\ 0 & \text{if } t \leq 0 \end{cases}$$

$K(\mathbf{x}, t)$ has many names; it is called **Gaussian kernel**, **Gauss-Weierstrass kernel** or **heat kernel**.

Can we use (10.1.14) even when $f \notin \mathcal{S}(\mathbb{R}^n)$? We answer this question in the next theorem.

Theorem 10.1.3

Assume that $f \in L^p(\mathbb{R}^n)$, where $1 \leq p \leq \infty$. Then

$$u(\mathbf{x}, t) = (f * K_t)(\mathbf{x}) \quad , \quad (\mathbf{x}, t) \in \mathbb{R}^n \times [0, \infty[\quad , \tag{10.1.15}$$

is a solution of (10.1.12).

If f is bounded and continuous, then u can be extended continuously to $\mathbb{R}^n \times [0, \infty[$ with $u(\mathbf{x}, 0) = f(\mathbf{x})$ for $\mathbf{x} \in \mathbb{R}^n$.

If $f \in L^p(\mathbb{R}^n)$, where $1 \leq p < \infty$, then $v_t(\mathbf{x}) = u(\mathbf{x}, t)$ for $\mathbf{x} \in \mathbb{R}^n$ satisfies $v_t \rightarrow f$ in $L^p(\mathbb{R}^n)$ as $t \rightarrow 0$.

Proof.

The proof is similar to the proof of Theorem 9.5.11.

¹Recall that the integral used in the definition of convolution in Section 6.6 was with respect to the Lebesgue measure m divided by $(2\pi)^{n/2}$.

i) First, we prove that u is well defined. We have that $K_t \in L^q(\mathbb{R}^n)$ for $1 \leq q \leq \infty$ and $t > 0$. Hence, $\mathbf{y} \mapsto f(\mathbf{y}) K_t(\mathbf{x} - \mathbf{y})$ for $\mathbf{y} \in \mathbb{R}^n$ is in $L^1(\mathbb{R}^n)$. For instance, if $1 < p < \infty$ we have

$$\int_{\mathbb{R}^n} |f(\mathbf{y}) K_t(\mathbf{x} - \mathbf{y})| d\mathbf{y} \leq \left(\int_{\mathbb{R}^n} |f(\mathbf{y})|^p d\mathbf{y} \right)^{1/p} \left(\int_{\mathbb{R}^n} |K_t(\mathbf{x} - \mathbf{y})|^q d\mathbf{y} \right)^{1/q} < \infty$$

for $\frac{1}{p} + \frac{1}{q} = 1$. Thus $u : \mathbb{R}^n \times]0, \infty[\rightarrow \mathbb{R}$ is well defined.

ii) We prove that u satisfies the heat equation. Proceeding as before, we can show that $\mathbf{y} \mapsto f(\mathbf{y}) \left(D_{\mathbf{x}}^{\alpha} \frac{\partial^k}{\partial t^k} K_t(\mathbf{x} - \mathbf{y}) \right)$ for $\mathbf{y} \in \mathbb{R}^n$ is in $L^1(\mathbb{R}^n)$ for $t > 0$ whatever the choice of $k \geq 0$ and the multi-index $\alpha \in \mathbb{N}^n$. Hence, we may interchange the integral with respect to $\mathbf{y}$ with the partial derivatives of any order with respect to the x_i 's and t .

Therefore,

$$\left(\frac{\partial}{\partial t} - \Delta_{\mathbf{x}} \right) u(\mathbf{x}, t) = \int_{\mathbb{R}^n} f(\mathbf{y}) \left(\frac{\partial}{\partial t} - \Delta_{\mathbf{x}} \right) K_t(\mathbf{x} - \mathbf{y}) d\mathbf{y} = 0, \quad (\mathbf{x}, t) \in \mathbb{R}^n \times]0, \infty[,$$

because $\left(\frac{\partial}{\partial t} - \Delta_{\mathbf{x}} \right) K_t(\mathbf{x} - \mathbf{y}) = 0$ for all $\mathbf{x} \in \mathbb{R}^n$ and $t > 0$.

iii) For the boundary conditions, we first note that

$$\int_{\mathbb{R}^n} K_t(\mathbf{y}) d\mathbf{y} = (2\pi)^{n/2} \hat{K}_t(\mathbf{0}) = e^{-\|\mathbf{y}\|^2 t} \Big|_{\mathbf{y}=\mathbf{0}} = 1 \quad (10.1.16)$$

and

$$K_t(\mathbf{x}) = t^{-n/2} K_1(t^{-1/2} \mathbf{x}) \quad (10.1.17)$$

If $f \in L^p(\mathbb{R}^n)$ with $1 \leq p < \infty$, it follows from Proposition 6.5.3 that $f * K_t \rightarrow f$ in $L^p(\mathbb{R}^n)$ as $t \rightarrow 0$.

If f is continuous and bounded on $\mathbb{R}^n$, then f is uniformly continuous on any bounded open set V because $\overline{V}$ is a compact subset of $\mathbb{R}^n$. It follows from Proposition 6.5.3 that $f * K_t \rightarrow f$ uniformly on compact subsets of $\mathbb{R}^n$ as $t \rightarrow 0$. In particular, $u(\mathbf{x}, 0) = f(\mathbf{x})$ for $\mathbf{x} \in \mathbb{R}^n$ defines a continuous extension of u to $\mathbb{R}^n \times [0, \infty[$. ■

Proposition 10.1.4

Assume that $f \in L^p(\mathbb{R}^n)$, where $1 \leq p \leq \infty$. The solution u of (10.1.12) given in (10.1.15) is of class $C^\infty(\mathbb{R}^n \times]0, \infty[)$.

Proof.

We have seen in the proof of the previous theorem that we may interchange integration with respect to $\mathbf{y}$ in

$$u(\mathbf{x}, t) = \int_{\mathbb{R}^n} f(\mathbf{y}) K_t(\mathbf{x} - \mathbf{y}) d\mathbf{y}$$

with partial derivatives of any order with respect to the x_i 's and t . Since $K \in C^\infty(\mathbb{R}^n \times]0, \infty[)$, we get that $u \in C^\infty(\mathbb{R}^n \times]0, \infty[)$. ■

Remark 10.1.5

1. The solution of (10.1.12) and (10.1.13) given by (10.1.15) smooths out any initial condition in $L^p(\mathbb{R}^n)$.
2. We may relax the condition on f in the previous theorem if we assume that $|f(\mathbf{x})| \leq Ce^{\|\mathbf{x}\|^2/(4T)}$ for $\mathbf{x} \in \mathbb{R}^n$, and some positive constants C and T . One can show that the function

$$u : \mathbb{R}^n \times]0, T[\rightarrow \mathbb{R} \\ (\mathbf{x}, t) \mapsto (f * K_t)(\mathbf{x})$$

is a solution of the heat equation on $\mathbb{R}^n \times]0, T[$. In addition, if f is also continuous, then $u(\cdot, t) \rightarrow f(\cdot)$ uniformly on compact subsets of $\mathbb{R}^n$ as $t \rightarrow 0$.

3. If no conditions are imposed on the growth at infinity of the solution u of (10.1.12) and (10.1.13), then the solution is not unique. Consider

$$g(t) = \begin{cases} e^{-1/t^2} & \text{if } t \neq 0 \\ 0 & \text{if } t = 0 \end{cases}$$

The function g is in $C^\infty(\mathbb{R})$. The function

$$u(x, t) = \sum_{k=0}^{\infty} \frac{g^{(k)}(t)}{(2k)!} x^{2k} \quad (10.1.18)$$

for $x \in \mathbb{R}$ and $t \in \mathbb{R}$ is a solution of the heat equation $\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} = 0$. This follows from the fact that the series (10.1.18), $\sum_{k=0}^{\infty} \frac{g^{(k+1)}(t)}{(2k)!} x^{2k}$ and $\sum_{k=1}^{\infty} \frac{g^{(k)}(t)}{(2k-2)!} x^{2k-2}$ converge uniformly on compact sets. Hence, we may interchange summation and partial derivatives in (10.1.18). Since $g^{(k)}(0) = 0$ for all k , $u(x, 0) = 0$ for all $x \in \mathbb{R}$ but u is not the trivial solution of the heat equation.

♠

The following theorem gives conditions on the growth at infinity of the solution u of (10.1.12) and (10.1.13) to provide unicity of solutions.

Proposition 10.1.6

Suppose that $u \in C(\mathbb{R}^n \times [0, \infty[) \cap C^2(\mathbb{R}^n \times]0, \infty[)$ is a solution of $\frac{\partial u}{\partial t} - \Delta u = 0$ on $\mathbb{R}^n \times]0, \infty[$ and $u(\mathbf{x}, 0) = 0$ for $\mathbf{x} \in \mathbb{R}^n$. If, for every $\epsilon > 0$, there exists $C > 0$ such that $|u(\mathbf{x}, t)| \leq Ce^{\epsilon\|\mathbf{x}\|^2}$ and $\|\nabla_{\mathbf{x}} u(\mathbf{x}, t)\| \leq Ce^{\epsilon\|\mathbf{x}\|^2}$ for all $\mathbf{x}$ and t , then $u \equiv 0$.

Proof.

Given $(\mathbf{x}_0, t_0) \in \mathbb{R}^n \times]0, \infty[$, we prove that $u(\mathbf{x}_0, t_0) = 0$.

We defined two functions, $f : \mathbb{R}^n \times]0, \infty[\rightarrow \mathbb{R}$ and $g : \mathbb{R}^n \times]-\infty, t_0[\rightarrow \mathbb{R}$, as follows. $f(\mathbf{x}, t) = u(\mathbf{x}, t)$ for $(\mathbf{x}, t) \in \mathbb{R}^n \times]0, \infty[$ and $g(\mathbf{x}, t) = K(\mathbf{x} - \mathbf{x}_0, t_0 - t)$ for $(\mathbf{x}, t) \in \mathbb{R}^n \times]-\infty, t_0[$.

We have that $\frac{\partial f}{\partial t} - \Delta f = 0$ on $\mathbb{R}^n \times]0, \infty[$ and $\frac{\partial g}{\partial t} + \Delta g = 0$ on $\mathbb{R}^n \times]-\infty, t_0[$.

Given $0 < a < b < t_0$ and $r > 0$, let $\Omega = \{(\mathbf{x}, t) : \|\mathbf{x}\| < r, a < t < b\}$. Moreover, let $F : \mathbb{R}^n \times]0, t_0[\rightarrow \mathbb{R}^{n+1}$ be the vector field defined by

$$F(\mathbf{x}, t) = \sum_{j=1}^n \left(f(\mathbf{x}, t) \frac{\partial g}{\partial x_j}(\mathbf{x}, t) - g(\mathbf{x}, t) \frac{\partial f}{\partial x_j}(\mathbf{x}, t) \right) \mathbf{e}_j + f(\mathbf{x}, t) g(\mathbf{x}, t) \mathbf{e}_{n+1}.$$

Since

$$\begin{aligned} \operatorname{div}_{\mathbf{x}, t} F(\mathbf{x}, t) &= \sum_{j=1}^n \frac{\partial}{\partial x_j} \left(f(\mathbf{x}, t) \frac{\partial g}{\partial x_j}(\mathbf{x}, t) - g(\mathbf{x}, t) \frac{\partial f}{\partial x_j}(\mathbf{x}, t) \right) + \frac{\partial}{\partial t} (f(\mathbf{x}, t) g(\mathbf{x}, t)) \\ &= f(\mathbf{x}, t) \left(\frac{\partial g}{\partial t}(\mathbf{x}, t) + \Delta g(\mathbf{x}, t) \right) + g(\mathbf{x}, t) \left(\frac{\partial f}{\partial t}(\mathbf{x}, t) - \Delta f(\mathbf{x}, t) \right) = 0 \end{aligned}$$

in $\mathbb{R}^n \times]0, t_0[$, we get from the divergence theorem that

$$\begin{aligned} 0 &= \int_{\Omega} \operatorname{div}_{\mathbf{x}, t} F(\mathbf{x}, t) \, d\mathbf{x} \, dt = \int_{\partial\Omega} F \cdot \nu \, dS \\ &= \int_{\|\mathbf{x}\| \leq r} f(\mathbf{x}, b) g(\mathbf{x}, b) \, d\mathbf{x} - \int_{\|\mathbf{x}\| \leq r} f(\mathbf{x}, a) g(\mathbf{x}, a) \, d\mathbf{x} \\ &\quad + \int_{\|\mathbf{x}\|=r} \int_a^b \sum_{j=1}^n \left(f(\mathbf{x}, t) \frac{\partial g}{\partial x_j}(\mathbf{x}, t) - g(\mathbf{x}, t) \frac{\partial f}{\partial x_j}(\mathbf{x}, t) \right) \frac{x_j}{r} \, dt \, dS_{\mathbf{x}}. \end{aligned} \tag{10.1.19}$$

The last integral in (10.1.19) converges to 0 as $r \rightarrow \infty$ because, given $0 < \epsilon < 1/(4(t_0 - a))$, there exist $C > 0$ such that

$$\begin{aligned} &\int_{\|\mathbf{x}\|=r} \int_a^b \left| f(\mathbf{x}, t) \frac{\partial g}{\partial x_j}(\mathbf{x}, t) \right| \underbrace{\left| \frac{x_j}{r} \right|}_{\leq 1} \, dt \, dS_{\mathbf{x}} \\ &\leq C e^{\epsilon r^2} \frac{2\pi}{(4\pi(t_0 - b))^{1+n/2}} (r + \|\mathbf{x}_0\|) e^{-(r - \|\mathbf{x}_0\|)^2/(4(t_0 - a))} (b - a) r^{n-1} \omega_n \rightarrow 0 \text{ as } r \rightarrow \infty \end{aligned}$$

and

$$\begin{aligned} &\int_{\|\mathbf{x}\|=r} \int_a^b \left| g(\mathbf{x}, t) \frac{\partial f}{\partial x_j}(\mathbf{x}, t) \right| \underbrace{\left| \frac{x_j}{r} \right|}_{\leq 1} \, dt \, dS_{\mathbf{x}} \\ &\leq C e^{\epsilon r^2} \frac{1}{(4\pi(t_0 - b))^{n/2}} e^{-(r - \|\mathbf{x}_0\|)^2/(4(t_0 - a))} (b - a) r^{n-1} \omega_n \rightarrow 0 \text{ as } r \rightarrow \infty. \end{aligned}$$

Hence, $r \rightarrow \infty$ in (10.1.19) yields

$$(u(\cdot, b) * K_{t_0-b}(\cdot))(\mathbf{x}_0) - (u(\cdot, a) * K_{t_0-a}(\cdot))(\mathbf{x}_0) = 0 \tag{10.1.20}$$

because $g(\mathbf{x}, t) = K(\mathbf{x} - \mathbf{x}_0, t_0 - t) = K(\mathbf{x}_0 - \mathbf{x}, t_0 - t)$.

i) We claim that $(u(\cdot, a) * K_{t_0-a}(\cdot))(\mathbf{x}_0) \rightarrow 0$ as $a \rightarrow 0^+$. We have that

$$\lim_{a \rightarrow 0^+} u(\mathbf{x}, a) K(\mathbf{x} - \mathbf{x}_0, t_0 - a) = u(\mathbf{x}, 0) K(\mathbf{x} - \mathbf{x}_0, t_0) = 0$$

for all $\mathbf{x} \in \mathbb{R}^n$ because $u(\mathbf{x}, 0) = 0$ for all $\mathbf{x} \in \mathbb{R}^n$. Moreover, given $\epsilon < 1/(4t_0)$, there exist $C > 0$ such that

$$|u(\mathbf{x}, a) K(\mathbf{x} - \mathbf{x}_0, t_0 - a)| < C e^{\epsilon \|\mathbf{x}\|^2} \frac{e^{-\|\mathbf{x}-\mathbf{x}_0\|^2/(4t_0)}}{(\pi t_0)^{n/2}} \in L^1(\mathbb{R}^n)$$

for $0 < a < t_0/2$. We then get from the Lebesgue Dominated Convergence Theorem that

$$(u(\cdot, a) * K_{t_0-a}(\cdot))(\mathbf{x}_0) \rightarrow 0 \quad \text{as } a \rightarrow 0^+. \quad (10.1.21)$$

ii) Our second claim is that $(u(\cdot, b) * K_{t_0-b}(\cdot))(\mathbf{x}_0) \rightarrow u(\mathbf{x}_0, t_0)$ as $b \rightarrow t_0^-$.

We have that

$$\begin{aligned} |u(\cdot, b) * K_{t_0-b}(\cdot)(\mathbf{x}_0) - u(\mathbf{x}_0, t_0)| &\leq \int_{\mathbb{R}^n} |u(\mathbf{x}, b) - u(\mathbf{x}_0, b)| K_{t_0-b}(\mathbf{x} - \mathbf{x}_0) d\mathbf{x} \\ &\quad + \int_{\mathbb{R}^n} |u(\mathbf{x}_0, b) - u(\mathbf{x}_0, t_0)| K_{t_0-b}(\mathbf{x} - \mathbf{x}_0) d\mathbf{x}. \end{aligned} \quad (10.1.22)$$

Given $\mathbf{x} \in \mathbb{R}^n$, there exists $\theta \in]0, 1[$ such that

$$|u(\mathbf{x}, b) - u(\mathbf{x}_0, b)| \leq \|\nabla_{\mathbf{x}} u(\mathbf{x}_0 + \theta(\mathbf{x} - \mathbf{x}_0), b)\| \|\mathbf{x} - \mathbf{x}_0\|$$

according to the Mean Value Theorem for functions of several variables. Therefore, given $\epsilon > 0$, there exists $C > 0$ such that

$$|u(\mathbf{x}, b) - u(\mathbf{x}_0, b)| \leq C e^{\epsilon \|\mathbf{x}_0 + \theta(\mathbf{x} - \mathbf{x}_0)\|^2} \|\mathbf{x} - \mathbf{x}_0\|$$

by hypothesis. Choose $0 < c < t_0$ and $\epsilon < 1/(8(t_0 - c))$. Then,

$$\begin{aligned} |u(\mathbf{x}, b) - u(\mathbf{x}_0, b)| K_{t_0-b}(\mathbf{x} - \mathbf{x}_0) &\leq \frac{C}{(4\pi(t_0 - b))^{n/2}} e^{\epsilon \|\mathbf{x}_0 + \theta(\mathbf{x} - \mathbf{x}_0)\|^2} \|\mathbf{x} - \mathbf{x}_0\| e^{-\|\mathbf{x}-\mathbf{x}_0\|^2/(4(t_0-b))} \\ &\leq \frac{C}{(4\pi)^{n/2}} \underbrace{\left(\frac{1}{(t_0 - b)^{n/2}} e^{-\|\mathbf{x}-\mathbf{x}_0\|^2/(8(t_0-b))} \right)}_{\leq (4n)^{n/2} e^{-n/2} / \|\mathbf{x}-\mathbf{x}_0\|^n} e^{\epsilon \|\mathbf{x}_0 + \theta(\mathbf{x} - \mathbf{x}_0)\|^2} \|\mathbf{x} - \mathbf{x}_0\| e^{-\|\mathbf{x}-\mathbf{x}_0\|^2/(8(t_0-c))} \\ &\leq \frac{C n^{n/2} e^{-n/2}}{\pi^{n/2}} \|\mathbf{x} - \mathbf{x}_0\|^{n-1} e^{\epsilon \max\{\|\mathbf{x}_0\|^2, \|\mathbf{x}\|^2\}} e^{-\|\mathbf{x}-\mathbf{x}_0\|^2/(8(t_0-c))} \in L^1(\mathbb{R}^n) \end{aligned}$$

for all $b \in]c, t_0[$. Since $|u(\mathbf{x}, b) - u(\mathbf{x}_0, b)| K_{t_0-b}(\mathbf{x} - \mathbf{x}_0) \rightarrow 0$ as $b \rightarrow t_0^-$ for all $\mathbf{x} \neq \mathbf{x}_0$, we get from the Lebesgue Dominate Convergence Theorem that

$$\int_{\mathbb{R}^n} |u(\mathbf{x}, b) - u(\mathbf{x}_0, b)| K_{t_0-b}(\mathbf{x} - \mathbf{x}_0) d\mathbf{x} \rightarrow 0 \quad \text{as } b \rightarrow t_0^-. \quad (10.1.23)$$

Moreover,

$$\int_{\mathbb{R}^n} |u(\mathbf{x}_0, b) - u(\mathbf{x}_0, t_0)| K_{t_0-b}(\mathbf{x} - \mathbf{x}_0) d\mathbf{x} = |u(\mathbf{x}_0, b) - u(\mathbf{x}_0, t_0)| \int_{\mathbb{R}^n} K_{t_0-b}(\mathbf{x} - \mathbf{x}_0) d\mathbf{x}$$

$$= |u(\mathbf{x}_0, b) - u(\mathbf{x}_0, t_0)| \rightarrow 0 \quad \text{as } b \rightarrow t_0^- . \quad (10.1.24)$$

It follows from (10.1.22), (10.1.23) and (10.1.24) that

$$(u(\cdot, b) * K_{t_0-b}(\cdot))(\mathbf{x}_0) \rightarrow u(\mathbf{x}_0, t_0) \quad \text{as } b \rightarrow t_0^- . \quad (10.1.25)$$

This complete the prove of the second claim.

iii) If we let $a \rightarrow 0^+$ and $b \rightarrow t_0^-$ in (10.1.20), we get from (10.1.21) and (10.1.25) that $u(\mathbf{x}_0, t_0) = 0$. ■

10.2 Fundamental Solution

As before, let $K : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$ be the function defined by

$$K(\mathbf{x}, t) = \begin{cases} (4\pi t)^{-n/2} e^{-\|\mathbf{x}\|^2/(4t)} & \text{if } t > 0 \\ 0 & \text{if } t \leq 0 \end{cases}$$

We now generalize the discussion of the previous section to the solution of the heat equation in the sense of distributions.

Theorem 10.2.1

K is a fundamental solution of the heat equation $\frac{\partial u}{\partial t} - \Delta u = 0$ on $\mathbb{R}^n \times \mathbb{R}$.

Proof.

i) We prove that $K \in \mathcal{D}'(\mathbb{R}^n \times \mathbb{R})$ by proving that $K \in L^1_{loc}(\mathbb{R}^n \times \mathbb{R})$. We prove in fact a stronger result. Since $\int_{\mathbb{R}^n} |K(\mathbf{x}, t)| d\mathbf{x} = 1$ for all $t > 0$, we get

$$\int_I \int_{\mathbb{R}^n} |K(\mathbf{x}, t)| d\mathbf{x} dt \leq \int_I dt < \infty$$

for any bounded interval $I \subset \mathbb{R}$. Thus K defines a proper distribution on $\mathbb{R}^n \times \mathbb{R}$.

ii) We show that $\frac{\partial K}{\partial t} - \Delta K = \delta$. For $\epsilon > 0$, let

$$K^{[\epsilon]}(\mathbf{x}, t) = \begin{cases} K(\mathbf{x}, t) & \text{if } t > \epsilon, \mathbf{x} \in \mathbb{R}^n \\ 0 & \text{if } t \leq \epsilon, \mathbf{x} \in \mathbb{R}^n \end{cases}$$

It is easy to show, as we have done in (i), that $K^{[\epsilon]} \in \mathcal{D}'(\mathbb{R}^n \times \mathbb{R})$.

Given $\phi \in \mathcal{D}(\mathbb{R}^n \times \mathbb{R})$, we have that $K^{[\epsilon]}(\mathbf{x}, t)\phi(\mathbf{x}, t) \rightarrow K(\mathbf{x}, t)\phi(\mathbf{x}, t)$ as $\epsilon \rightarrow 0^+$ for all $(\mathbf{x}, t) \in \mathbb{R}^n \times \mathbb{R}$. Moreover, since ϕ has a compact support, $M = \sup_{(\mathbf{x}, t) \in \mathbb{R}^n \times \mathbb{R}} |\phi(\mathbf{x}, t)| < \infty$ and there exists $I \subset \mathbb{R}$ such that $\text{supp } \phi \subset \mathbb{R}^n \times I$. Since

$$|K^{[\epsilon]}(\mathbf{x}, t)\phi(\mathbf{x}, t)| \leq M \chi_{\mathbb{R}^n \times I}(\mathbf{x}, t) K(\mathbf{x}, t)$$

for all $(\mathbf{x}, t) \in \mathbb{R}^n \times \mathbb{R}$ and $M\mathcal{X}_{\mathbb{R}^n \times I}K \in L^1(\mathbb{R}^n \times \mathbb{R})$, we get from the Lebesgue Dominate Convergence Theorem that

$$\int_{\mathbb{R}} \int_{\mathbb{R}^n} K^{[\epsilon]}(\mathbf{x}, t) \phi(\mathbf{x}, t) d\mathbf{x} dt \rightarrow \int_{\mathbb{R}} \int_{\mathbb{R}^n} K(\mathbf{x}, t) \phi(\mathbf{x}, t) d\mathbf{x} dt \quad \text{as } \epsilon \rightarrow 0^+.$$

Since $\phi \in \mathcal{D}(\mathbb{R}^n \times \mathbb{R})$ is arbitrary, we have that $K^{[\epsilon]} \rightarrow K$ in $\mathcal{D}'(\mathbb{R}^n \times \mathbb{R})$ as $\epsilon \rightarrow 0^+$.

To complete the proof, we only have to show that $\frac{\partial}{\partial t} K^{[\epsilon]} - \Delta_{\mathbf{x}} K^{[\epsilon]} \rightarrow \delta$ in $\mathcal{D}'(\mathbb{R}^n \times \mathbb{R})$ as $\epsilon \rightarrow 0^+$. Namely, for $\phi \in \mathcal{D}(\mathbb{R}^n \times \mathbb{R})$, we have to show that

$$\left\langle K^{[\epsilon]}, -\frac{\partial}{\partial t} \phi - \Delta_{\mathbf{x}} \phi \right\rangle \rightarrow \phi(\mathbf{0}, 0) \quad \text{as } \epsilon \rightarrow 0^+.$$

Integration by parts yields

$$\begin{aligned} \left\langle K^{[\epsilon]}, -\frac{\partial}{\partial t} \phi - \Delta_{\mathbf{x}} \phi \right\rangle &= \int_{\epsilon}^{\infty} \int_{\mathbb{R}^n} K(\mathbf{x}, t) \left(-\frac{\partial \phi}{\partial t}(\mathbf{x}, t) - \Delta_{\mathbf{x}} \phi(\mathbf{x}, t) \right) d\mathbf{x} dt \\ &= \int_{\epsilon}^{\infty} \int_{\mathbb{R}^n} \left(\frac{\partial K}{\partial t}(\mathbf{x}, t) - \Delta_{\mathbf{x}} K(\mathbf{x}, t) \right) \phi(\mathbf{x}, t) d\mathbf{x} dt + \int_{\mathbb{R}^n} K(\mathbf{x}, \epsilon) \phi(\mathbf{x}, \epsilon) d\mathbf{x} \end{aligned}$$

because ϕ has a compact support in $\mathbb{R}^n \times \mathbb{R}$. Since $\frac{\partial K}{\partial t}(\mathbf{x}, t) - \Delta_{\mathbf{x}} K(\mathbf{x}, t) = 0$ and $K(\mathbf{x}, t) = K(-\mathbf{x}, t)$ for $\mathbf{x} \in \mathbb{R}^n$ and $t > 0$, we have that

$$\left\langle K^{[\epsilon]}, -\frac{\partial}{\partial t} \phi - \Delta_{\mathbf{x}} \phi \right\rangle = \int_{\mathbb{R}^n} K(-\mathbf{x}, \epsilon) \phi(\mathbf{x}, \epsilon) d\mathbf{x} = (K(\cdot, \epsilon) * \phi(\cdot, \epsilon))(\mathbf{0}).$$

Finally, we prove that

$$\begin{aligned} (K(\cdot, \epsilon) * \phi(\cdot, \epsilon))(\mathbf{0}) &= (K(\cdot, \epsilon) * \phi(\cdot, 0))(\mathbf{0}) + (K(\cdot, \epsilon) * (\phi(\cdot, \epsilon) - \phi(\cdot, 0)))(\mathbf{0}) \\ &\rightarrow \phi(\mathbf{0}, 0) \quad \text{as } \epsilon \rightarrow 0^+. \end{aligned}$$

According to (10.1.16), (10.1.17) and Proposition 6.5.3, we have that

$$K(\cdot, \epsilon) * \phi(\cdot, 0) \rightarrow \phi(\cdot, 0) \quad \text{as } \epsilon \rightarrow 0^+.$$

Thus

$$(K(\cdot, \epsilon) * \phi(\cdot, 0))(\mathbf{0}) \rightarrow \phi(\mathbf{0}, 0) \quad \text{as } \epsilon \rightarrow 0^+.$$

Since $\phi : \mathbb{R}^n \times \mathbb{R}$ is a continuous function with compact support, ϕ is uniformly continuous. In particular, $\phi(\cdot, \epsilon) - \phi(\cdot, 0) \rightarrow 0$ uniformly on $\mathbb{R}^n$. Since $K(\mathbf{x}, \epsilon) > 0$ for all $\mathbf{x} \in \mathbb{R}^n$, it follows from (10.1.16) that $\|K(\cdot, \epsilon)\|_1 = 1$ for $\epsilon > 0$. Hence

$$\begin{aligned} |(K(\cdot, \epsilon) * (\phi(\cdot, \epsilon) - \phi(\cdot, 0)))(\mathbf{0})| &\leq \|K(\cdot, \epsilon)\|_1 \|\phi(\cdot, \epsilon) - \phi(\cdot, 0)\|_{\infty} \\ &= \|\phi(\cdot, \epsilon) - \phi(\cdot, 0)\|_{\infty} \rightarrow 0 \quad \text{as } \epsilon \rightarrow 0^+. \end{aligned} \quad \blacksquare$$

Corollary 10.2.2

The heat operator $L(\mathbf{x}, y, D) = \frac{\partial}{\partial t} - \Delta_{\mathbf{x}}$ is hypoelliptic.

Proof.

This follows from Theorem 6.7.11 because $K \in C^\infty((\mathbb{R}^n \times \mathbb{R}) \setminus \{(\mathbf{0}, 0)\})$. ■

If $f \in \mathcal{E}'(\mathbb{R}^n \times \mathbb{R})$, then $u = K * f$ is a solution (in the sense of distributions) of the non-homogeneous heat equation $\frac{\partial u}{\partial t} - \Delta_{\mathbf{x}} u = f$ on $\mathbb{R}^n \times \mathbb{R}$. The conclusion of the next theorem states that $K * f$ can define a function almost everywhere if f is nice enough.

Theorem 10.2.3

If $f \in L^1(\mathbb{R}^n \times \mathbb{R})$, then $u = f * K$ is well defined almost everywhere and $\frac{\partial u}{\partial t} - \Delta_{\mathbf{x}} u = f$ in the sense of distribution.

Proof.

Since $f \in L^1(\mathbb{R}^n \times \mathbb{R})$, we have by Fubini's Theorem that $f(\cdot, s) \in L^1(\mathbb{R}^n)$ for almost all $s \in \mathbb{R}$. We have also seen that $\|K(\cdot, t-s)\|_1 = 1$ for all $s < t$. Hence, from Young's Inequality, Theorem 6.5.2, we have

$$\| |f|(\cdot, s) * K(\cdot, t-s) \|_1 \leq \|f(\cdot, s)\|_1 \|K(\cdot, t-s)\|_1 = \|f(\cdot, s)\|_1 < \infty$$

for almost all s such that $s < t$. Thus $\mathbf{x} \mapsto (|f|(\cdot, s) * K(\cdot, t-s))(\mathbf{x})$ for $\mathbf{x} \in \mathbb{R}^n$ is in $L^1(\mathbb{R}^n)$ for almost all s such that $s < t$. The convolution is obviously on $\mathbb{R}^n$. We then have

$$\begin{aligned} |u(\mathbf{x}, t)| &\leq \int_{-\infty}^{\infty} \int_{\mathbb{R}^n} |f|(\mathbf{y}, s) K(\mathbf{x} - \mathbf{y}, t-s) d\mathbf{y} ds = \int_{-\infty}^t \int_{\mathbb{R}^n} |f|(\mathbf{y}, s) K(\mathbf{x} - \mathbf{y}, t-s) d\mathbf{y} ds \\ &= \int_{-\infty}^t (|f|(\cdot, s) * K(\cdot, t-s))(\mathbf{x}) ds \end{aligned}$$

because $K(\mathbf{z}, t-s) = 0$ for $\mathbf{z} \in \mathbb{R}^n$ and $s \geq t$. Hence,

$$\begin{aligned} \int_{\mathbb{R}^n} |u(\mathbf{x}, t)| d\mathbf{x} &\leq \int_{-\infty}^t \int_{\mathbb{R}^n} (|f|(\cdot, s) * K(\cdot, t-s))(\mathbf{x}) d\mathbf{x} ds = \int_{-\infty}^t \| |f|(\cdot, s) * K(\cdot, t-s) \|_1 ds \\ &\leq \int_{-\infty}^t \|f(\cdot, s)\|_1 ds = \int_{-\infty}^t \int_{\mathbb{R}^n} |f(\mathbf{x}, s)| d\mathbf{x} ds \leq \|f\|_1 < \infty, \end{aligned}$$

where the last L^1 -norm is on $L^1(\mathbb{R}^n \times \mathbb{R})$. It follows that $u(\mathbf{x}, t)$ exists for all $t \in \mathbb{R}$ and for almost every $\mathbf{x} \in \mathbb{R}^n$.

The rest of the proof is similar to the proof of Theorem 9.3.3 with N and Δ in that proof replaced by K and $\frac{\partial}{\partial t} - \Delta_{\mathbf{x}}$ respectively. ■

10.3 Maximum Principle

Theorem 10.3.1 (Weak Maximum Principle)

Let Ω be a bounded open subset of $\mathbb{R}^n$. Suppose that $u \in C(\overline{\Omega} \times [0, T])$ is a solution of the heat equation $\frac{\partial u}{\partial t} - \Delta u = 0$ in $\Omega \times]0, T[$. Then, u reaches its maximum on $(\Omega \times \{0\}) \cup (\partial\Omega \times [0, T])$.

Proof.

Since $\frac{\partial}{\partial t} - \Delta_{\mathbf{x}}$ is hypoelliptic, we have that $u \in C^\infty(\Omega \times]0, T[)$.

Consider $\epsilon > 0$ and T' such that $0 < T' < T$. let $v(\mathbf{x}, t) = u(\mathbf{x}, t) + \epsilon \|\mathbf{x}\|^2$ for $(\mathbf{x}, t) \in \overline{\Omega} \times [0, T]$. We have

$$\frac{\partial v}{\partial t} - \Delta_{\mathbf{x}} v = -2n\epsilon < 0 \quad , \quad (\mathbf{x}, t) \in \Omega \times]0, T[\quad . \quad (10.3.1)$$

If the maximum of v on $\overline{\Omega} \times [0, T']$ is $(\mathbf{x}_0, t_0) \in \Omega \times]0, T'[$, then $(\mathbf{x}_0, t_0)$ is a critical point and $\nabla_{\mathbf{x}, t} v(\mathbf{x}_0, t_0) = \mathbf{0} \in \mathbb{R}^n \times \mathbb{R}$. Moreover, $\frac{\partial^2 v}{\partial x_j^2}(\mathbf{x}_0, t_0) \leq 0$ for $1 \leq j \leq n$ because, in each x_j -direction, v has a local maximum or v is constant near $(\mathbf{x}_0, t_0)$. Hence, $\frac{\partial v}{\partial t}(\mathbf{x}_0, t_0) = 0$ and $\Delta_{\mathbf{x}} v(\mathbf{x}_0, t_0) \leq 0$ implies that $\frac{\partial v}{\partial t} - \Delta_{\mathbf{x}} v \geq 0$. This contradicts (10.3.1). We need to add the term $\epsilon \|\mathbf{x}\|^2$ to u to get the strict inequality in (10.3.1).

If the maximum of v on $\overline{\Omega} \times [0, T']$ is $(\mathbf{x}_0, T')$ with $\mathbf{x}_0 \in \Omega$, we have that $\frac{\partial v}{\partial t}(\mathbf{x}_0, T') \geq 0$ because of the result in the previous paragraph; namely, v is constant or increasing in the direction of t . We still have $\frac{\partial^2 v}{\partial x_j^2}(\mathbf{x}_0, T') \leq 0$ for all j for the same reason than before. We need to introduce $T' < T$ because there is no guarantee that the second order partial derivatives $\frac{\partial^2 u}{\partial x_j^2}(\mathbf{x}, t)$ exist for $t = T$. Hence, $\frac{\partial v}{\partial t}(\mathbf{x}_0, T') \geq 0$ and $\Delta_{\mathbf{x}} v(\mathbf{x}_0, T') \leq 0$ implies that $\frac{\partial v}{\partial t} - \Delta_{\mathbf{x}} v \geq 0$. This again contradicts (10.3.1).

We have found that

$$\begin{aligned} \max\{u(\mathbf{x}, t) : (\mathbf{x}, t) \in \overline{\Omega} \times [0, T']\} &\leq \max\{v(\mathbf{x}, t) : (\mathbf{x}, t) \in \overline{\Omega} \times [0, T']\} \\ &= \max\{v(\mathbf{x}, t) : (\mathbf{x}, t) \in (\Omega \times \{0\}) \cup (\partial\Omega \times [0, T'])\} \\ &\leq \max\{u(\mathbf{x}, t) : (\mathbf{x}, t) \in (\Omega \times \{0\}) \cup (\partial\Omega \times [0, T'])\} + \epsilon \max\{\|\mathbf{x}\|^2 : \mathbf{x} \in \overline{\Omega}\} . \end{aligned}$$

If $\epsilon \rightarrow 0$, we get

$$\max\{u(\mathbf{x}, t) : (\mathbf{x}, t) \in \overline{\Omega} \times [0, T']\} \leq \max\{u(\mathbf{x}, t) : (\mathbf{x}, t) \in (\Omega \times \{0\}) \cup (\partial\Omega \times [0, T'])\}$$

for all $T' < T$. Thus

$$\max\{u(\mathbf{x}, t) : (\mathbf{x}, t) \in \overline{\Omega} \times [0, T]\} \leq \max\{u(\mathbf{x}, t) : (\mathbf{x}, t) \in (\Omega \times \{0\}) \cup (\partial\Omega \times [0, T])\} . \quad \blacksquare$$

Corollary 10.3.2

There is only one solution $u \in C(\overline{\Omega} \times [0, T])$ (if any) of

$$\begin{aligned} \frac{\partial u}{\partial t}(\mathbf{x}, t) - \Delta_{\mathbf{x}}(\mathbf{x}, t)u &= 0 \quad , \quad (\mathbf{x}, t) \in \Omega \times]0, T[\\ u(\mathbf{x}, 0) &= f(\mathbf{x}) \quad , \quad \mathbf{x} \in \Omega \\ u(\mathbf{x}, t) &= g(\mathbf{x}, t) \quad , \quad (\mathbf{x}, t) \in \partial\Omega \times]0, T[\end{aligned} \quad (10.3.2)$$

Proof.

If u_1 and u_2 are two solutions of (10.3.2) in $C(\overline{\Omega} \times [0, T])$, we may apply the weak maximum principle above to $u_1 - u_2$ and $u_2 - u_1$ to conclude that $u_1 - u_2 \leq 0$ and $u_2 - u_1 \leq 0$ on $\Omega \times [0, T]$ respectively. Thus $u_1 = u_2$ on $\Omega \times [0, T]$. ■

10.4 Separation of Variables

10.4.1 Cartesian Coordinates

Example 10.4.1

Find the solution of the heat equation

$$\frac{\partial u}{\partial t} = c^2 \frac{\partial^2 u}{\partial x^2} \quad , \quad 0 < x < \pi \text{ and } t > 0, \quad (10.4.1)$$

with $c^2 = 4$, the boundary conditions $u(0, t) = u(\pi, t) = 0$ for $t > 0$ and the initial condition

$$u(x, 0) = f(x) = \begin{cases} x & 0 < x \leq \pi/2 \\ \pi - x & \pi/2 < x < \pi \end{cases} \quad (10.4.2)$$

Separation of Variables: If we substitute $u(x, t) = F(x)G(t)$ in (10.4.1), we get

$$F(x) \frac{dG}{dt}(t) = 4 \frac{d^2 F}{dx^2}(x) G(t) \quad , \quad t > 0 \text{ and } 0 < x < \pi .$$

Thus, after dividing both sides by $4F(x)G(t)$, we get

$$\frac{1}{4G(t)} \frac{dG}{dt}(t) = \frac{1}{F(x)} \frac{d^2 F}{dx^2}(x) \quad , \quad t > 0 \text{ and } 0 < x < \pi .$$

Since the right hand side is independent of t and the left hand side is independent of x , we get

$$\frac{1}{4G(t)} \frac{dG}{dt}(t) = \frac{1}{F(x)} \frac{d^2 F}{dx^2}(x) = k \quad , \quad t > 0 \text{ and } 0 < x < \pi ,$$

for some constant k . We end up with two ordinary differential equations.

$$\frac{d^2 F}{dx^2}(x) - kF(x) = 0 \quad \text{and} \quad \frac{dG}{dt}(t) - 4kG(t) = 0. \quad (10.4.3)$$

Two boundary conditions are associated to the first ordinary differential equation in (10.4.3). From $u(0, t) = 0$, we get $F(0)G(t) = 0$. Since we assume that G is not null, we get $F(0) = 0$. Similarly, from $u(\pi, t) = 0$, we get $F(\pi)G(t) = 0$. Again, since we assume that G is not null, we get $F(\pi) = 0$. The boundary conditions for the first ordinary differential equation in (10.4.3) are $F(0) = F(\pi) = 0$.

Simple Functions: We first consider the boundary value problem

$$\frac{d^2 F}{dx^2}(x) - kF(x) = 0 \quad , \quad 0 \leq x \leq \pi, \quad (10.4.4)$$

with $F(0) = F(\pi) = 0$. The form of the general solution of this ordinary differential equation is determined by the roots of the characteristic equation $\lambda^2 - k = 0$.

If $k > 0$, the roots of the characteristic equation are $\pm\sqrt{k}$. Since the roots are real, the solution of the ordinary differential equation is of the form $F(x) = Ae^{\sqrt{k}x} + Be^{-\sqrt{k}x}$. However, $F(0) = 0$ implies that $A + B = 0$ and $F(\pi) = 0$ implies $Ae^{\pi\sqrt{k}} + Be^{-\pi\sqrt{k}} = 0$. The only solution of the system

$$\begin{pmatrix} 1 & 1 \\ e^{\pi\sqrt{k}} & e^{-\pi\sqrt{k}} \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

is $A = B = 0$ because

$$\det \begin{pmatrix} 1 & 1 \\ e^{\pi\sqrt{k}} & e^{-\pi\sqrt{k}} \end{pmatrix} = e^{-\pi\sqrt{k}} - e^{\pi\sqrt{k}} = e^{-\pi\sqrt{k}}(1 - e^{2\pi\sqrt{k}}) \neq 0$$

for $k \neq 0$. Therefore, the null solution is the only solution of the boundary value problem (10.4.4) for $k > 0$.

If $k = 0$, the solution of the ordinary differential equation is $F(x) = Ax + B$. However, $F(0) = 0$ implies that $B = 0$. Hence, $F(\pi) = 0$ implies that $A\pi = 0$. Thus $A = B = 0$ and again the null solution is the only solution of the boundary value problem (10.4.4) for $k = 0$.

If $k < 0$, the roots of the characteristic equation are $\pm i\sqrt{-k}$. Since the roots are complex, the solution of the ordinary differential equation is of the form $F(x) = A \cos(\sqrt{-k}x) + B \sin(\sqrt{-k}x)$. The condition $F(0) = 0$ implies that $A = 0$. Hence, $F(\pi) = 0$ implies that $B \sin(\pi\sqrt{-k}) = 0$. If we take $B = 0$, we get the null solution. We must therefore have $\sin(\pi\sqrt{-k}) = 0$. This implies that $k = k_n \equiv -n^2$ for $n > 0$. The boundary value problem (10.4.4) has non-null solutions only for $k = k_n = -n^2 < 0$ with n a positive integer, and the solutions associated to k_n are of the form $F(x) = F_n(x) \equiv B_n \sin(nx)$.

The second ordinary differential equation in (10.4.3) is therefore

$$\frac{dG}{dt}(t) + 4n^2 G(t) = 0,$$

where n is a positive integer. For each n , the general solution of this first order separable ordinary differential equation is $G(t) = G_n(t) \equiv C_n e^{-4n^2 t}$.

We have found the simple functions

$$u_n(x, t) \equiv F_n(x)G_n(t) = b_n \sin(nx) e^{-4n^2 t} \quad , \quad n > 0 \quad .$$

The constant b_n comes from the product of the constants B_n and C_n .

Initial condition: We seek a solution of the form

$$u(x, t) = \sum_{n=1}^{\infty} u_n(x, t) = \sum_{n=1}^{\infty} b_n \sin(nx) e^{-4n^2 t} \quad .$$

From $u(x, 0) = f(x)$, we get $f(x) = \sum_{n=1}^{\infty} b_n \sin(nx)$ for $0 < x < \pi$. This is the Fourier sine series of f . The coefficients of this series are given by

$$b_n = \frac{2}{\pi} \int_0^{\pi} f(x) \sin(nx) dx = \frac{2}{\pi} \int_0^{\pi/2} x \sin(nx) dx + \frac{2}{\pi} \int_{\pi/2}^{\pi} (\pi - x) \sin(nx) dx \quad .$$

The first integral can be evaluated using integration by parts. We get

$$\begin{aligned} \int_0^{\pi/2} x \sin(nx) dx &= -\frac{1}{n} x \cos(nx) \Big|_0^{\pi/2} + \frac{1}{n} \int_0^{\pi/2} \cos(nx) dx = -\frac{\pi}{2n} \cos\left(\frac{n\pi}{2}\right) + \frac{1}{n^2} \sin(nx) \Big|_0^{\pi/2} \\ &= -\frac{\pi}{2n} \cos\left(\frac{n\pi}{2}\right) + \frac{1}{n^2} \sin\left(\frac{n\pi}{2}\right) \end{aligned}$$

Similarly, the second integral can be evaluated using integration by parts

$$\begin{aligned} \int_{\pi/2}^{\pi} (\pi - x) \sin(nx) dx &= -\frac{1}{n} (\pi - x) \cos(nx) \Big|_{\pi/2}^{\pi} - \frac{1}{n} \int_{\pi/2}^{\pi} \cos(nx) dx \\ &= \frac{\pi}{2n} \cos\left(\frac{n\pi}{2}\right) - \frac{1}{n^2} \sin(nx) \Big|_{\pi/2}^{\pi} = \frac{\pi}{2n} \cos\left(\frac{n\pi}{2}\right) + \frac{1}{n^2} \sin\left(\frac{n\pi}{2}\right) \end{aligned}$$

Thus

$$b_n = \frac{4}{n^2 \pi} \sin\left(\frac{n\pi}{2}\right) = \begin{cases} 0 & \text{if } n \text{ is even} \\ (-1)^{(n+3)/2} \frac{4}{n^2 \pi} & \text{if } n \text{ is odd} \end{cases}$$

The solution of the heat equation is

$$u(x, t) = \sum_{m=1}^{\infty} (-1)^{m+1} \frac{4}{(2m-1)^2 \pi} \sin((2m-1)x) e^{-4(2m-1)^2 t} \quad . \quad (10.4.5)$$

Since

$$\left| (-1)^{m+1} \frac{4}{(2m-1)^2 \pi} \sin((2m-1)x) e^{-4(2m-1)^2 t} \right| \leq \frac{4}{\pi(2m-1)^2}$$

and the series

$$\sum_{m=1}^{\infty} \frac{4}{\pi(2m-1)^2}$$

converges, we have that the series solution converge absolutely and uniformly for $(x, t) \in [0, \pi] \times [0, \infty[$. Since each term of the series are continuous functions on $[0, \pi] \times [0, \infty[$, it follow from the uniform convergence of the series that u is continuous on $[0, \pi] \times [0, \infty[$. Moreover, u satisfies the boundary and initial conditions.

To prove that the method of separation of variables is valid in the classical sense, it remains to prove that the differential equation is satisfied by u when standard derivatives of order two in x and of order one in t are used. Since each term of the series solution of u satisfies the differential equation in the classical sense, it is enough to prove that the series (10.4.5),

$$\begin{aligned} \frac{\partial u}{\partial x} &\sim \sum_{m=1}^{\infty} \left((-1)^{m+1} \frac{4}{(2m-1)\pi} \cos((2m-1)x) e^{-4(2m-1)^2 t} \right), \\ \frac{\partial^2 u}{\partial x^2} &\sim \sum_{m=1}^{\infty} \left((-1)^m \frac{4}{\pi} \sin((2m-1)x) e^{-4(2m-1)^2 t} \right), \\ \frac{\partial u}{\partial t} &\sim \sum_{m=1}^{\infty} \left((-1)^m \frac{16}{\pi} \sin((2m-1)x) e^{-4(2m-1)^2 t} \right) \end{aligned} \quad (10.4.6)$$

converge absolutely and uniformly for $(x, t) \in [0, \pi] \times]\epsilon, \infty[$ whatever $\epsilon > 0$ arbitrary but fixed. To prove the absolute and uniform convergence of these series, it suffices to prove that the series

$$\sum_{m=1}^{\infty} (2m-1)^k e^{-4(2m-1)^2 t} \quad (10.4.7)$$

converges uniformly for $t \geq \epsilon$ whatever the value of $k \geq -2$. Since $e^{-4(2m-1)^2 t} \leq e^{-4(2m-1)^2 \epsilon}$ for $t \geq \epsilon$, we get as a consequence of the comparison test that (10.4.7) converges uniformly because

$$\sum_{m=1}^{\infty} (2m-1)^k e^{-4(2m-1)^2 \epsilon}$$

converges whatever the value of k . That this series converges can be deduced from the ratio test since

$$\lim_{m \rightarrow \infty} \frac{(2m+1)^k e^{-4(2m+1)^2 \epsilon}}{(2m-1)^k e^{-4(2m-1)^2 \epsilon}} = \lim_{m \rightarrow \infty} \left(\frac{2m+1}{2m-1} \right)^k e^{-32m\epsilon} = 0 < 1.$$

♣

Remark 10.4.2

In the previous example, the proof of the absolute and uniform convergence of the series like those in (10.4.6) is valid for any series obtained from the series solutions of u by deriving the terms of this series. This shows that the solution u is in fact infinitely differentiable with respect to x and t in $]0, \pi[\times]0, \infty[$. That was expected since $\frac{\partial}{\partial t} - \Delta_x$ is hypoelliptic.

Moreover, the previous discussion is valid whatever the choice of f and is not limited to the function f given in (10.4.2).

Even if the initial condition was not differentiable everywhere on the interval $[0, \pi]$, the solution is smooth for $t > 0$. This is typical of the parabolic equation. For the heat equation, we have seen in Remark 10.1.5 that this property is not limited to one-dimension heat equations. ♠

From now on, we will not verify that the series solution and its derivatives converge absolutely and (locally) uniformly on the domain, and that the series solution satisfy the partial differential equation. In fact, the series solution is a strong solution of the partial differential equation.

Example 10.4.3

Find the solution of the heat equation

$$\frac{\partial u}{\partial t} = c^2 \frac{\partial^2 u}{\partial x^2} \quad , \quad 0 < x < 5 \text{ and } t > 0 \quad , \quad (10.4.8)$$

with $c^2 = 4$, the boundary conditions $u(0, t) = 0$ and $\frac{\partial u}{\partial x}(5, t) = 0$ for $t > 0$, and the following initial condition.

$$\text{a) } u(x, 0) = 8 \sin\left(\frac{13\pi x}{10}\right) \quad \text{b) } u(x, 0) = x(5 - x)^2$$

Separation of Variables: If we substitute $u(x, t) = F(x)G(t)$ in (10.4.8), we get

$$F(x) \frac{dG}{dt}(t) = 4 \frac{d^2 F}{dx^2}(x) G(t) \quad , \quad t > 0 \text{ and } 0 < x < 5 \quad .$$

Thus, after dividing both sides by $4F(x)G(t)$, we get

$$\frac{1}{4G(t)} \frac{dG}{dt}(t) = \frac{1}{F(x)} \frac{d^2 F}{dx^2}(x) \quad , \quad t > 0 \text{ and } 0 < x < 5 \quad .$$

Since the right hand side is independent of t and the left hand side is independent of x , we get

$$\frac{1}{4G(t)} \frac{dG}{dt}(t) = \frac{1}{F(x)} \frac{d^2 F}{dx^2}(x) = k \quad , \quad t > 0 \text{ and } 0 < x < 5 \quad ,$$

for some constant k . We end up with two ordinary differential equations.

$$\frac{d^2 F}{dx^2}(x) - kF(x) = 0 \quad \text{and} \quad \frac{dG}{dt}(t) - 4kG(t) = 0 \quad . \quad (10.4.9)$$

Two boundary conditions are associated to the first ordinary differential equation in (10.4.9). From $u(0, t) = 0$, we get $F(0)G(t) = 0$. Since we assume that G is not null, we get $F(0) = 0$. Similarly, from $\frac{\partial u}{\partial x}(5, t) = 0$, we get $F'(5)G(t) = 0$. Again, since we assume that G is not null, we get $F'(5) = 0$. The boundary conditions for the first ordinary differential equation in (10.4.9) are $F(0) = F'(5) = 0$.

Simple Functions: We first consider the boundary value problem

$$\frac{d^2 F}{dx^2}(x) - kF(x) = 0 \quad , \quad 0 \leq x \leq 5 \quad , \quad (10.4.10)$$

with $F(0) = F'(5) = 0$. The form of the general solution of this ordinary differential equation is determined by the roots of the characteristic equation $\lambda^2 - k = 0$.

If $k > 0$, the roots of the characteristic equation are $\pm\sqrt{k}$. Since the roots are real, the solution of the ordinary differential equation is of the form $F(x) = Ae^{\sqrt{k}x} + Be^{-\sqrt{k}x}$. However, $F(0) = 0$ implies that $A + B = 0$ and $F'(5) = 0$ implies $A\sqrt{k}e^{5\sqrt{k}} - B\sqrt{k}e^{-5\sqrt{k}} = 0$. The only solution of the system

$$\begin{pmatrix} 1 & 1 \\ \sqrt{k}e^{5\sqrt{k}} & -\sqrt{k}e^{-5\sqrt{k}} \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

is $A = B = 0$ because

$$\det \begin{pmatrix} 1 & 1 \\ \sqrt{k}e^{5\sqrt{k}} & -\sqrt{k}e^{-5\sqrt{k}} \end{pmatrix} = -\sqrt{k}(e^{-5\sqrt{k}} + e^{5\sqrt{k}}) < 0$$

for $k > 0$. Therefore, the null solution is the only solution of the boundary value problem (10.4.10) for $k > 0$.

If $k = 0$, the solution of the ordinary differential equation is $F(x) = Ax + B$. However, $F(0) = 0$ yields $B = 0$ and $F'(5) = 0$ yields $A = 0$. Thus $A = B = 0$ and again the null solution is the only solution of the boundary value problem (10.4.10) for $k = 0$.

If $k < 0$, the roots of the characteristic equation are $\pm i\sqrt{-k}$. Since the roots are complex, the solution of the ordinary differential equation is of the form $F(x) = A\cos(\sqrt{-k}x) + B\sin(\sqrt{-k}x)$. The condition $F(0) = 0$ implies that $A = 0$. Hence, $F'(5) = 0$ implies that $B\sqrt{-k}\cos(5\sqrt{-k}) = 0$. If we take $B = 0$, we get the null solution. We must therefore have $\cos(5\sqrt{-k}) = 0$. This implies that $k = k_n \equiv -((2n-1)\pi/10)^2$ for $n > 0$. The boundary value problem (10.4.10) has non-null solutions only for $k = k_n = -((2n-1)\pi/10)^2$ with n a positive integer, and the solutions associated to k_n are of the form $F(x) = F_n(x) \equiv B_n \sin((2n-1)\pi x/10)$.

The second ordinary differential equation in (10.4.9) is then

$$\frac{dG}{dt}(t) + 4\left(\frac{(2n-1)\pi}{10}\right)^2 G(t) = 0 \quad ,$$

where n is a positive integer. For each n , the general solution of this first order separable ordinary differential equation is $G(t) = G_n(t) \equiv C_n e^{-t((2n-1)\pi/5)^2}$.

We have found the simple functions

$$u_n(x, t) \equiv F_n(x)G_n(t) = b_n \sin\left(\frac{(2n-1)\pi x}{10}\right) e^{-t((2n-1)\pi/5)^2} \quad , \quad n > 0 \quad .$$

The constant b_n comes from the product of the constants B_n and C_n .

Initial condition: We seek a solution of the form

$$u(x, t) = \sum_{n=1}^{\infty} u_n(x, t) = \sum_{n=1}^{\infty} b_n \sin\left(\frac{(2n-1)\pi x}{10}\right) e^{-t((2n-1)\pi/5)^2}.$$

a) From $u(x, 0) = 8 \sin\left(\frac{13\pi x}{10}\right)$, we get

$$8 \sin\left(\frac{13\pi x}{10}\right) = \sum_{n=1}^{\infty} b_n \sin\left(\frac{(2n-1)\pi x}{10}\right).$$

Thus,

$$b_n = \begin{cases} 8 & \text{if } n = 7 \\ 0 & \text{if } n \neq 7 \end{cases}$$

The solution is

$$u(x, t) = 8 \sin\left(\frac{13\pi x}{10}\right) e^{-t(13\pi/5)^2}.$$

b) From $u(x, 0) = x(5-x)^2$, we get

$$x(5-x)^2 = \sum_{n=1}^{\infty} b_n \sin\left(\frac{(2n-1)\pi x}{10}\right)$$

for $0 \leq x \leq 5$. Expanding $f(x) = x(5-x)^2$ on the interval $[0, 5]$ to a odd function of period 20 as suggested in Section 4.3.2, we find that

$$b_n = \frac{2}{5} \int_0^5 x(5-x)^2 \sin\left(\frac{(2n-1)\pi x}{10}\right) dx, \quad n > 0.$$

A long computation using integration by parts yields

$$b_n = 4000 \left(\frac{4n\pi - 2\pi + (-1)^n 6}{\pi^4 (16n^4 - 32n^3 + 24n^2 - 8n + 1)} \right), \quad n > 0.$$

The solution of the heat equation is

$$u(x, t) = 4000 \sum_{n=1}^{\infty} \frac{4n\pi - 2\pi + (-1)^n 6}{\pi^4 (16n^4 - 32n^3 + 24n^2 - 8n + 1)} \sin\left(\frac{(2n-1)\pi x}{10}\right) e^{-t((2n-1)\pi/5)^2}.$$

♣

Remark 10.4.4

By separation of variables, we find that the solutions of the heat equation for a finite interval with null boundary conditions are linear combinations of the following simple functions.

$$u(x, t) = e^{-c^2 k^2 t} (A \cos(kx) + B \sin(kx)).$$

The solutions of the heat equation for a rectangular domain with null boundary conditions are linear combinations of the following simple functions.

$$u(x, y, t) = e^{-c^2 k^2 t} (A \cos(\mu x) + B \sin(\mu x))(C \cos(\eta y) + D \sin(\eta y))$$

with $k^2 = \mu^2 + \eta^2$. ♠

Example 10.4.5

Solve the heat equation

$$\frac{\partial u}{\partial t} = c^2 \Delta u = c^2 \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right) \quad , \quad 0 < x, y < 5 \text{ and } t > 0$$

with $c = 3$, the boundary conditions

$$\begin{aligned} u(0, y, t) = u(5, y, t) = 0 \quad , \quad 0 < y < 5 \quad \text{and} \quad t > 0 \quad , \\ \frac{\partial u}{\partial y}(x, 0, t) = \frac{\partial u}{\partial y}(x, 5, t) = 0 \quad , \quad 0 < x < 5 \quad \text{and} \quad t > 0 \quad , \end{aligned}$$

and the initial conditions

$$u(x, y, 0) = 11 \sin\left(\frac{3\pi x}{5}\right) + 17 \sin(2\pi x) \cos\left(\frac{7\pi y}{5}\right) \quad , \quad 0 \leq x, y \leq 5 \quad .$$

The simple functions are of the form

$$u(x, y, t) = e^{-c^2 k^2 t} (A \cos(\mu x) + B \sin(\mu x)) (C \cos(\eta y) + D \sin(\eta y)) \quad ,$$

where $k^2 = \mu^2 + \eta^2$. From $u(0, y, t) = 0$, we get $Ae^{-c^2 k^2 t} (C \cos(\eta y) + D \sin(\eta y)) = 0$. Since the expression between parentheses cannot be zero for all $0 \leq y \leq 5$ (i.e. we cannot have $C = D = 0$), we must have $A = 0$. From $u(5, y, t) = 0$, we then get $B \sin(5\mu) e^{-c^2 k^2 t} (C \cos(\eta y) + D \sin(\eta y)) = 0$. Again, since the expression between parentheses cannot be zero for all $0 \leq y \leq 5$ and B cannot be 0, we must have $\sin(5\mu) = 0$. This yields $\mu = m\pi/5$ for $m > 0$.

Since $\frac{\partial u}{\partial y} = B \sin\left(\frac{m\pi x}{5}\right) e^{-c^2 k^2 t} (-C\eta \sin(\eta y) + D\eta \cos(\eta y))$, we get from $\frac{\partial u}{\partial y}(x, 0, t) = 0$ that $BD\eta \sin\left(\frac{m\pi x}{5}\right) e^{-c^2 k^2 t} = 0$ for all $0 \leq x \leq 5$ and $t > 0$, where $B \neq 0$. Thus $D = 0$. From $\frac{\partial u}{\partial y}(x, 5, t) = 0$, we then get $-BC\eta \sin(5\eta) \sin\left(\frac{m\pi x}{5}\right) e^{-c^2 k^2 t} = 0$ for all $0 \leq x \leq 5$ and $t > 0$, where $B \neq 0$. Since we cannot have $C = 0$, we must have $\sin(5\eta) = 0$. Thus $\eta = n\pi/5$ for $n \geq 0$.

We have found the simple functions

$$u_{n,m}(x, y, t) \equiv A_{n,m} e^{-9k_{n,m}^2 t} \sin\left(\frac{m\pi x}{5}\right) \cos\left(\frac{n\pi y}{5}\right) \quad ,$$

where m is a positive integer, n is a non-negative integer (so $n = 0$ is possible), and

$$k_{n,m}^2 = \left(\frac{m\pi}{5}\right)^2 + \left(\frac{n\pi}{5}\right)^2 = \frac{\pi^2}{5^2} (m^2 + n^2) \quad .$$

We seek a solution of the form

$$u(x, y, t) = \sum_{m=1}^{\infty} \sum_{n=0}^{\infty} u_{n,m}(x, y, t) = \sum_{m=1}^{\infty} \sum_{n=0}^{\infty} A_{n,m} e^{-9k_{n,m}^2 t} \sin\left(\frac{m\pi x}{5}\right) \cos\left(\frac{n\pi y}{5}\right) \quad .$$

From the initial condition

$$u(x, y, 0) = 11 \sin\left(\frac{3\pi x}{5}\right) + 17 \sin\left(\frac{10\pi x}{5}\right) \cos\left(\frac{7\pi y}{5}\right),$$

we get

$$\sum_{m=1}^{\infty} \sum_{n=0}^{\infty} A_{n,m} \sin\left(\frac{m\pi x}{5}\right) \cos\left(\frac{n\pi y}{5}\right) = 11 \sin\left(\frac{3\pi x}{5}\right) + 17 \sin\left(\frac{10\pi x}{5}\right) \cos\left(\frac{7\pi y}{5}\right).$$

Thus,

$$A_{n,m} = \begin{cases} 11 & \text{if } m = 3, n = 0 \\ 17 & \text{if } m = 10, n = 7 \\ 0 & \text{otherwise} \end{cases}$$

Since $k_{n,m} = 9\pi^2/25$ for $m = 3$ and $n = 0$, and $k_{n,m} = 149\pi^2/25$ for $m = 10$ and $n = 7$, we get the solution

$$u(x, y, t) = 11e^{-81\pi^2 t/25} \sin\left(\frac{3\pi x}{5}\right) + 17e^{-1341\pi^2 t/25} \sin\left(\frac{10\pi x}{5}\right) \cos\left(\frac{7\pi y}{5}\right).$$

♣

Example 10.4.6

Find the solution of the heat equation

$$\frac{\partial u}{\partial t} - c^2 \frac{\partial^2 u}{\partial x^2} = 0, \quad 0 < x < L \text{ and } t > 0, \quad (10.4.11)$$

with the boundary conditions $u(0, t) = t^2$ and $u(L, t) = 1$ for $t > 0$, and the initial condition $u(x, 0) = x/L$ for $0 \leq x \leq L$.

We cannot readily use our formulae for the simple functions because the boundary conditions are not null. We have first to transform the problem into one where the formulae can be used.

Let $\phi(x, t) = (1 - t^2)x/L + t^2 = (x + (L - x)t^2)/L$. This is the equation of the line in the x, u -plane that contains the points $(0, t^2)$ and $(L, 1)$. The substitution $u(x, t) = w(x, t) + \phi(x, t)$ in (10.4.11) yields

$$\frac{\partial w}{\partial t} - c^2 \frac{\partial^2 w}{\partial x^2} = \frac{2(x - L)t}{L}, \quad 0 < x < L \text{ and } t > 0, \quad (10.4.12)$$

with the boundary conditions $w(0, t) = 0$ and $w(L, t) = 0$ for $t > 0$, and the initial condition $w(x, 0) = 0$ for $0 \leq x \leq L$.

The simple functions of the heat equation (i.e. when the right hand side of (10.4.12) is 0) are of the form $B_n e^{-c^2 n^2 \pi^2 t/L^2} \sin\left(\frac{n\pi x}{L}\right)$ due to the boundary conditions. So, we seek a solution of the form

$$w(x, t) = \sum_{n=1}^{\infty} T_n(t) \sin\left(\frac{n\pi x}{L}\right). \quad (10.4.13)$$

The boundary conditions are satisfied by each term of the series.

We need the Fourier sine series of $h(x, t) = \frac{2(x-L)t}{L}$ with respect to x . Using integration by parts, we find that the coefficients of the series are

$$\begin{aligned} b_n &= \frac{2}{L} \int_0^L \frac{2(x-L)t}{L} \sin\left(\frac{n\pi x}{L}\right) dx = \left(\frac{4t(L-x)}{n\pi L} \cos\left(\frac{n\pi x}{L}\right) \right) \Big|_0^L + \frac{4t}{n\pi L} \int_0^L \cos\left(\frac{n\pi x}{L}\right) dx \\ &= -\frac{4t}{n\pi} + \frac{4t}{n^2\pi^2} \sin\left(\frac{n\pi x}{L}\right) \Big|_0^L = -\frac{4t}{n\pi} . \end{aligned}$$

Thus,

$$h(x, t) = - \sum_{n=1}^{\infty} \frac{4t}{n\pi} \sin\left(\frac{n\pi x}{L}\right) . \quad (10.4.14)$$

If we substitute (10.4.13) and (10.4.14) into (10.4.12), we get

$$\sum_{n=1}^{\infty} \left(T'_n(t) + \left(\frac{cn\pi}{L} \right)^2 T_n(t) \right) \sin\left(\frac{n\pi x}{L}\right) = - \sum_{n=1}^{\infty} \frac{4t}{n\pi} \sin\left(\frac{n\pi x}{L}\right) .$$

Each T_n must satisfy the ordinary differential equation.

$$T'_n(t) + \left(\frac{cn\pi}{L} \right)^2 T_n(t) = -\frac{4t}{n\pi} , \quad n > 0 .$$

Those are linear first order ordinary differential equations; namely, ordinary differential equations of the form $\frac{dT}{dt} + P(t)T = Q(t)$. The general solution of this ordinary differential equation is $T(t) = e^{-\int P(t) dt} \left(\int Q(t) e^{\int P(t) dx} dt + C \right)$. Hence,

$$T_n(t) = e^{-c^2 n^2 \pi^2 t / L^2} \left(- \int \frac{4t}{n\pi} e^{c^2 n^2 \pi^2 t / L^2} dt + C_n \right)$$

Using integration by parts to evaluate the integral in the expression of T_n above, we obtain

$$\begin{aligned} \int \frac{4t}{n\pi} e^{c^2 n^2 \pi^2 t / L^2} dt &= \left(\frac{4t}{n\pi} \right) \left(\frac{L^2}{c^2 n^2 \pi^2} \right) e^{c^2 n^2 \pi^2 t / L^2} - \left(\frac{4}{n\pi} \right) \left(\frac{L^2}{c^2 n^2 \pi^2} \right) \int e^{c^2 n^2 \pi^2 t / L^2} dt \\ &= \left(\frac{4L^2 t}{c^2 n^3 \pi^3} \right) e^{c^2 n^2 \pi^2 t / L^2} - \left(\frac{4L^4}{c^4 n^5 \pi^5} \right) e^{c^2 n^2 \pi^2 t / L^2} , \end{aligned}$$

where we may ignore the constant of integration since we already have C_n . Hence,

$$T_n(t) = - \left(\frac{4L^2 t}{c^2 n^3 \pi^3} \right) + \left(\frac{4L^4}{c^4 n^5 \pi^5} \right) + C_n e^{-c^2 n^2 \pi^2 t / L^2} , \quad n > 0 .$$

We have found that

$$w(x, t) = \sum_{n=1}^{\infty} \left(- \left(\frac{4L^2 t}{c^2 n^3 \pi^3} \right) + \left(\frac{4L^4}{c^4 n^5 \pi^5} \right) + C_n e^{-c^2 n^2 \pi^2 t / L^2} \right) \sin\left(\frac{n\pi x}{L}\right) .$$

To determine the C_n , we use the initial condition $w(x, 0) = 0$ to get

$$\left(-\left(\frac{4L^2 t}{c^2 n^3 \pi^3} \right) + \left(\frac{4L^4}{c^4 n^5 \pi^5} \right) + C_n e^{-c^2 n^2 \pi^2 t / L^2} \right) \Big|_{t=0} = \left(\frac{4L^4}{c^4 n^5 \pi^5} \right) + C_n = 0$$

for all n . Thus $C_n = -\left(\frac{4L^4}{c^4 n^5 \pi^5} \right)$ for $n > 0$ and

$$w(x, t) = \sum_{n=1}^{\infty} \left(\frac{4L^4}{c^4 n^5 \pi^5} \right) \left(-\frac{c^2 n^2 \pi^2}{L^2} t + 1 - e^{-c^2 n^2 \pi^2 t / L^2} \right) \sin\left(\frac{n\pi x}{L} \right).$$

The solution of the problem is

$$u(x, t) = \frac{(1 - t^2)x}{L} + t^2 + \sum_{n=1}^{\infty} \left(\frac{4L^4}{c^4 n^5 \pi^5} \right) \left(-\frac{c^2 n^2 \pi^2}{L^2} t + 1 - e^{-c^2 n^2 \pi^2 t / L^2} \right) \sin\left(\frac{n\pi x}{L} \right).$$

♣

10.4.2 Cylindrical and Spherical Coordinates

Remark 10.4.7

The solutions of the heat equation in a cylindrical domain with null boundary conditions (this includes the planar circular domain if the solution is independent of the z -coordinate) are linear combinations of the following simple functions. If the problem is independent of θ and z ,

$$u(r, t) = A e^{-c^2 k^2 t} J_0(kr).$$

If the problem is independent of z only

$$u(r, \theta, t) = e^{-c^2 k^2 t} (A \cos(n\theta) + B \sin(n\theta)) J_n(kr) \quad , \quad n \geq 0.$$

If the problem is independent of θ only

$$u(r, z, t) = e^{-c^2 k^2 t} (A \cos(\eta z) + B \sin(\eta z)) J_0(\mu r)$$

with $k^2 = \mu^2 + \eta^2$.

♠

Remark 10.4.8

The solutions of the heat equation on a spherical domain with null boundary conditions are linear combinations of the following simple functions. If the problem is independent of θ and ϕ ,

$$u(r, t) = e^{-c^2 k^2 t} \left(\frac{A}{kr} \cos(kr) + \frac{B}{kr} \sin(kr) \right).$$

♠

10.5 Addendum

The heat equation is the archetypal example of a parabolic equation. It is also used as an example for an evolution or dissipating partial differential equation. We are talking of irreversible time-dependent partial differential equations. Namely, the partial differential equation is not invariant under the substitution $t \rightarrow -t$. We cannot go back in time.

There are many techniques used to study parabolic partial differential equations. Some of the classical techniques are energy methods, and Laplace transforms. There is also a more modern approach using semi-group theory. One of the fundamental results of this approach is the Hille-Yosida theorem used to determine if a semi-group is strongly continuous which is an important property to solve parabolic partial differential equations. Unfortunately, all these techniques will not be presented in this book. A good reference for the semi-group approach to parabolic partial differential equations is [12].

For the study of nonlinear (parabolic) partial differential equations, we recommend [35] and, in particular, [12].

10.6 Exercises

Suggested exercises:

- In [16]: numbers 4 to 6, 8, 11 and 12 in Section 7.1(a); number 1 in Section 7.1(c).
- In [22]: numbers 6 and 7 in Sections 5.1; numbers 2, 3, 5 to 8, and 11 in Section 5.2.
- In [24]: numbers 5.1, 5.5 to 5.10, 5.12, 5.13 and 5.20 in Section 5.7; numbers 6.13 to 6.17, 6.21 to 6.23, and 6.25 in Section 6.7.
- In [38]: all the numbers in Section 2.3; numbers 1 to 4, 8, 11, 16 to 19 in Section 2.4; numbers 1 to 3 in Section 3.1; all the numbers in Section 3.3; numbers 1 and 4 in Section 4.2; numbers 1 and 8 in Section 5.6.

Chapter 11

Sobolev Spaces

This chapter is a brief introduction to Sobolev spaces. We give some of the most important results about these spaces. In particular, we give those which will be useful for our study of elliptic partial differential equations in the next chapter. As we said, we present only a few results about Sobolev spaces and often not in their most general form. We refer the reader to [2] for a thorough coverage of the subject.

11.1 The $W^{k,p}$ spaces

Definition 11.1.1

Let Ω be an open subset of $\mathbb{R}^n$. For $k \in \mathbb{N}$ and p a positive integer or ∞ , the **Sobolev space** $W^{k,p}(\Omega)$ is the space defined by

$$W^{k,p}(\Omega) = \{f \in L^p(\Omega) : D^\alpha f \in L^p(\Omega), \alpha \in \mathbb{N}^n \text{ and } |\alpha| \leq k\},$$

where the derivatives D^α are derivatives in the sense of distributions.

We can define a norm on $W^{k,p}(\Omega)$ by

$$\|f\|_{k,p} = \begin{cases} \left(\sum_{|\alpha| \leq k} \|D^\alpha f\|_p^p \right)^{1/p} & \text{if } 0 < p < \infty \\ \max_{|\alpha| \leq k} \|D^\alpha f\|_\infty & \text{if } p = \infty \end{cases} \quad (11.1.1)$$

for $f \in W^{k,p}(\Omega)$, where $\|\cdot\|_p$ is the L^p -norm on $L^p(\Omega)$. We will sometime use the notation $\|f\|_{p,\Omega}$ and $\|f\|_{k,p,\Omega}$ when you need to clearly identify the domain to be considered for the norm.

Note that $\|f\|_\infty = \operatorname{ess\,sup}_{\mathbf{x} \in \Omega} |f(\mathbf{x})|$. However, to simplify the notation, we will only write $\|f\|_\infty = \sup_{\mathbf{x} \in \Omega} |f(\mathbf{x})|$ but the reader must understand that the essential supremum is used when f is only defined almost everywhere.

Theorem 11.1.2

Let Ω be an open subset of $\mathbb{R}^n$. For $k \in \mathbb{N}$ and p a positive integer or ∞ , the space $W^{k,p}(\Omega)$ equipped with the norm (11.1.1) is a Banach space. If $p \neq \infty$, $W^{k,p}(\Omega)$ is also separable.

Proof.

Let $\{\alpha_1, \alpha_2, \dots, \alpha_N\}$, where $N = \sum_{|\alpha| \leq k} 1$, be an ordering of the multi-indices $\alpha \in \mathbb{N}^n$ such that $|\alpha| \leq k$. We assume that $\alpha_1 = (0, 0, \dots, 0)$. Consider the mapping

$$\begin{aligned} T : W^{k,p}(\Omega) &\rightarrow (L^p(\Omega))^N \\ f &\mapsto (D^{\alpha_1} f, D^{\alpha_2} f, \dots, D^{\alpha_N} f) \end{aligned}$$

This is an isometry if the norm on $(L^p(\Omega))^N$ is the norm

$$\|f\|_p = \begin{cases} \left(\sum_{j=1}^N \|f_{\alpha_j}\|_p^p \right)^{1/p} & \text{if } 0 < p < \infty \\ \max_{1 \leq j \leq N} \|f_{\alpha_j}\|_\infty & \text{if } p = \infty \end{cases}$$

for $f \in (L^p(\Omega))^N$.

i) We consider the case $1 < p < \infty$ and leave the cases $p = 1$ and $p = \infty$ to the reader.

We prove that $W^{k,p}(\Omega)$ is complete by proving that T is a closed mapping. Suppose that $\{f_j\}_{j=1}^\infty \subset W^{k,p}(\Omega)$ is such that $\{T(f_j)\}_{j=1}^\infty$ converges to g in $(L^p(\Omega))^N$. Then $\{D^{\alpha_i} f_j\}_{j=1}^\infty$ converge to g_{α_i} in $L^p(\Omega)$ for $1 \leq i \leq N$. In particular, for α_1 , we have that $\{f_j\}_{j=1}^\infty$ converges to $g \equiv g_{\alpha_1}$ in $L^p(\Omega)$.

Since L^p convergence implies convergence in the sense of distributions, we have that $g_{\alpha_i} = D^{\alpha_i} g$ almost everywhere for all $0 \leq i \leq N$. To prove this statement, we use the definition of the derivative in the sense of distributions to write

$$\int_{\Omega} D^{\alpha_i} f_j(\mathbf{x}) \phi(\mathbf{x}) \, d\mathbf{x} = (-1)^{|\alpha_i|} \int_{\Omega} f_j(\mathbf{x}) D^{\alpha_i} \phi(\mathbf{x}) \, d\mathbf{x}$$

for all $\phi \in \mathcal{D}(\Omega)$. As $j \rightarrow \infty$, the integral on the left converges to $\int_{\Omega} g_{\alpha_i}(\mathbf{x}) \phi(\mathbf{x}) \, d\mathbf{x}$ for all $\phi \in \mathcal{D}(\Omega)$ because

$$\begin{aligned} \left| \int_{\Omega} D^{\alpha_i} f_j(\mathbf{x}) \phi(\mathbf{x}) \, d\mathbf{x} - \int_{\Omega} g_{\alpha_i}(\mathbf{x}) \phi(\mathbf{x}) \, d\mathbf{x} \right| &\leq \int_{\Omega} |D^{\alpha_i} f_j(\mathbf{x}) - g_{\alpha_i}(\mathbf{x})| |\phi(\mathbf{x})| \, d\mathbf{x} \\ &\leq \left(\int_{\Omega} |D^{\alpha_i} f_j(\mathbf{x}) - g_{\alpha_i}(\mathbf{x})|^p \, d\mathbf{x} \right)^{1/p} \left(\int_{\Omega} |\phi(\mathbf{x})|^q \, d\mathbf{x} \right)^{1/q} \rightarrow 0 \quad \text{as } j \rightarrow \infty \end{aligned}$$

by assumption, where $1/p + 1/q = 1$. Similarly, the integral on the right converges to

$$(-1)^{|\alpha_i|} \int_{\Omega} g(\mathbf{x}) D^{\alpha_i} \phi(\mathbf{x}) \, d\mathbf{x} = \int_{\Omega} D^{\alpha_i} g(\mathbf{x}) \phi(\mathbf{x}) \, d\mathbf{x}$$

for all $\phi \in \mathcal{D}(\Omega)$. Hence,

$$\int_{\Omega} g_{\alpha_i}(\mathbf{x}) \phi(\mathbf{x}) d\mathbf{x} = \int_{\Omega} D^{\alpha_i} g(\mathbf{x}) \phi(\mathbf{x}) d\mathbf{x}$$

for all $\phi \in \mathcal{D}(\Omega)$ and all i . Thus, $g_{\alpha_i} = D^{\alpha_i} g$ almost everywhere for all i .

ii) To prove that $W^{k,p}(\Omega)$ is separable for $0 \leq p < \infty$, it is enough to note that $T : W^{k,p}(\Omega) \rightarrow T(W^{k,p}(\Omega))$ is an isometry and that $T(W^{k,p}(\Omega))$ is a subset of the separable space $(L^p(\Omega))^N$. ■

The most important Sobolev spaces (for the study of partial differential equations) are those with $p = 2$. There is a nice equivalent way to define the spaces $W^{k,2}(\mathbb{R}^n)$ for $0 \leq k \leq \infty$.

Theorem 11.1.3

For $k \in \mathbb{N}$, $f \in W^{k,2}(\mathbb{R}^n)$ if and only if $\hat{f} \in L^2(\mathbb{R}^n, \rho)$, where the measure ρ is defined by $d\rho(\mathbf{x}) = (1 + \|\mathbf{x}\|_2^2)^k d\mathbf{x}$. Moreover, the norm $\|\cdot\|_{k,2}$ defined in (11.1.1) is equivalent to the norm

$$\|f\|_{k,\rho} = \left(\int_{\mathbb{R}^n} |\hat{f}(\mathbf{y})|^2 d\rho(\mathbf{y}) \right)^{1/2}, \quad f \in W^{k,2}(\mathbb{R}^n). \quad (11.1.2)$$

Proof.

i) We first find two positive constants, C_1 and C_2 , such that

$$C_1 (1 + \|\mathbf{y}\|_2^2)^k \leq \sum_{|\alpha| \leq k} |\mathbf{y}^{2\alpha}| \leq C_2 (1 + \|\mathbf{y}\|_2^2)^k$$

for $\mathbf{y} \in \mathbb{R}^n$.

i.a) Since $|\mathbf{y}^\alpha| \leq 1$ for $\|\mathbf{y}\|_2 \leq 1$, and $|\mathbf{y}^\alpha| \leq \|\mathbf{y}\|_2^{|\alpha|} \leq \|\mathbf{y}\|_2^k$ for $\|\mathbf{y}\|_2 \geq 1$ and $|\alpha| \leq k$, we get

$$\sum_{|\alpha| \leq k} |\mathbf{y}^{2\alpha}| = \sum_{|\alpha| \leq k} |\mathbf{y}^\alpha|^2 \leq C_2 \max\{1, \|\mathbf{y}\|_2^{2k}\} \leq C_2 (1 + \|\mathbf{y}\|_2^2)^k$$

for $\mathbf{y} \in \mathbb{R}^n$, where $C_2 = \sum_{|\alpha| \leq k} 1$.

i.b) Let

$$C'_1 = \min_{\|\mathbf{y}\|=1} \sum_{|\alpha|=k} |\mathbf{y}^{2\alpha}|. \quad (11.1.3)$$

Since $\sum_{|\alpha|=k} |\mathbf{y}^{2\alpha}|$ is continuous on the compact set $\{\mathbf{y} : \|\mathbf{y}\|_2 = 1\}$, the minimum is reached at

a point $\mathbf{w} \neq \mathbf{0}$ in $\mathbb{R}^n$. If $w_j \neq 0$, then $|w_j^{2k}| > 0$ and thus $C'_1 = \sum_{|\alpha|=k} |\mathbf{w}^{2\alpha}| > 0$. Given $\mathbf{y} \neq \mathbf{0}$, let

$\mathbf{z} = \|\mathbf{y}\|_2^{-1} \mathbf{y}$. Since $\|\mathbf{z}\|_2 = 1$, we get from $\sum_{|\alpha|=k} |\mathbf{z}^{2\alpha}| \geq C'_1 > 0$ that $\sum_{|\alpha|=k} |\mathbf{y}^{2\alpha}| \geq C'_1 \|\mathbf{y}\|_2^{2k}$ because

$\sum_{|\alpha|=k} |\mathbf{z}^{2\alpha}|$ is an homogeneous function of degree $2k$. We have that

$$(1 + \|\mathbf{y}\|_2^2)^k \leq 2^k \max\{1, \|\mathbf{y}\|_2^{2k}\} \leq 2^k (1 + \|\mathbf{y}\|_2^{2k}) \leq 2^k \left(1 + \frac{1}{C'_1} \sum_{|\alpha|=k} |\mathbf{y}^{2\alpha}| \right)$$

$$\leq 2^k C_1'' \left(1 + \sum_{|\alpha|=k} |\mathbf{y}^{2\alpha}| \right) \leq 2^k C_1'' \sum_{|\alpha| \leq k} |\mathbf{y}^{2\alpha}|$$

for $\mathbf{y} \in \mathbb{R}^n$, where $C_1'' = \max\{1, 1/C_1'\}$. Finally, this is

$$C_1 (1 + \|\mathbf{y}\|^2)^k \leq \sum_{|\alpha| \leq k} |\mathbf{y}^{2\alpha}|$$

for $\mathbf{y} \in \mathbb{R}^n$, where $C_1 = 1/(2^k C_1'')$.

ii) We have that

$$\begin{aligned} \|f\|_{k,\rho} < \infty &\Leftrightarrow \int_{\mathbb{R}^n} |\hat{f}(\mathbf{y})|^2 (1 + \|\mathbf{y}\|_2^2)^k d\mathbf{y} < \infty \Leftrightarrow \int_{\mathbb{R}^n} |\mathbf{y}^\alpha \hat{f}(\mathbf{y})|^2 < \infty \text{ for all } |\alpha| \leq k \\ &\Leftrightarrow \mathbf{y}^\alpha \hat{f}(\mathbf{y}) \in L^2(\mathbb{R}^n) \text{ for all } |\alpha| \leq k \Leftrightarrow D^\alpha f \in L^2(\mathbb{R}^n) \text{ for all } |\alpha| \leq k \\ &\Leftrightarrow \|f\|_{k,2} < \infty, \end{aligned}$$

where the second to last equivalence is due to Plancherel theorem, Theorem 6.6.6. To be more precise, we saw in Proposition 6.6.12 that $(D^\alpha u)^\wedge = (i\mathbf{y})^\alpha \hat{u}$ in the sense of distributions for all $u \in \mathcal{S}'(\mathbb{R}^n)$ and multi-index α . Therefore, $(D^\alpha u)^\wedge \in L^2(\mathbb{R}^n)$ if and only if $(i\mathbf{y})^\alpha \hat{u} \in L^2(\mathbb{R}^n)$. It follows from Plancherel theorem that there exists $g \in L^2(\mathbb{R}^n)$ such that $(D^\alpha f)^\wedge = \hat{g}$ in the sense of distributions if and only if $(i\mathbf{y})^\alpha \hat{f} \in L^2(\mathbb{R}^n)$. But $(D^\alpha f)^\wedge = \hat{g}$ in the sense of distributions means that $(D^\alpha f)(\hat{\phi}) = g(\hat{\phi})$ for all $\phi \in \mathcal{S}(\mathbb{R}^n)$. From Proposition 6.6.11, this means that $(D^\alpha f)(\phi) = g(\phi)$ for all $\phi \in \mathcal{S}(\mathbb{R}^n)$. Therefore, $D^\alpha f = g \in L^2(\mathbb{R}^n)$ in the sense of distributions. Thus, the two norms describe the same space $W^{k,2}(\mathbb{R}^n)$.

Moreover, since $(D^\alpha f)^\wedge(\mathbf{y}) = (i\mathbf{y})^\alpha \hat{f}(\mathbf{y})$ for $\mathbf{y} \in \mathbb{R}^n$ and $f \in W^{k,2}(\mathbb{R}^n)$, we get from Plancherel theorem that

$$\sum_{|\alpha| \leq k} \|D^\alpha f\|_2^2 = \sum_{|\alpha| \leq k} \|(D^\alpha f)^\wedge\|_2^2 = \sum_{|\alpha| \leq k} \|\mathbf{y}^\alpha \hat{f}(\mathbf{y})\|_2^2 = \int_{\mathbb{R}^n} \left(\sum_{|\alpha| \leq k} |\mathbf{y}^{2\alpha}| \right) |\hat{f}(\mathbf{y})|^2 d\mathbf{y}.$$

Thus, it follows from (i) that $C_1 \|f\|_{k,\rho}^2 \leq \|f\|_{k,2}^2 \leq C_2 \|f\|_{k,\rho}^2$. ■

Hence, we can expand the definition of $W^{k,2}(\mathbb{R}^n)$ from $k \in \mathbb{N}$ to $k \in \mathbb{R}$ as follows.

Definition 11.1.4

For $s \in \mathbb{R}$, the **Sobolev space** $W^{s,2}(\mathbb{R}^n)$ is the space defined by

$$W^{s,2}(\mathbb{R}^n) = \{f \in L^2(\mathbb{R}^n) : \hat{f} \in L^2(\mathbb{R}^n, \rho)\},$$

where the measure ρ is defined by $d\rho(\mathbf{x}) = (1 + \|\mathbf{x}\|^2)^s d\mathbf{x}$.

It follows from Theorem 11.1.3 that the space $W^{k,2}(\mathbb{R}^n)$ for $k \in \mathbb{N}$ is an Hilbert space equipped with the scalar product

$$\langle f, g \rangle_{k,\rho} = \int_{\mathbb{R}^n} \hat{f}(\mathbf{y}) \overline{\hat{g}(\mathbf{y})} d\rho(\mathbf{y}) \quad (11.1.4)$$

for $f, g \in W^{k,2}(\mathbb{R}^n)$.

A direct consequence of Definition 11.1.4 is the following proposition.

Proposition 11.1.5

For $s \in \mathbb{R}$, $f \in W^{s,2}(\mathbb{R}^n)$ if and only if $D^\alpha f \in W^{s-|\alpha|,2}(\mathbb{R}^n)$ for all multi-indices α such that $|\alpha| \leq s$. Moreover, D^α is a bounded operator from $W^{s,2}(\mathbb{R}^n)$ to $W^{s-|\alpha|,2}(\mathbb{R}^n)$ for $|\alpha| \leq s$. To be more precise $\|D^\alpha f\|_{s-|\alpha|,\rho} \leq \|f\|_{s,\rho}$ for $f \in W^{s,2}(\mathbb{R}^n)$.

Proof.

Note that the result is trivial if s is a non-negative integer because it is obvious that $\|D^\alpha f\|_{s-|\alpha|,2} \leq \|D^\alpha f\|_{s,2}$ for all $f \in W^{s,2}(\mathbb{R}^n)$ by definition of the norm $\|\cdot\|_{s,2}$.

For $f \in W^{s,2}(\mathbb{R}^n)$ and a multi-index α such that $|\alpha| \leq s$, we have that

$$\begin{aligned} \|D^\alpha f\|_{s-|\alpha|,\rho}^2 &= \int_{\mathbb{R}^n} |(D^\alpha f)^\wedge(\mathbf{y})|^2 (1 + \|\mathbf{y}\|^2)^{s-|\alpha|} d\mathbf{y} = \int_{\mathbb{R}^n} |\hat{f}(\mathbf{y})|^2 |\mathbf{y}^\alpha|^2 (1 + \|\mathbf{y}\|^2)^{s-|\alpha|} d\mathbf{y} \\ &\leq \int_{\mathbb{R}^n} |\hat{f}(\mathbf{y})|^2 (1 + \|\mathbf{y}\|^2)^s d\mathbf{y} = \|f\|_{s,\rho}^2 < \infty. \end{aligned} \quad (11.1.5)$$

It follows from (11.1.5) that $f \in W^{s,2}(\mathbb{R}^n)$ implies $D^\alpha f \in W^{s-|\alpha|,2}(\mathbb{R}^n)$ for all multi-indices α such that $|\alpha| \leq s$.

Moreover, (11.1.5) proves that D^α is a bounded operator of norm less than 1 from $W^{s,2}(\mathbb{R}^n)$ to $W^{s-|\alpha|,2}(\mathbb{R}^n)$ for $|\alpha| \leq s$.

Conversely, if $D^\alpha f \in W^{s-|\alpha|,2}(\mathbb{R}^n)$ for all multi-indices α such that $|\alpha| \leq s$, then $f \in W^{s,2}(\mathbb{R}^n)$ if we take $\alpha = (0, 0, \dots, 0)$. ■

11.1.1 Dense Subsets**Definition 11.1.6**

An open set $\Omega \subset \mathbb{R}^n$ satisfies the **segment property** if, for each $\mathbf{x} \in \partial\Omega$, there exists a bounded open neighbourhood $U_{\mathbf{x}} \subset \mathbb{R}^n$ of $\mathbf{x}$ and an element $\mathbf{y} \in \mathbb{R}^n$ such that $\mathbf{z} + \delta\mathbf{y} \in \Omega$ for $0 < \delta < 1$ if $\mathbf{z} \in \overline{\Omega} \cap U_{\mathbf{x}}$ (Figure 11.1).

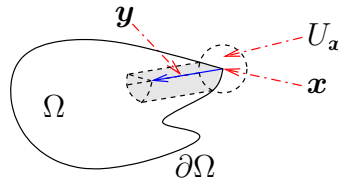


Figure 11.1: Graphical representation of the segment property for an arbitrary set Ω . The region in grey represents the set $\bigcup_{0 \leq \delta \leq 1} (\overline{\Omega} \cap U_{\mathbf{x}} + \delta\mathbf{y})$.

If $\Omega \subset \mathbb{R}^n$ is an open set with a bounded boundary of class C^2 , then Ω has the segment property. Open sets with only continuous boundaries may still have the segment property if the “corners” are not too numerous. A simple example is a square in $\mathbb{R}^2$.

We need a few lemmas before being in a position to prove the main results of this section.

Lemma 11.1.7

Suppose that $\Omega_1 \subset \Omega_2$ are two open subsets of $\mathbb{R}^n$, $k \in \mathbb{N}$ and p is a positive integer. Moreover, suppose that $f \in W^{k,p}(\Omega_1)$ has a compact support in Ω_1 . Let $\underline{f} : \Omega_2 \rightarrow \mathbb{R}$ be the function defined by

$$\underline{f}(\mathbf{x}) = \begin{cases} f(\mathbf{x}) & \text{if } \mathbf{x} \in \Omega_1 \\ 0 & \text{if } \mathbf{x} \in \Omega_2 \setminus \Omega_1 \end{cases}$$

Then $\underline{f} \in W^{k,p}(\Omega_2)$. In particular, $D^\alpha \underline{f} = \underline{D^\alpha f}$ almost everywhere and in the sense of distributions on Ω_2 for $|\alpha| \leq k$.

Proof.

Since $K = \text{supp } f$ is a compact subset of the open set Ω_1 , there exists an open set U such that $K \subset U \subset \bar{U} \subset \Omega_1$. Choose $\psi \in \mathcal{D}(\Omega_2)$ such that $\psi(\mathbf{x}) = 1$ for $\mathbf{x} \in K$ and $\psi(\mathbf{x}) = 0$ for $\mathbf{x} \in \Omega_2 \setminus U$ ¹.

Since

$$\int_{\Omega_2} |\underline{D^\alpha f}(\mathbf{x})|^p d\mathbf{x} = \int_{\Omega_1} |D^\alpha f(\mathbf{x})|^p d\mathbf{x} < \infty$$

for $|\alpha| \leq k$, we get that $\underline{D^\alpha f}$ defines a distribution in $L^p(\Omega_2)$ for $|\alpha| \leq k$. Hence,

$$\begin{aligned} (D^\alpha \underline{f})(\phi) &= (-1)^{|\alpha|} \int_{\Omega_2} \underline{f}(\mathbf{x}) D^\alpha \phi(\mathbf{x}) d\mathbf{x} = (-1)^{|\alpha|} \int_{\Omega_1} f(\mathbf{x}) D^\alpha (\psi\phi)(\mathbf{x}) d\mathbf{x} \\ &= \int_{\Omega_1} D^\alpha f(\mathbf{x}) (\psi\phi)(\mathbf{x}) d\mathbf{x} = \int_{\Omega_1} D^\alpha f(\mathbf{x}) \phi(\mathbf{x}) d\mathbf{x} \\ &= \int_{\Omega_2} \underline{D^\alpha f}(\mathbf{x}) \phi(\mathbf{x}) d\mathbf{x} = (\underline{D^\alpha f})(\phi) \end{aligned}$$

for $\phi \in \mathcal{D}(\Omega_2)$ and $|\alpha| \leq k$. The second and fourth equality are a consequence of $\phi\psi = \phi$ on K . The third equality comes from $f \in W^{k,p}(\Omega_1)$ and $\psi\phi \in \mathcal{D}(\Omega_1)$. Thus $D^\alpha \underline{f} = \underline{D^\alpha f}$ almost everywhere in $\mathbb{R}^n$. ■

It is also interesting to note that, as a distribution, $D^\alpha \underline{f}|_{\Omega_1} = D^\alpha f$.

¹To obtain such a function ψ , let $\{V_j\}_{j=1}^J$ be a finite open cover of $\text{supp } f$ with $V_j \subset \Omega_1$. Such a finite open cover exists because f has a compact support in Ω_1 . Select a partition of unity $\{\psi_j\}_{j=1}^J$ for $\text{supp } f$ subordinate to $\{V_j\}_{j=1}^J$ and set $\psi = \sum_{j=1}^J \psi_j$.

Lemma 11.1.8

Suppose that Ω is an open subset of $\mathbb{R}^n$, $k \in \mathbb{N}$ and p is a positive integer. Moreover, suppose that $f \in W^{k,p}(\Omega)$ is such that $S = \text{supp } f \cap \partial\Omega \neq \emptyset$ and $S \not\subseteq \partial\Omega$ ². Let $\underline{f} : \mathbb{R}^n \rightarrow \mathbb{R}$ be the function defined by

$$\underline{f}(\mathbf{x}) = \begin{cases} f(\mathbf{x}) & \text{if } \mathbf{x} \in \Omega \\ 0 & \text{if } \mathbf{x} \in \mathbb{R}^n \setminus \Omega \end{cases}$$

Then $\underline{f} \in W^{k,p}(V)$, where $V = \mathbb{R}^n \setminus S$. In particular, $D^\alpha \underline{f} = \underline{D^\alpha f}$ almost everywhere and in the sense of distributions on V for $|\alpha| \leq k$.

Proof.

Since

$$\int_V |D^\alpha \underline{f}(\mathbf{x})|^p d\mathbf{x} = \int_\Omega |D^\alpha f(\mathbf{x})|^p d\mathbf{x} < \infty$$

for $|\alpha| \leq k$, we get that $\underline{D^\alpha f}$ defines a distribution included in $L^p(V)$ for $|\alpha| \leq k$.

Given $\phi \in \mathcal{D}(V)$ and $|\alpha| \leq k$, let $K = \text{supp } f \cap \text{supp } \phi$. Since $\text{supp } \phi \subset V$, so $\text{supp } \phi \cap S = \emptyset$, and $\text{supp } f \subset \overline{\Omega}$ with $\text{supp } f \cap \partial\Omega = S$, we get that $K \subset \Omega$. Since K is a closed subset of the compact set $\text{supp } \phi$, we have that K is a compact subset of Ω . Thus, there exists an open set U such that $K \subset U \subset \overline{U} \subset \Omega$. Choose a function $\psi \in \mathcal{D}(\mathbb{R}^n)$ such that $\psi(\mathbf{x}) = 1$ for $\mathbf{x} \in K$ and $\psi(\mathbf{x}) = 0$ for $\mathbf{x} \in \mathbb{R}^n \setminus U$ ³.

We have that

$$\begin{aligned} (D^\alpha \underline{f})(\phi) &= (-1)^{|\alpha|} \int_V \underline{f}(\mathbf{x}) D^\alpha \phi(\mathbf{x}) d\mathbf{x} = (-1)^{|\alpha|} \int_\Omega f(\mathbf{x}) D^\alpha (\psi\phi)(\mathbf{x}) d\mathbf{x} \\ &= \int_\Omega D^\alpha f(\mathbf{x}) (\psi\phi)(\mathbf{x}) d\mathbf{x} = \int_\Omega D^\alpha f(\mathbf{x}) \phi(\mathbf{x}) d\mathbf{x} \\ &= \int_V \underline{D^\alpha f}(\mathbf{x}) \phi(\mathbf{x}) d\mathbf{x} = (\underline{D^\alpha f})(\phi). \end{aligned}$$

The second and fourth equality are a consequence of $\phi\psi = \phi$ in K . The third equality comes from $f \in W^{k,p}(\Omega)$ and $\psi\phi \in \mathcal{D}(\Omega)$. Since $\phi \in \mathcal{D}(V)$ and $|\alpha| \leq k$ are arbitrary, we get $D^\alpha \underline{f} = \underline{D^\alpha f}$ almost everywhere in V . ■

We need a special version of Proposition 6.5.3.

Lemma 11.1.9

Suppose that Ω is an open subset of $\mathbb{R}^n$, U is an open subset such that $\overline{U} \subset \Omega$ is compact, $k \in \mathbb{N}$, p is a positive integer, and $f \in W^{k,p}(\Omega)$.

Given $\psi \in \mathcal{D}(\mathbb{R}^n)$ such that $\psi(\mathbf{x}) = 0$ for $\|\mathbf{x}\| > 1$ and $\int_{\mathbb{R}^n} \psi(\mathbf{x}) d\mathbf{x} = 1$ ⁴, let $\psi_\epsilon(\mathbf{x}) =$

²The lemma is obvious if $\text{supp } f \cap \partial\Omega = \partial\Omega$.

³See footnote on page 402

$\epsilon^{-n}\psi(\epsilon^{-1}\mathbf{x})$ for $\mathbf{x} \in \mathbb{R}^n$ and $\epsilon > 0$. Let $\underline{f} : \mathbb{R}^n \rightarrow \mathbb{R}$ be the function defined by

$$\underline{f}(\mathbf{x}) = \begin{cases} f(\mathbf{x}) & \text{if } \mathbf{x} \in \Omega \\ 0 & \text{if } \mathbf{x} \in \mathbb{R}^n \setminus \Omega \end{cases}$$

Then, $\underline{f} * \psi_\epsilon \in C^\infty(\mathbb{R}^n)$ for all ϵ and $(\underline{f} * \psi_\epsilon)|_U \in W^{k,p}(U)$ for $\epsilon < \text{dist}(\partial\Omega, U)$. Moreover, $\|(\underline{f} * \psi_\epsilon) - f\|_{k,p,U} \rightarrow 0$ as $\epsilon \rightarrow 0^+$.

If f has a compact support, then $\underline{f} * \psi_\epsilon \in \mathcal{D}(\mathbb{R}^n)$ for all ϵ . In particular, $\text{supp}(\underline{f} * \psi_\epsilon) \subset \text{supp } f + \overline{B_\epsilon(\mathbf{0})} \subset \Omega$ for $\epsilon < \text{dist}(\partial U, \text{supp } f)$.

Proof.

We have that $\underline{f}$ is a distribution on $\mathbb{R}^n$ because $\underline{f} \in L^p(\mathbb{R}^n)$. It follows from Proposition 6.5.5 that $\underline{f} * \psi_\epsilon \in C^\infty(\mathbb{R}^n)$.

Given $|\alpha| \leq k$, since $\underline{D^\alpha f} \in L^p(\mathbb{R}^n)$ because $\text{supp } \underline{D^\alpha f} \subset \text{supp } f \subset \Omega$ and $D^\alpha f \in L^p(\Omega)$, we get from Theorem 6.5.2 that $\underline{D^\alpha f} * \psi_\epsilon \in L^p(\mathbb{R}^n)$.

For $\epsilon < \text{dist}(\partial\Omega, U)$, $\phi \in \mathcal{D}(U)$ and $|\alpha| \leq k$, we have that

$$\begin{aligned} D^\alpha(\underline{f} * \psi_\epsilon)(\phi) &= (-1)^{|\alpha|} \int_{\mathbb{R}^n} (\underline{f} * \psi_\epsilon) D^\alpha \phi \, d\mathbf{x} \\ &= (-1)^{|\alpha|} \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} \underline{f}(\mathbf{y} - \mathbf{x}) \psi_\epsilon(\mathbf{y}) \, d\mathbf{y} \right) D^\alpha \phi(\mathbf{x}) \, d\mathbf{x} \\ &= (-1)^{|\alpha|} \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} \underline{f}(\mathbf{y} - \mathbf{x}) D^\alpha \phi(\mathbf{x}) \, d\mathbf{x} \right) \psi_\epsilon(\mathbf{y}) \, d\mathbf{y} \\ &= (-1)^{|\alpha|} \int_{B_\epsilon(\mathbf{0})} \left(\int_U f(\mathbf{y} - \mathbf{x}) D^\alpha \phi(\mathbf{x}) \, d\mathbf{x} \right) \psi_\epsilon(\mathbf{y}) \, d\mathbf{y} \\ &= \int_{B_\epsilon(\mathbf{0})} \left(\int_U D^\alpha f(\mathbf{y} - \mathbf{x}) \phi(\mathbf{x}) \, d\mathbf{x} \right) \psi_\epsilon(\mathbf{y}) \, d\mathbf{y} = \int_U \left(\int_{B_\epsilon(\mathbf{0})} D^\alpha f(\mathbf{y} - \mathbf{x}) \psi_\epsilon(\mathbf{y}) \, d\mathbf{y} \right) \phi(\mathbf{x}) \, d\mathbf{x} \\ &= \int_U \left(\int_{\mathbb{R}^n} \underline{D^\alpha f}(\mathbf{y} - \mathbf{x}) \psi_\epsilon(\mathbf{y}) \, d\mathbf{y} \right) \phi(\mathbf{x}) \, d\mathbf{x} = \int_U (\underline{D^\alpha f} * \psi_\epsilon) \phi \, d\mathbf{x}, \end{aligned}$$

where the fourth and seventh equality are a consequence of $\underline{f}(\mathbf{y} - \mathbf{x}) = f(\mathbf{y} - \mathbf{x})$ and $\underline{D^\alpha f}(\mathbf{y} - \mathbf{x}) = D^\alpha f(\mathbf{y} - \mathbf{x})$ respectively for $\mathbf{x} \in U$ and $\mathbf{y} \in B_\epsilon(\mathbf{0})$ because $U + B_\epsilon(\mathbf{0}) \subset \Omega$ for our choice of ϵ . We have use integration by parts for the fifth equality with $\text{supp } \phi \subset U$.

Since the previous relation is true for all $\phi \in \mathcal{D}(U)$, we get that $D^\alpha(\underline{f} * \psi_\epsilon) = \underline{D^\alpha f} * \psi_\epsilon$ almost everywhere in U . It follows that $D^\alpha(\underline{f} * \psi_\epsilon)|_U \in L^p(U)$. Since this is true for all $|\alpha| \leq k$, we get that $D^\alpha(\underline{f} * \psi_\epsilon)|_U \in W^{p,k}(U)$.

⁴One such function is

$$\psi(\mathbf{x}) = \begin{cases} k e^{-1/(1-\|\mathbf{x}\|^2)} & \text{for } \|\mathbf{x}\| < 1 \\ 0 & \text{for } \|\mathbf{x}\| \geq 1 \end{cases}$$

where $k^{-1} = \int_{\|\mathbf{x}\| < 1} e^{-1/(1-\|\mathbf{x}\|^2)} \, d\mathbf{x}$.

We get from Proposition 6.5.3 that

$$\|D^\alpha(\underline{f} * \psi_\epsilon) - D^\alpha \underline{f}\|_{p,U} = \|\underline{D^\alpha f} * \psi_\epsilon - D^\alpha \underline{f}\|_{p,U} \leq \|\underline{D^\alpha f} * \psi_\epsilon - \underline{D^\alpha f}\|_{p,\mathbb{R}^n} \rightarrow 0 \quad \text{as } \epsilon \rightarrow 0^+.$$

Since this is true for all $|\alpha| \leq k$, we get that $\|\underline{f} * \psi_\epsilon - \underline{f}\|_{p,k,U} \rightarrow 0$ as $\epsilon \rightarrow 0^+$.

If f has a compact support, then $\underline{f} * \psi_\epsilon$ has a compact support because $\text{supp}(\underline{f} * \psi_\epsilon) \subset \text{supp } \underline{f} + \text{supp } \psi_\epsilon \subset \text{supp } f + \overline{B_\epsilon(\mathbf{0})}$. ■

Lemma 11.1.10

Suppose that Ω is an open subset of $\mathbb{R}^n$, U is a bounded open subset of Ω , $k \in \mathbb{N}$, p is a positive integer, and $f \in W^{k,p}(\Omega)$. Given $\mathbf{y} \in \mathbb{R}^n$, suppose that there exists $d > 0$ such that $\delta \mathbf{y} + U \subset \Omega$ for $0 \leq \delta < d$, and define $f_\delta : U \rightarrow \mathbb{R}$ by $f_\delta(\mathbf{x}) = f(\mathbf{x} + \delta \mathbf{y})$ for $\mathbf{x} \in U$ and $0 \leq \delta < d$. Then $f_\delta \in W^{k,p}(U)$ for $0 \leq \delta < d$ and $f_\delta \rightarrow f_0 = f$ in $W^{k,p}(U)$ as $\delta \rightarrow 0^+$.

Proof.

i) We first note that if $g \in L^p(\Omega)$, then $g_\delta \in L^p(U)$ for $0 \leq \delta < d$ where $g_\delta : U \rightarrow \mathbb{R}$ is defined by $g_\delta(\mathbf{x}) = g(\mathbf{x} + \delta \mathbf{y})$ for $\mathbf{x} \in U$. This follows from

$$\int_U |g_\delta|^p d\mathbf{x} = \int_U |g(\mathbf{x} + \delta \mathbf{y})|^p d\mathbf{x} \leq \int_\Omega |g|^p d\mathbf{x} < +\infty$$

because $\delta \mathbf{y} + U \subset \Omega$.

Given $\phi \in \mathcal{D}(U)$ and $|\alpha| \leq k$, we can use integration by parts to get

$$\begin{aligned} D^\alpha f_\delta(\phi) &= (-1)^{|\alpha|} \int_U f(\mathbf{x} + \delta \mathbf{y}) D^\alpha \phi(\mathbf{x}) d\mathbf{x} = \int_U D^\alpha f(\mathbf{x} + \delta \mathbf{y}) \phi(\mathbf{x}) d\mathbf{x} \\ &= \int_U (D^\alpha f)_\delta(\mathbf{x}) \phi(\mathbf{x}) d\mathbf{x} \end{aligned}$$

because $\mathbf{x} + \delta \mathbf{y} \in \Omega$ for all $\mathbf{x} \in U$ and $\text{supp } \phi \subset U$. Thus, $D^\alpha f_\delta = (D^\alpha f)_\delta \in L^p(U)$ ⁵ according to the remark at the beginning of the proof with $g = D^\alpha f$. Since this is true for all multi-indices α such that $|\alpha| \leq k$, we get that $f_\delta \in W^{k,p}(U)$.

ii) We fix the multi-index α such that $|\alpha| \leq k$. Given $\epsilon > 0$, since $\mathcal{D}(\Omega)$ is dense in $L^p(\Omega)$, there exists $\phi \in \mathcal{D}(\Omega)$ such that $\|D^\alpha f - \phi\|_{p,\Omega} < \epsilon/3$. We then have that $\|D^\alpha f - \phi\|_{p,U} \leq \|D^\alpha f - \phi\|_{p,\Omega} < \epsilon/3$ and $\|D^\alpha f_\delta - \phi_\delta\|_{p,U} = \|(D^\alpha f)_\delta - \phi_\delta\|_{p,U} \leq \|D^\alpha f - \phi\|_{p,\Omega} < \epsilon/3$ for all $0 \leq \delta < d$ because $\mathbf{x} + \delta \mathbf{y} \in \Omega$ for all $\mathbf{x} \in U$ when $0 \leq \delta < d$.

Since ϕ has a compact support, it is uniformly continuous. So, there exists δ_0 such that $\|\phi_\delta - \phi\|_{\infty,U} < \epsilon/(3Q)$ for $0 \leq \delta < \delta_0$ and $Q^p = \int_U d\mathbf{x}$. This implies that $\|\phi_\delta - \phi\|_{p,U} < \epsilon/3$ for $0 \leq \delta < \delta_0$.

Hence,

$$\|D^\alpha f_\delta - D^\alpha f\|_{p,U} \leq \|D^\alpha f_\delta - \phi_\delta\|_{p,U} + \|\phi_\delta - \phi\|_{p,U} + \|\phi - D^\alpha f\|_{p,U} < \epsilon$$

⁵As it is so for a sufficiently differentiable function f .

for $0 \leq \delta < \delta_0$. Thus, $\|D^\alpha f_\delta - D^\alpha f\|_{p,U} \rightarrow 0$ as $\delta \rightarrow 0^+$. Since this is true for all multi-indices α such that $|\alpha| \leq k$, we get that $\|D^\alpha f_\delta - D^\alpha f\|_{k,p,U} \rightarrow 0$ as $\delta \rightarrow 0^+$. ■

Theorem 11.1.11

Suppose that $k \in \mathbb{N}$ and p is a positive integer. If $\Omega \subset \mathbb{R}^n$ satisfies the segment property, then $\mathcal{D}(\mathbb{R}^n)|_\Omega = \{\phi|_\Omega : \phi \in \mathcal{D}(\mathbb{R}^n)\}$ is dense in $W^{k,p}(\Omega)$ with respect to the norm $\|\cdot\|_{k,p,\Omega}$.

Remark 11.1.12

Since elements of $W^{k,p}(\mathbb{R}^n)$ are equivalence class of functions which are equal almost everywhere, the statement $\mathcal{D}(\mathbb{R}^n) \subset W^{k,p}(\mathbb{R}^n)$ has a special meaning. When view as an element of $W^{k,p}(\mathbb{R}^n)$, a function $\phi \in \mathcal{D}(\mathbb{R}^n)$ is associated to its equivalence class in $W^{k,p}(\mathbb{R}^n)$. We will regularly make this type of interpretation from now on. ♠

Proof.

i) We first shown that every member of $W^{k,p}(\Omega)$ can be approximated by members of $W^{k,p}(\Omega)$ with bounded support.

Choose $\phi \in \mathcal{D}(\mathbb{R}^n)$ such that $\phi(\mathbf{x}) = 1$ for $\|\mathbf{x}\|_2 \leq 1$ and $\phi(\mathbf{x}) = 0$ for $\|\mathbf{x}\|_2 > 2$, let M be a constant such that $|D^\alpha \phi(\mathbf{x})| \leq M$ for all $\mathbf{x} \in \mathbb{R}^n$ and $|\alpha| \leq k$.

Given $\delta > 0$, let $\phi_\delta \in \mathcal{D}(\mathbb{R}^n)$ be the function defined $\phi_\delta(\mathbf{x}) = \phi(\delta\mathbf{x})$ for $\mathbf{x} \in \mathbb{R}^n$. For $\delta \leq 1$, we have that $|D^\alpha \phi_\delta(\mathbf{x})| \leq \delta^{|\alpha|} M \leq M$ for all $\mathbf{x} \in \mathbb{R}^n$ and $|\alpha| \leq k$.

Given $f \in W^{k,p}(\Omega)$, let $f_\delta = \phi_\delta f$ on Ω . We have that f_δ has a bounded support; namely, $\phi_\delta f \subset \Omega \cap B_{2/\delta}(\mathbf{0})$. Moreover, $\phi_\delta f \in W^{k,p}(\Omega)$ because

$$D^\alpha f_\delta = \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} D^{\alpha-\beta} \phi_\delta D^\beta f ,$$

where $\beta \leq \alpha$ means $\beta_j \leq \alpha_j$ for all j and $\binom{\alpha}{\beta} = \prod_{j=1}^n \binom{\alpha_j}{\beta_j}$. We are obviously talking about derivative in the sense of distributions. In particular, we have that

$$\|D^\alpha f_\delta\|_p \leq M \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} \|D^\beta f\|_p < \infty$$

for $0 < \delta \leq 1$ and $|\alpha| \leq k$. Moreover, it follows from the Lebesgue Dominated Convergence Theorem that

$$\|f - f_\delta\|_{k,p,\Omega} \leq 2 \left(\sum_{|\alpha| \leq k} \int_{\mathbf{x} \in \Omega \setminus B_{1/\delta}(\mathbf{0})} |D^\alpha f(\mathbf{x})|^p d\mathbf{x} \right)^{1/p} \rightarrow 0 \quad \text{as } \delta \rightarrow 0^+$$

because $D^\alpha f \in L^p(\Omega)$ for all $|\alpha| \leq k$. Hence, we have shown that every member of $W^{k,p}(\Omega)$ can be approximated by members of $W^{k,p}(\Omega)$ with bounded support.

ii) From now on, we assume that $f \in W^{k,p}(\Omega)$ has a bounded support. Let $K = \text{supp } f \subset \overline{\Omega}$. For each bounded open set U_x provided in the definition of segment property, choose an open

set $V_{\mathbf{x}}$ such that $\overline{V_{\mathbf{x}}} \subset U_{\mathbf{x}}$. Let $Q = K \setminus \bigcup_{\mathbf{x} \in \partial\Omega} V_{\mathbf{x}}$. We have that $Q \subset \Omega$ because $\partial\Omega \subset \bigcup_{\mathbf{x} \in \partial\Omega} V_{\mathbf{x}}$. Hence, Q is a compact subset of Ω because it is bounded and closed. Choose an open set U such that $Q \subset U \subset \overline{U} \subset \Omega$. We get an open cover $\{U\} \cup \left\{ \bigcup_{\mathbf{x} \in \partial\Omega} V_{\mathbf{x}} \right\}$ of K . Being closed and bounded, K is compact. So, there exists a finite subcover of K that we denote $\{V_i\}_{i=1}^I$. At most one of the V_i may have an empty intersection with $\partial\Omega$; namely, the open set U introduced before.

Let $\{\phi_i\}_{i=1}^I$ be a partition of unity of K subordinate to $\{V_i\}_{i=1}^I$ and let $f_i = \phi_i f$. We have that $f_i \in W^{k,p}(\Omega)$ with $\text{supp } f_i \subset V_i \cap \overline{\Omega}$. Given $\epsilon > 0$, if we can find $\xi_i \in \mathcal{D}(\mathbb{R}^n)$ such that $\|f_i - \xi_i\|_{k,p,\Omega} < \epsilon/I$ for all i , then we will have that $\xi = \sum_{i=1}^I \xi_i \in \mathcal{D}(\mathbb{R}^n)$ and $\|f - \xi\|_{k,p,\Omega} \leq \sum_{i=1}^I \|f_i - \xi_i\|_{k,p,\Omega} < \epsilon$. Because ϵ is arbitrary, this will prove that $\mathcal{D}(\mathbb{R}^n)|_{\Omega}$ is dense in $W^{k,p}(\Omega)$.

iii) Consider $\underline{f}_i : \mathbb{R}^n \rightarrow \mathbb{R}$ defined by

$$\underline{f}_i(\mathbf{x}) = \begin{cases} f_i(\mathbf{x}) & \text{if } \mathbf{x} \in \Omega \\ 0 & \text{if } \mathbf{x} \in \mathbb{R}^n \setminus \Omega \end{cases}$$

If $\text{supp } f_i \subset \Omega$, then we have from Lemma 11.1.7 that $\underline{f}_i \in W^{k,p}(\mathbb{R}^n)$ and from Lemma 11.1.9 (with Ω and U replaced by $\mathbb{R}^n$ and Ω respectively) that $\underline{f}_i * \psi_{\epsilon} \in \mathcal{D}(\mathbb{R}^n)$ and

$$\|\underline{f}_i * \psi_{\epsilon} - f_i\|_{k,p,\Omega} = \|\underline{f}_i * \psi_{\epsilon} - \underline{f}_i\|_{k,p,\Omega} \quad \text{as } \epsilon \rightarrow 0^+.$$

We choose ξ_i to be $\underline{f}_i * \psi_{\epsilon}$ for ϵ small enough to get $\|\xi_i - f_i\|_{k,p,\Omega} = \|\underline{f}_i * \psi_{\epsilon} - f_i\|_{k,p,\Omega} < \epsilon/I$.

We now suppose that $S = \text{supp } f_i \cap \partial\Omega \neq \emptyset$. It follows from Lemma 11.1.8 that $\underline{f}_i \in W^{k,p}(V)$, where $V = \mathbb{R}^n \setminus S$.

Let $\mathbf{y}_i$ be the vector provided in the definition of the segment property such that $\delta\mathbf{y} + (\overline{\Omega} \cap U_i) \subset \Omega$ for $0 < \delta < 1$. We define $\underline{f}_{i,\delta} \in W^{k,p}(V - \delta\mathbf{y}_i)$ by $\underline{f}_{i,\delta}(\mathbf{x}) = \underline{f}_i(\mathbf{x} + \delta\mathbf{y}_i)$ for $\mathbf{x} \in V - \delta\mathbf{y}_i$.

Let $d = \min\{1, \|\mathbf{y}_i\|^{-1} \text{dist}(V_i, \partial U_i)\}$. We then have that $\delta\mathbf{y}_i + \overline{\Omega} \subset V$ for $0 < \delta < d$ because $S \subset V_i \cap \partial\Omega$ and $(V_i \cap \partial\Omega) - \delta\mathbf{y}_i \subset U_i \setminus \overline{\Omega}$ for $0 < \delta < d$. Thus $S - \delta\mathbf{y}_i \subset U_i \setminus \overline{\Omega}$ for $0 < \delta < d$ (Figure 11.2); namely, $S \cap (\delta\mathbf{y} + \overline{\Omega}) = \emptyset$.

From Lemma 11.1.10 (with Ω and U replaced by V and Ω respectively), we get that $\underline{f}_{i,\delta} \rightarrow \underline{f}_i = f_i$ in $W^{k,p}(\Omega)$ as $\delta \rightarrow 0$. So, there exists δ_i such that $\|\underline{f}_{i,\delta_i} - f_i\|_{k,p,\Omega} < \epsilon/(2I)$.

Since $\overline{\Omega} \subset V - \delta_i\mathbf{y}_i$, it follows from Lemma 11.1.9 (with Ω and U replaced by $V - \delta_i\mathbf{y}_i$ and Ω respectively) that $\underline{f}_{i,\delta_i} * \psi_{\epsilon} \rightarrow \underline{f}_{i,\delta_i}$ in $W^{k,p}(\Omega)$ as $\epsilon \rightarrow 0$. Choose ξ_i to be $\underline{f}_{i,\delta_i} * \psi_{\epsilon}$ for ϵ small enough to get $\|\underline{f}_{i,\delta_i} - \xi_i\|_{k,p,\Omega} < \epsilon/(2I)$. Then $\|f_i - \xi_i\|_{k,p,\Omega} \leq \|f_i - \underline{f}_{i,\delta_i}\|_{k,p,\Omega} + \|\underline{f}_{i,\delta_i} - \xi_i\|_{k,p,\Omega} < \epsilon/I$. ■

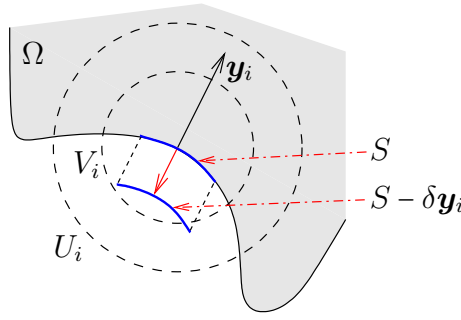


Figure 11.2: Illustration of the translation of the region S mentioned in the proof of Theorem 11.1.11 for $0 < \delta < d$.

As the following example shows, $\mathcal{D}(\Omega)$ is not dense in $W^{k,2}(\Omega)$ for k a positive integer.

Example 11.1.13

Let k be a positive integer, and p be a positive integer or ∞ . If $\Omega \subset \mathbb{R}^n$ is a bounded open set, then $\mathcal{D}(\Omega)$ is not dense in $W^{k,p}(\Omega)$ with respect to the norm $\|\cdot\|_{k,p,\Omega}$.

To prove this statement, consider the continuous linear functional

$$T : W^{k,p}(\Omega) \rightarrow \mathbb{R}$$

$$v \mapsto \int_{\Omega} \left(\sum_{|\alpha| \leq k} D^{\alpha} u(\mathbf{y}) D^{\alpha} v(\mathbf{y}) \right) d\mathbf{y}$$

where $u \in C^{\infty}(\overline{\Omega})$ is a non-trivial function satisfying

$$\sum_{|\alpha| \leq k} (-1)^{\alpha} D^{2\alpha} u(\mathbf{y}) = 0, \quad (11.1.6)$$

for $\mathbf{y} \in \Omega$. Such function u can be found by the method of separation of variables applied to 11.1.6 on an open rectangular region containing Ω for instance.

It follows from integration by parts that $T(v) = 0$ for all $v \in \mathcal{D}(\Omega)$. However, $u \in W^{k,p}(\Omega)$ because Ω is bounded and $u \in C^{\infty}(\overline{\Omega})$. Thus, T is not the trivial functional because $T(u) > 0$.

If $\mathcal{D}(\Omega)$ was dense in $W^{k,p}(\Omega)$, we would have that $T = 0$ on $W^{k,p}(\Omega)$ because $T = 0$ on $\mathcal{D}(\Omega)$. But this is not the case. ♣

The following lemma is a special case of Lemma 11.1.9 with $\Omega = U = \mathbb{R}^n$ if we assume that the distance between a set and the empty set is ∞ . Nevertheless, we state it because there is an easy proof in this special case.

Lemma 11.1.14

Suppose that $k \in \mathbb{N}$, p is a positive integer, and $f \in W^{k,p}(\mathbb{R}^n)$. Given $\psi \in \mathcal{D}(\mathbb{R}^n)$ such that $\psi(\mathbf{x}) = 0$ for $\|\mathbf{x}\| > 1$ and $\int_{\mathbb{R}^n} \psi(\mathbf{x}) d\mathbf{x} = 1$, let $\psi_{\epsilon}(\mathbf{x}) = \epsilon^{-n} \psi(\epsilon^{-1} \mathbf{x})$ for $\mathbf{x} \in \mathbb{R}^n$ and

$\epsilon > 0$. Then $f * \psi_\epsilon \in W^{k,p}(\mathbb{R}^n) \cap C^\infty(\mathbb{R}^n)$ for all ϵ and $\|(f * \psi_\epsilon) - f\|_{k,p} \rightarrow 0$ as $\epsilon \rightarrow 0^+$. If f has a compact support, then $f * \psi_\epsilon \in \mathcal{D}(\mathbb{R}^n)$ for all ϵ ; more precisely, $\text{supp } f * \psi_\epsilon \subset \text{supp } f + B_\epsilon(\mathbf{0})$.

Proof.

We have from Young's inequality, Theorem 6.5.2, that $D^\alpha f * \psi_\epsilon \in L^p(\mathbb{R}^n)$ for all $|\alpha| \leq k$ because $f \in W^{k,p}(\mathbb{R}^n)$ implies that $D^\alpha f \in L^p(\mathbb{R}^n)$ for all $|\alpha| \leq k$. Moreover, from Proposition 6.5.6, we have $D^\alpha(f * \psi_\epsilon) = D^\alpha f * \psi_\epsilon$ in the sense of distributions. Thus $f * \psi_\epsilon \in W^{k,p}(\mathbb{R}^n)$.

We have from Proposition 6.5.3 that

$$\|D^\alpha(f * \psi_\epsilon) - D^\alpha f\|_p = \|D^\alpha f * \psi_\epsilon - D^\alpha f\|_p \rightarrow 0 \quad \text{as } \epsilon \rightarrow 0^+$$

for $|\alpha| \leq k$. Thus, $\|(f * \psi_\epsilon) - f\|_{k,p} \rightarrow 0$ as $\epsilon \rightarrow 0^+$.

It follows from Proposition 6.5.5 that $f * \phi_\epsilon \in C^\infty(\mathbb{R}^n)$ for all $\epsilon > 0$. If f has a compact support, then $f * \phi_\epsilon$ has a compact support because ϕ_ϵ has a compact support. To be precise, $\text{supp}(f * \phi_\epsilon) \subset \text{supp } f + \text{supp } \phi_\epsilon$. ■

The next proposition is a special case of Theorem 11.1.11 if $\Omega = \mathbb{R}^n$ and if we assume that a set with an empty boundary automatically satisfy the segment property. The proof for this special case is a lot simpler than the proof for Theorem 11.1.11. So, we give it.

Proposition 11.1.15

For $k \in \mathbb{N}$ and $0 < p < \infty$, $\mathcal{D}(\mathbb{R}^n)$ is dense in $W^{k,p}(\mathbb{R}^n)$ with respect to the norm $\|\cdot\|_{k,p}$.

Proof.

As in the proof of Theorem 11.1.11, we have that every member of $W^{k,p}(\mathbb{R}^n)$ can be approximated by members of $W^{k,p}(\mathbb{R}^n)$ with bounded support.

We now show that every member of $W^{k,p}(\mathbb{R}^n)$ with compact support can be approximated by members of $\mathcal{D}(\mathbb{R}^n)$.

Suppose that $f \in W^{k,p}(\mathbb{R}^n)$ has compact support. Choose $\phi \in \mathcal{D}(\mathbb{R}^n)$ such that $\int_{\mathbb{R}^n} \phi(\mathbf{x}) d\mathbf{x} = 1$ and set $\phi_\epsilon(\mathbf{x}) = \epsilon^{-n} \phi(\epsilon^{-1} \mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^n$. It follows from Proposition 6.5.5 that $f * \phi_\epsilon \in C^\infty(\mathbb{R}^n)$ for all $\epsilon > 0$. Since both f and ϕ_ϵ have compact support, $f * \phi_\epsilon$ has compact support. Thus, $f * \phi_\epsilon \in \mathcal{D}(\mathbb{R}^n)$. Moreover, from Proposition 6.5.3, we have that

$$D^\alpha(f * \phi_\epsilon) = (D^\alpha f) * \phi_\epsilon \rightarrow D^\alpha f \quad \text{as } \epsilon \rightarrow 0^+$$

in $L^p(\mathbb{R}^n)$ for all $|\alpha| \leq k$. Thus, $f * \phi_\epsilon \rightarrow f$ in $W^{k,p}(\mathbb{R}^n)$ as $\epsilon \rightarrow 0^+$. ■

The next theorem is a version of Proposition 11.1.15 for $W^{s,2}(\mathbb{R}^n)$ with s real. Its proof is remarkably simple.

Theorem 11.1.16

For $s \in \mathbb{R}$, $\mathcal{D}(\mathbb{R}^n)$ is dense in $W^{s,2}(\mathbb{R}^n)$ with respect to the norm $\|\cdot\|_{s,\rho}$.

Proof.

We consider the operator

$$\begin{aligned} T^{[s]} : W^{s,2}(\mathbb{R}^n) &\rightarrow L^2(\mathbb{R}^n) \\ f &\mapsto \mathcal{F}^{-1}\left((1 + \|\mathbf{y}\|_2^2)^{s/2} \mathcal{F}(f)\right) \end{aligned}$$

This is an isometry because

$$\|T^{[s]}(f)\|_2 = \|\mathcal{F}^{-1}\left((1 + \|\mathbf{y}\|_2^2)^{s/2} \mathcal{F}(f)\right)\|_2 = \|(1 + \|\mathbf{y}\|_2^2)^{s/2} \mathcal{F}(f)\|_2 = \|f\|_{s,\rho},$$

where the second equality is a consequence of Plancherel theorem, Theorem 6.6.6, and the last one is just the definition of $\|\cdot\|_{s,\rho}$.

It follows from Proposition 6.6.11 that the Fourier transform is a homeomorphism from $\mathcal{S}(\mathbb{R}^n)$ onto itself. It is also clear that $\phi \mapsto (1 + \|\mathbf{y}\|_2^2)^{s/2} \phi$ is a homeomorphism from $\mathcal{S}(\mathbb{R}^n)$ onto itself. Thus, $T^{[s]}$ is a homeomorphism from $\mathcal{S}(\mathbb{R}^n)$ onto itself.

Since $\mathcal{S}(\mathbb{R}^n)$ is dense in $L^2(\mathbb{R}^n)$ according to the norm $\|g\|_2 = \left(\int_{\mathbb{R}^n} |g(\mathbf{x})|^2 d\mathbf{x}\right)^{1/2}$ for $g \in L^2(\mathbb{R}^n)$, it follows from the isometry $T^{[s]}$ that $\mathcal{S}(\mathbb{R}^n)$ is dense in $W^{s,2}(\mathbb{R}^n)$.

Moreover, $\mathcal{D}(\mathbb{R}^n)$ is dense in $\mathcal{S}(\mathbb{R}^n)$ according to the norm $\|\cdot\|_{s,\rho}$. To prove this, choose $\phi \in \mathcal{D}(\mathbb{R}^n)$ such that $|\phi(\mathbf{x})| \leq 1$ for all $\mathbf{x} \in \mathbb{R}^n$, $\phi(\mathbf{x}) = 1$ for $\|\mathbf{x}\| < 1$ and $\phi(\mathbf{x}) = 0$ for $\|\mathbf{x}\| > 2$. Let $\phi_\delta(\mathbf{x}) = \phi(\delta\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^n$ and $\delta > 0$. If $\psi \in \mathcal{S}(\mathbb{R}^n)$, then $\psi_\delta = \psi\phi_\delta \in \mathcal{D}(\mathbb{R}^n)$ and $\psi_\delta - \psi \rightarrow 0$ as $\delta \rightarrow 0^+$ in $L^2(\mathbb{R}^n)$ follows from the Lebesgue Dominate Convergence Theorem since $\psi_\delta(\mathbf{x}) \rightarrow \psi(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^n$ as $\delta \rightarrow 0^+$ and $\|\psi_\delta\|_2 \leq \|\psi\|_2 < \infty$ for all δ . Since

$$\begin{aligned} P : \mathcal{S}(\mathbb{R}^n) &\rightarrow \mathcal{S}(\mathbb{R}^n) \\ \zeta &\mapsto (1 + \|\mathbf{y}\|_2^2)^{s/2} \mathcal{F}(\zeta)(\mathbf{y}) \end{aligned}$$

is continuous, we get that $P(\psi_\delta - \psi) \rightarrow 0$ in $\mathcal{S}(\mathbb{R}^n)$ as $\delta \rightarrow 0^+$ according to the norm $\|\cdot\|_2$. This implies that $\|\psi - \psi_\delta\|_{s,\rho} = \|P(\psi - \psi_\delta)\|_2 \rightarrow 0$ as $\delta \rightarrow 0^+$. ■

Proposition 11.1.17

Let $\Omega \subset \mathbb{R}^n$ be an open set and $\phi \in \mathcal{D}(\Omega)$. For $k \in \mathbb{N}$ and p a positive integer or ∞ ,

$$\begin{aligned} T : W^{k,p}(\Omega) &\rightarrow W^{k,p}(\Omega) \\ f &\mapsto \phi f \end{aligned}$$

is a bounded linear mapping.

Proof.

The result follows from the Leibniz formula ⁶, and the fact that $\phi \in \mathcal{D}(\Omega)$ and all its derivatives are bounded on Ω . ■

Lemma 11.1.18

Let $\Omega_1 \subset \Omega_2$ be two open subsets of $\mathbb{R}^n$. Then, $W^{k,p}(\Omega_2)|_{\Omega_1} \subset W^{k,p}(\Omega_1)$. More precisely, if $f \in W^{k,p}(\Omega_2)$, then $f|_{\Omega_1} \in W^{k,p}(\Omega_1)$ and $D^\alpha(f|_{\Omega_1}) = (D^\alpha f)|_{\Omega_1}$ for $|\alpha| \leq k$.

Proof.

Given $f \in W^{k,p}(\Omega_2)$, we have that $(D^\alpha f)|_{\Omega_1} \in L^p(\Omega_1)$ for all multi-indices α such that $|\alpha| \leq k$ because $\int_{\Omega_1} |D^\alpha f|^p d\mathbf{x} \leq \int_{\Omega_2} |D^\alpha f|^p d\mathbf{x} < \infty$ for all α such that $|\alpha| \leq k$. In particular, this implies that $f|_{\Omega_1}$ defines a proper distribution on Ω_1 .

Moreover, for $|\alpha| \leq k$, we have that

$$\begin{aligned} D^\alpha(f|_{\Omega_1})(\phi) &= (-1)^{|\alpha|} \int_{\Omega_1} f|_{\Omega_1}(\mathbf{x}) D^\alpha \phi(\mathbf{x}) d\mathbf{x} = (-1)^{|\alpha|} \int_{\Omega_2} f(\mathbf{x}) D^\alpha \phi(\mathbf{x}) d\mathbf{x} \\ &= \int_{\Omega_2} D^\alpha f(\mathbf{x}) \phi(\mathbf{x}) d\mathbf{x} = \int_{\Omega_1} (D^\alpha f(\mathbf{x}))|_{\Omega_1} \phi(\mathbf{x}) d\mathbf{x} \end{aligned}$$

for all $\phi \in \mathcal{D}(\Omega_1) \subset \mathcal{D}(\Omega_2)$, where we assume that a test function ϕ on Ω_1 is extended to a test function ϕ on Ω_2 by $\phi(\mathbf{x}) = 0$ for $\mathbf{x} \in \Omega_2 \setminus \Omega_1$. Hence, $D^\alpha(f|_{\Omega_1}) = (D^\alpha f)|_{\Omega_1}$. ■

Theorem 11.1.19

Let Ω be an open subset of $\mathbb{R}^n$, $k \in \mathbb{N}$ and $0 < p < \infty$. Then $C^\infty(\Omega) \cap W^{k,p}(\Omega)$ is dense in $W^{k,p}(\Omega)$.

Proof.

We prove that $S = \{f \in C^\infty(\Omega) : \|f\|_{k,p} < \infty\}$ is dense in $W^{k,p}(\Omega)$. Consider $f \in W^{k,p}(\Omega)$ and $\delta > 0$.

i) Let $\Omega_{-1} = \Omega_0 = \emptyset$ and

$$\Omega_j = \left\{ \mathbf{x} \in \Omega : \|\mathbf{x}\| < j \text{ and } \text{dist}(\mathbf{x}, \partial\Omega) > \frac{1}{j} \right\}$$

for $j > 0$. Moreover, let $V_j = \Omega_{j+1} \setminus \overline{\Omega_{j-1}}$ for $j > 0$. The collection $\{V_j\}_{j=1}^\infty$ is an open cover of Ω . Let $\{\phi_j\}_{j=1}^\infty$ be a partition of unity of Ω subordinate to $\{V_j\}_{j=1}^\infty$. We are using here the Definition 9.9.13 of a partition of unity.

⁶The reader can easily verify that the Leibnitz formula that we already know for the product of two differentiable functions is also true for the product ϕu with $\phi \in \mathcal{D}(\Omega)$ and $u \in \mathcal{D}'(\Omega)$.

Choose $\psi \in \mathcal{D}(\mathbb{R}^n)$ such that $\psi(\mathbf{x}) = 0$ for $\|\mathbf{x}\| > 1$ and $\int_{\mathbb{R}^n} \psi(\mathbf{x}) d\mathbf{x} = 1$. We consider $\psi_\epsilon(\mathbf{x}) = \epsilon^{-n} \psi(\epsilon^{-1} \mathbf{x})$ for $\mathbf{x} \in \mathbb{R}^n$ and $\epsilon > 0$.

ii) Choose $\delta > 0$. Consider $\underline{f} : \mathbb{R}^n \rightarrow \mathbb{R}$ defined by

$$\underline{f}(\mathbf{x}) = \begin{cases} f(\mathbf{x}) & \text{if } \mathbf{x} \in \Omega \\ 0 & \text{if } \mathbf{x} \in \mathbb{R}^n \setminus \Omega \end{cases}$$

We show that, for some set $\{\epsilon_j\}_{j=1}^\infty \subset \mathbb{R}$ to be specified later,

$$\phi = \sum_{j=1}^\infty (\phi_j \underline{f}) * \psi_{\epsilon_j} \quad (11.1.7)$$

is in $C^\infty(\Omega) \cap W^{k,p}(\Omega)$ and satisfies $\|f - \phi\|_{k,p} \leq \delta$ on Ω . As usual, we may assume that ϕ_j is defined on $\mathbb{R}^n$ by $\phi_j(\mathbf{x}) = 0$ for $\mathbf{x} \in \mathbb{R}^n \setminus \Omega$. It follows from Proposition 11.1.17 that $\phi_j \underline{f} \in W^{k,p}(\mathbb{R}^n)$ and from Lemma 11.1.14 that $(\phi_j \underline{f}) * \psi_{\epsilon_j} \in W^{k,p}(\mathbb{R}^n) \cap C^\infty(\mathbb{R}^n)$.

If $\epsilon_j < \frac{1}{(j+1)(j+2)}$ and $U_j = \Omega_{j+2} \setminus \overline{\Omega_{j-2}}$ for $j > 0$, then

$$\text{supp}((\phi_j \underline{f}) * \psi_{\epsilon_j}) \subset \text{supp}(\phi_j \underline{f}) + \text{supp} \psi_{\epsilon_j} \subset \{\mathbf{x} : \text{dist}(\mathbf{x}, V_j) < \epsilon_j\} \subset U_j$$

as can be visualized in Figure 11.3. Note that this also implies that $\text{supp} \phi \subset \overline{\Omega}$. If $K \subset \Omega$ is a compact set, then there are only a finite number of U_j that intersect K . In fact, $U_j \cap K = \emptyset$ for $j > q+1$ such that $K \subset \Omega_q$. Thus, the sum in 11.1.7 is finite on K . Therefore, ϕ is C^∞ on the interior of K since each term of the series is in $C^\infty(\mathbb{R}^n)$. Finally, since this is true for any compact subset of Ω , then $\phi \in C^\infty(\Omega)$. Since $\phi \in L^1_{loc}(\Omega)$, we also get that $\phi \in \mathcal{D}'(\Omega)$.

We use Lemma 11.1.9 to choose $\epsilon_j < \frac{1}{(j+1)(j+2)}$ such that

$$\|((\phi_j \underline{f}) * \psi_{\epsilon_j}) - (\phi_j f)\|_{k,p,\Omega} < \frac{\delta}{2^j Q}$$

for $j > 0$, where $Q^p = \sum_{|\alpha| \leq k} 1$.

Consider $h_{j-1} : \Omega \rightarrow \mathbb{R}$ defined by

$$h_{j-1}(\mathbf{x}) = \begin{cases} (f - \phi)(\mathbf{x}) & \text{if } \mathbf{x} \in \Omega_{j-1} \\ 0 & \text{if } \mathbf{x} \in \Omega \setminus \Omega_{j-1} \end{cases}$$

Since

$$h_{j-1}|_{\Omega_{j-1}} = (f - \phi)|_{\Omega_{j-1}} = f - \sum_{i=1}^j (\phi_i \underline{f}) * \psi_{\epsilon_i} = \sum_{i=1}^j \phi_i f - \sum_{i=1}^j (\phi_i \underline{f}) * \psi_{\epsilon_i}$$

on Ω_{j-1} because $\sum_{i=1}^j \phi_i(\mathbf{x}) = 1$ for all $\mathbf{x} \in \Omega_{j-1}$, we get

$$\|D^\alpha h_{j-1}\|_{p,\Omega} = \|D^\alpha h_{j-1}\|_{p,\Omega_{j-1}} \leq \|f - \phi\|_{k,p,\Omega_{j-1}} = \left\| \sum_{i=1}^j \phi_i f - \sum_{i=1}^j (\phi_i \underline{f}) * \psi_{\epsilon_i} \right\|_{k,p,\Omega_{j-1}}$$

$$\leq \left\| \sum_{i=1}^j (\phi_i f - (\phi_i \underline{f}) * \psi_{\epsilon_i}) \right\|_{k,p,\Omega} \leq \sum_{i=1}^j \|(\phi_i f - (\phi_i \underline{f}) * \psi_{\epsilon_i})\|_{k,p,\Omega} \leq \sum_{i=1}^{\infty} \frac{\delta}{2^i Q} = \frac{\delta}{Q} \quad (11.1.8)$$

for all multi-indices α such that $|\alpha| \leq k$. Thus $h_{j-1} \in W^{k,p}(\Omega)$ for all j . Note that $\sum_{i=1}^j (\phi_i \underline{f}) * \psi_{\epsilon_i} \in C^\infty(\mathbb{R}^n)$ and has a compact support in Ω_{j-1} .

Since $D^\alpha h_{j-1}(\mathbf{x}) \rightarrow D^\alpha(f - \phi)(\mathbf{x})$ as $j \rightarrow \infty$ for all $\mathbf{x} \in \Omega$ and multi-indices α with $|\alpha| \leq k$, we get from (11.1.8) and the Lebesgue Monotone Convergence Theorem that $\|D^\alpha(f - \phi)\|_{p,\Omega} \leq \delta/Q$ for all multi-indices α with $|\alpha| \leq k$. Hence, $\|f - \phi\|_{k,p,\Omega} \leq \delta$.

iii) We have that $\phi \in S$ because $f \in W^{k,p}(\Omega)$ and $\phi - f \in W^{k,p}(\Omega)$ by (ii), thus $\phi = (\phi - f) + f \in W^{k,p}(\Omega)$. In particular,

$$\|\phi\|_{k,p} = \|(\phi - f) + f\|_{k,p} \leq \|(\phi - f)\|_{k,p} + \|f\|_{k,p} \leq \delta + \|f\|_{k,p} < \infty. \quad \blacksquare$$

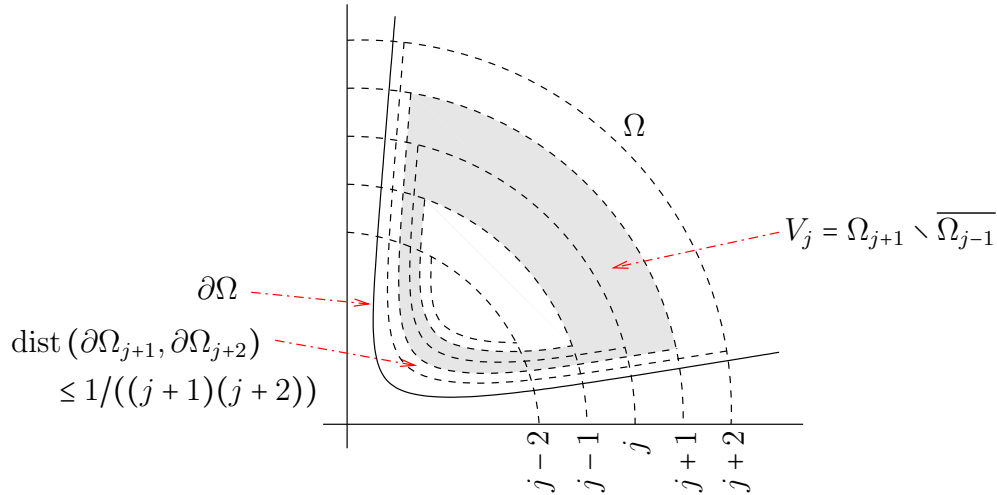


Figure 11.3: The open cover used in the proof of Theorem 11.1.19.

11.1.2 Localized Sobolev Spaces

Definition 11.1.20

Let Ω be an open subset of $\mathbb{R}^n$. If $k \in \mathbb{N}$ and p is a positive integer or ∞ , the **localized Sobolev space** $W_{loc}^{k,p}(\Omega)$ is the set of all distributions $u \in \mathcal{D}'(\Omega)$ with the following property. For every bounded open set V such that $\overline{V} \subset \Omega$, there exists $g \in W^{k,p}(\mathbb{R}^n)$ such that $u = g$ on V in the sense of distributions ⁷.

Similarly, if $s \in \mathbb{R}$, the **localized Sobolev space** $W_{loc}^{s,2}(\Omega)$ is the set of all distributions $u \in \mathcal{D}'(\Omega)$ with the following property. For every bounded open set V such that $\overline{V} \subset \Omega$,

there exists $g \in W^{s,2}(\mathbb{R}^n)$ such that $u = g$ on V in the sense of distributions ⁷.

Proposition 11.1.17 is also true for the Sobolev spaces $W^{s,2}(\mathbb{R}^n)$ with $s \in [0, \infty[$. To prove it, the following technical lemma is needed first.

Lemma 11.1.21

For all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $m \in \mathbb{R}$, we have $\left(\frac{1 + \|\mathbf{y}\|^2}{1 + \|\mathbf{x}\|^2} \right)^m \leq 2^{|m|} (1 + \|\mathbf{y} - \mathbf{x}\|^2)^{|m|}$.

Proof.

It is enough to prove the lemma for $m \geq 0$. If $m < 0$, it suffices to interchange $\mathbf{x}$ and $\mathbf{y}$.

It follows from $(\|\mathbf{y} - \mathbf{x}\| - \|\mathbf{x}\|)^2 \geq 0$ that $2\|\mathbf{x}\| \|\mathbf{y} - \mathbf{x}\| \leq \|\mathbf{y} - \mathbf{x}\|^2 + \|\mathbf{x}\|^2$. Hence, from $\|\mathbf{y}\| \leq \|\mathbf{y} - \mathbf{x}\| + \|\mathbf{x}\|$, we get that

$$\|\mathbf{y}\|^2 \leq (\|\mathbf{y} - \mathbf{x}\| + \|\mathbf{x}\|)^2 = \|\mathbf{y} - \mathbf{x}\|^2 + \|\mathbf{x}\|^2 + 2\|\mathbf{y} - \mathbf{x}\| \|\mathbf{x}\| \leq 2(\|\mathbf{y} - \mathbf{x}\|^2 + \|\mathbf{x}\|^2) .$$

Then,

$$1 + \|\mathbf{y}\|^2 \leq 1 + 2(\|\mathbf{y} - \mathbf{x}\|^2 + \|\mathbf{x}\|^2) \leq 2(1 + \|\mathbf{y} - \mathbf{x}\|^2)(1 + \|\mathbf{x}\|^2) .$$

It follows that

$$\frac{1 + \|\mathbf{y}\|^2}{1 + \|\mathbf{x}\|^2} \leq 2(1 + \|\mathbf{y} - \mathbf{x}\|^2) . \quad \blacksquare$$

Proposition 11.1.22

For $\phi \in \mathcal{S}(\mathbb{R}^n)$ and $s \in [0, \infty[$, the mapping

$$\begin{aligned} T : W^{s,2}(\mathbb{R}^n) &\rightarrow W^{s,2}(\mathbb{R}^n) \\ f &\mapsto \phi f \end{aligned}$$

is a bounded linear mapping.

Proof.

To prove that there exists a constant C such that $\|\phi f\|_{s,\rho} \leq C\|f\|_{s,\rho}$ for $f \in W^{s,2}(\mathbb{R}^n)$, we first prove that the mapping $T : L^2(\mathbb{R}^n, \rho) \rightarrow L^2(\mathbb{R}^n, \rho)$ defined by

$$T(h)(\mathbf{y}) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} (1 + \|\mathbf{z}\|^2)^{-s} \hat{\phi}(\mathbf{y} - \mathbf{z}) h(\mathbf{z}) d\rho(\mathbf{z})$$

for $h \in L^2(\mathbb{R}^n, \rho)$ is a continuous linear mapping, where the measure ρ is defined by $d\rho(\mathbf{x}) = (1 + \|\mathbf{x}\|^2)^s d\mathbf{x}$. In particular, $\|h\|_{2,\rho} \equiv \left(\int_{\mathbb{R}^n} |h(\mathbf{z})|^2 d\rho(\mathbf{z}) \right)^{1/2}$ for $h \in L^2(\mathbb{R}^n, \rho)$

⁷ $u = g$ on V means that $u(\phi) = g(\phi) = \int_{\Omega} g\phi d\mathbf{x}$ for all $\phi \in \mathcal{D}(V)$. If $u \in L^1_{loc}(\Omega)$, this implies that $u = g$ almost everywhere on V .

Let

$$K(\mathbf{y}, \mathbf{z}) = (2\pi)^{-n/2} (1 + \|\mathbf{z}\|^2)^{-s} \hat{\phi}(\mathbf{y} - \mathbf{z}) .$$

According to the generalized Young's inequality, Theorem 6.5.1, we only have to prove that

$$\sup_{\mathbf{z} \in \mathbb{R}^n} \int_{\mathbb{R}^n} |K(\mathbf{y}, \mathbf{z})| d\rho(\mathbf{y}) < C \quad \text{and} \quad \sup_{\mathbf{y} \in \mathbb{R}^n} \int_{\mathbb{R}^n} |K(\mathbf{y}, \mathbf{z})| d\rho(\mathbf{z}) < C$$

for some constant C to get $\|T(h)\|_{2,\rho} \leq C\|h\|_{2,\rho}$ for $h \in L^2(\mathbb{R}^n, \rho)$. From Lemma 11.1.21, we get

$$\begin{aligned} \int_{\mathbb{R}^n} |K(\mathbf{y}, \mathbf{z})| d\rho(\mathbf{y}) &= (2\pi)^{-n/2} \int_{\mathbb{R}^n} (1 + \|\mathbf{z}\|^2)^{-s} (1 + \|\mathbf{y}\|^2)^s |\hat{\phi}(\mathbf{y} - \mathbf{z})| d\mathbf{y} \\ &\leq 2^s (2\pi)^{-n/2} \int_{\mathbb{R}^n} (1 + \|\mathbf{y} - \mathbf{z}\|^2)^s |\hat{\phi}(\mathbf{y} - \mathbf{z})| d\mathbf{y} \\ &\leq 2^s (2\pi)^{-n/2} \int_{\mathbb{R}^n} (1 + \|\mathbf{y}\|^2)^s |\hat{\phi}(\mathbf{y})| d\mathbf{y} \equiv C_1 < \infty \end{aligned}$$

for all $\mathbf{z} \in \mathbb{R}^n$ because $\hat{\phi} \in \mathcal{S}(\mathbb{R}^n)$ by Proposition 6.6.11. Moreover,

$$\int_{\mathbb{R}^n} |K(\mathbf{y}, \mathbf{z})| d\rho(\mathbf{z}) = \int_{\mathbb{R}^n} |\hat{\phi}(\mathbf{y} - \mathbf{z})| d\mathbf{z} = \int_{\mathbb{R}^n} |\hat{\phi}(\mathbf{y})| d\mathbf{y} \equiv C_2 < \infty$$

for all $\mathbf{y} \in \mathbb{R}^n$ because $\hat{\phi} \in \mathcal{S}(\mathbb{R}^n)$ by Proposition 6.6.11. It suffices to take $C = \max\{C_1, C_2\}$ to reach the conclusion.

However, We have that

$$\begin{aligned} (\phi f)^\wedge(\mathbf{y}) &= (\hat{\phi} * \hat{f})(\mathbf{y}) = \int_{\mathbb{R}^n} \hat{\phi}(\mathbf{y} - \mathbf{z}) \hat{f}(\mathbf{z}) d\mu(\mathbf{z}) \\ &= (2\pi)^{-n/2} \int_{\mathbb{R}^n} (1 + \|\mathbf{z}\|^2)^{-s} \hat{\phi}(\mathbf{y} - \mathbf{z}) \hat{f}(\mathbf{z}) d\rho(\mathbf{z}) = T(\hat{f}) , \end{aligned}$$

where item (4) of Proposition 6.6.11 was used for the first equality. Hence

$$\|\phi f\|_{s,\rho} = \|(\phi f)^\wedge\|_{2,\rho} = \|T(\hat{f})\|_{2,\rho} \leq C\|\hat{f}\|_{2,\rho} = \|f\|_{s,\rho}$$

for all $f \in W^{s,2}(\mathbb{R}^n)$. ■

The reader should be careful when reading the next proposition. For instance, if $u \in \mathcal{D}'(\Omega)$ is represented by a function in $W^{s,2}(\Omega)$, then we also denote this function as u .

Proposition 11.1.23

Let Ω be an open subset of $\mathbb{R}^n$ and k be a non-negative integer. Suppose that $u \in \mathcal{D}'(\Omega)$. Then $u \in W_{loc}^{k,p}(\Omega)$ if and only if $\phi u \in W^{k,p}(\mathbb{R}^n)$ ⁸ for all $\phi \in \mathcal{D}(\Omega)$. Similarly, if $s \in \mathbb{R}$, then $u \in W_{loc}^{s,2}(\Omega)$ if and only if $\phi u \in W^{s,2}(\mathbb{R}^n)$ ⁸ for all $\phi \in \mathcal{D}(\Omega)$.

Remark 11.1.24

As usual, $\phi \in \mathcal{D}(\Omega)$ can be expanded to $\phi \in \mathcal{D}(\mathbb{R}^n)$ with $\text{supp } \phi \subset \Omega$ by setting $\phi(\mathbf{x}) = 0$ for $\mathbf{x} \in \mathbb{R}^n \setminus \Omega$. ♠

⁸Recall that $\phi u \in \mathcal{D}'(\mathbb{R}^n)$ is defined by $(\phi u)(\psi) = u(\phi\psi)$ for all $\psi \in \mathcal{D}(\mathbb{R}^n)$ because $\text{supp}(\phi\psi) \subset \text{supp } \phi \subset \Omega$ for all $\psi \in \mathcal{D}(\mathbb{R}^n)$ and so we may assume that $\psi\psi \in \mathcal{D}(\Omega)$. Hence $\phi u \in W^{k,p}(\mathbb{R}^n)$ (resp. $\phi u \in W^{s,2}(\mathbb{R}^n)$) means that there exists $g \in W^{k,p}(\mathbb{R}^n)$ (resp. $g \in W^{s,2}(\mathbb{R}^n)$) such that $\phi u = g$ in the sense of distributions on $\mathbb{R}^n$.

Proof.

i) Suppose that $u \in W_{loc}^{s,2}(\Omega)$ and $\phi \in \mathcal{D}(\Omega)$. Choose an open set V such that $\text{supp } \phi \subset V \subset \overline{V} \subset \Omega$. We get from the definition of $W_{loc}^{s,2}(\Omega)$ that there exists $g \in W^{s,2}(\mathbb{R}^n)$ such that $u = g$ on $V \supset \text{supp } \phi$ in the sense of distributions. Hence $\phi u = \phi g$ in the sense of distribution on $\mathbb{R}^n$ because

$$(\phi u)(\psi) = u(\phi\psi) = g(\phi\psi) = (\phi g)(\psi)$$

for all $\psi \in \mathcal{D}'(\mathbb{R}^n)$ since $\text{supp}(\phi\psi) \subset \text{supp } \phi \subset V$. Therefore, $\phi u = \phi g \in W^{s,2}(\mathbb{R}^n)$ by Proposition 11.1.22.

Suppose that $\phi u \in W^{s,2}(\mathbb{R}^n)$ for all $\phi \in \mathcal{D}(\Omega)$. Let V be an open subset of Ω such that $\overline{V} \subset \Omega$. Choose $\phi \in \mathcal{D}(\Omega)$ such that $\phi(\mathbf{x}) = 1$ for all $\mathbf{x} \in V$, then $u = \phi u$ on V in the sense of distributions with $\phi u \in W^{s,2}(\mathbb{R}^n)$.

ii) The proof in the case of $W^{k,p}(\mathbb{R}^n)$ is basically identical. ■

It follows from Propositions 11.1.22 and 11.1.23 that $W^{s,2}(\mathbb{R}^n) \subset W_{loc}^{s,2}(\Omega)$ for every open set $\Omega \subset \mathbb{R}^n$.

11.2 The $W_0^{k,p}$ spaces

Definition 11.2.1

Let Ω be an open subset of $\mathbb{R}^n$. For $k \in \mathbb{N}$ and p a positive integer, the **Sobolev space** $W_0^{k,p}(\Omega)$ is the closure in $W^{k,p}(\Omega)$ of $\mathcal{D}(\Omega)$ with respect to the norm $\|\cdot\|_{k,p}$. For $s \in \mathbb{R}$, the **Sobolev space** $W_0^{s,2}(\Omega)$ is the closure in $W^{s,2}(\Omega)$ of $\mathcal{D}(\Omega)$ with respect to the norm $\|\cdot\|_{s,2}$.

If $\Omega \subset \mathbb{R}^n$ is a bounded open set, we will show when studying the trace operators in Section 11.8 that the space $W_0^{k,p}(\Omega)$ generalizes the space of functions $f \in C^k(\Omega) \cap C^{k-1}(\overline{\Omega})$ such that $\frac{\partial^j f}{\partial \nu^j}(\mathbf{x}) = 0$ for all $\mathbf{x} \in \partial\Omega$ and $0 \leq j < k$, where $\nu(\mathbf{x})$ is the outward unit normal to $\partial\Omega$ at $\mathbf{x} \in \partial\Omega$. See Theorem 11.8.12.

The next proposition generalizes Lemma 11.1.7.

Proposition 11.2.2

Consider $k \in \mathbb{N}$ and p a positive integer. Let Ω be an open subset of $\mathbb{R}^n$. Given a function $f : \Omega \rightarrow \mathbb{C}$, we define a function $\underline{f} : \mathbb{R}^n \rightarrow \mathbb{C}$ by

$$\underline{f}(\mathbf{x}) = \begin{cases} f(\mathbf{x}) & \text{if } \mathbf{x} \in \Omega \\ 0 & \text{if } \mathbf{x} \in \mathbb{R}^n \setminus \Omega \end{cases}$$

If $f \in W_0^{k,p}(\Omega)$, then $D^\alpha \underline{f} = \underline{D^\alpha f}$ in the sense of distributions on $\mathbb{R}^n$ for $|\alpha| \leq k$. In particular, $\underline{f} \in W^{k,p}(\mathbb{R}^n)$.

Proof.

By definition, there exists a sequence $\{\phi_i\}_{i=0}^\infty$ in $\mathcal{D}(\Omega)$ converging to f with respect to the norm $\|\cdot\|_{k,p}$.

i) We first prove that

$$\int_{\Omega} D^{\alpha} \phi_i(\mathbf{x}) \psi(\mathbf{x}) d\mathbf{x} \rightarrow \int_{\Omega} D^{\alpha} f(\mathbf{x}) \psi(\mathbf{x}) d\mathbf{x} \quad \text{as } i \rightarrow \infty$$

for any bounded function ψ with bounded support and $|\alpha| \leq k$. According to Hölders' inequality,

$$\begin{aligned} \left| \int_{\Omega} D^{\alpha} \phi_i(\mathbf{x}) \psi(\mathbf{x}) d\mathbf{x} - \int_{\Omega} D^{\alpha} f(\mathbf{x}) \psi(\mathbf{x}) d\mathbf{x} \right| &\leq \int_{\Omega} |D^{\alpha} \phi_i(\mathbf{x}) - D^{\alpha} f(\mathbf{x})| |\psi(\mathbf{x})| d\mathbf{x} \\ &\leq \left(\int_{\Omega} |D^{\alpha} \phi_i(\mathbf{x}) - D^{\alpha} f(\mathbf{x})|^p d\mathbf{x} \right)^{1/p} \left(\int_{\Omega} |\psi(\mathbf{x})|^q d\mathbf{x} \right)^{1/q} \\ &= C \left(\int_{\Omega} |D^{\alpha} \phi_i(\mathbf{x}) - D^{\alpha} f(\mathbf{x})|^p d\mathbf{x} \right)^{1/p} \leq C \|\phi_i - f\|_{k,p} \rightarrow 0 \quad \text{as } i \rightarrow \infty, \end{aligned}$$

where $\frac{1}{p} + \frac{1}{q} = 1$ and $C = \left(\int_{\Omega} |\psi(\mathbf{x})|^q d\mathbf{x} \right)^{1/q} < \infty$.

ii) Since

$$\int_{\mathbb{R}^n} |f(\mathbf{x})|^p d\mathbf{x} = \int_{\Omega} |f(\mathbf{x})|^p d\mathbf{x} < \infty,$$

the function $\underline{f}$ defines a distribution on $\mathbb{R}^n$. Hence, for $|\alpha| \leq k$, we get from (i) that

$$\begin{aligned} (D^{\alpha} \underline{f})(\phi) &= (-1)^{|\alpha|} \int_{\mathbb{R}^n} \underline{f}(\mathbf{x}) D^{\alpha} \phi(\mathbf{x}) d\mathbf{x} = (-1)^{|\alpha|} \int_{\Omega} f(\mathbf{x}) D^{\alpha} \phi(\mathbf{x}) d\mathbf{x} \\ &= (-1)^{|\alpha|} \lim_{i \rightarrow \infty} \int_{\Omega} \phi_i(\mathbf{x}) D^{\alpha} \phi(\mathbf{x}) d\mathbf{x} = \lim_{i \rightarrow \infty} \int_{\Omega} D^{\alpha} \phi_i(\mathbf{x}) \phi(\mathbf{x}) d\mathbf{x} \\ &= \int_{\Omega} D^{\alpha} f(\mathbf{x}) \phi(\mathbf{x}) d\mathbf{x} = \int_{\mathbb{R}^n} \underline{D^{\alpha} f}(\mathbf{x}) \phi(\mathbf{x}) d\mathbf{x} = (\underline{D^{\alpha} f})(\phi) \end{aligned}$$

for $\phi \in \mathcal{D}(\mathbb{R}^n)$. Thus $D^{\alpha} \underline{f} = \underline{D^{\alpha} f}$ almost everywhere in $\mathbb{R}^n$.

iii) We have that $D^{\alpha} \underline{f} = \underline{D^{\alpha} f} \in L^p(\mathbb{R}^n)$ for $|\alpha| \leq k$ because

$$\left(\int_{\mathbb{R}^n} |\underline{D^{\alpha} f}(\mathbf{x})|^p d\mathbf{x} \right)^{1/p} = \left(\int_{\Omega} |D^{\alpha} f(\mathbf{x})|^p d\mathbf{x} \right)^{1/p} < \infty$$

since $f \in W^{k,p}(\Omega)$. Thus, $\underline{f} \in W^{k,p}(\mathbb{R}^n)$.

It is also interesting to note that, as a distribution, $D^{\alpha} \underline{f}|_{\Omega} = D^{\alpha} f$. ■

Lemma 11.2.3

Let Ω be an open subset of $\mathbb{R}^n$. If $f \in W^{k,p}(\Omega)$ has a compact support in Ω , then there exists a sequence $\{\psi_i\}_{i=1}^\infty \subset \mathcal{D}(\Omega)$ such that $\psi_i \rightarrow f$ in $W^{k,p}(\Omega)$ as $i \rightarrow \infty$; namely,

$$f \in W_0^{k,p}(\Omega).$$

Proof.

Since $K = \text{supp } f$ is a compact subset of the open set Ω , there exists an open set U such that $K \subset U \subset \overline{U} \subset \Omega$. Hence, $d = \text{dist}(K, \partial U) > 0$.

Choose $\phi \in \mathcal{D}(\mathbb{R}^n)$ such that $\phi(\mathbf{x}) = 0$ for $\|\mathbf{x}\| > 1$ and $\int_{\mathbb{R}^n} \phi(\mathbf{x}) d\mathbf{x} = 1$. Let $\phi_i(\mathbf{x}) = i^n \phi(i\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^n$. It follows from Lemmas 11.1.7 and 11.1.14 that $\underline{f} * \phi_i \in \mathcal{D}(\mathbb{R}^n)$ for all $i > 0$ and $\underline{f} * \phi_i \rightarrow \underline{f}$ in $W^{k,p}(\mathbb{R}^n)$ as $i \rightarrow \infty$.

Since $1/i \rightarrow 0$ as $i \rightarrow \infty$, there exists $J > 0$ such that $1/i < d$ for $i > J$. Let $\psi_i = \underline{f} * \phi_{i+J}$ for $i > 0$. We have that

$$\text{supp } \psi_i \subset \text{supp } f + \text{supp } \phi_{i+J} \subset K + \overline{B_{1/(i+J)}(\mathbf{0})} \subset \overline{U} \subset \Omega$$

for all $i > 0$ and $\|\psi_i - \underline{f}\|_{k,p,\Omega} = \|\psi_i - \underline{f}\|_{k,p,\mathbb{R}^n} \rightarrow 0$ as $i \rightarrow \infty$. ■

Remark 11.2.4

To illustrate the fact that proofs in mathematics are often not unique. We give another proof of the previous lemma.

We expand f to $\underline{f} : \mathbb{R}^n \rightarrow \mathbb{C}$ by

$$\underline{f}(\mathbf{x}) = \begin{cases} f(\mathbf{x}) & \text{if } \mathbf{x} \in \Omega \\ 0 & \text{if } \mathbf{x} \in \mathbb{R}^n \setminus \Omega \end{cases}$$

It follows from Proposition 11.2.2 that $\underline{f} \in W^{k,p}(\mathbb{R}^n)$. Since $\mathcal{D}(\mathbb{R}^n)$ is dense in $W^{k,p}(\mathbb{R}^n)$, there exists a sequence $\{\phi_i\}_{i=0}^\infty$ in $\mathcal{D}(\mathbb{R}^n)$ converging to $\underline{f}$ with respect to the norm $\|\cdot\|_{k,p,\mathbb{R}^n}$.

Choose $\phi \in \mathcal{D}(\Omega)$ such that $\phi(\mathbf{x}) = 1$ for all $\mathbf{x} \in \text{supp } f$ ⁹. It follows from Proposition 11.1.17 that $\{\phi \phi_i\}_{i=0}^\infty$ converges to $\phi \underline{f}$ with respect to the norm $\|\cdot\|_{k,p,\mathbb{R}^n}$. Thus, $\{\phi \phi_i\}_{i=0}^\infty$ is a sequence in $\mathcal{D}(\Omega)$ converging to $\phi \underline{f}|_\Omega = f$ with respect to the norm $\|\cdot\|_{k,p,\Omega}$. Therefore $f \in W_0^{k,p}(\Omega)$ by definition. ♠

There is a converse to Proposition 11.2.2.

Proposition 11.2.5

Consider $k \in \mathbb{N}$ and p a positive integer. Let Ω be an open subset of $\mathbb{R}^n$ that satisfies the segment property. Given a function $f : \Omega \rightarrow \mathbb{C}$, we define a function $\underline{f} : \mathbb{R}^n \rightarrow \mathbb{C}$ by

$$\underline{f}(\mathbf{x}) = \begin{cases} f(\mathbf{x}) & \text{for } \mathbf{x} \in \Omega \\ 0 & \text{for } \mathbf{x} \notin \Omega \end{cases}$$

Then, $f \in W_0^{k,p}(\Omega)$ if and only if $\underline{f} \in W^{k,p}(\mathbb{R}^n)$.

⁹See footnote on page 402.

Proof.

We have proved in Proposition 11.2.2 that $\underline{f} \in W^{k,p}(\mathbb{R}^n)$ if $f \in W_0^{k,p}(\Omega)$. We now prove the converse.

i) We first assume that $K = \text{supp } f \subset \overline{\Omega}$ is compact. Let $Q = K \setminus \bigcup_{\mathbf{x} \in \partial\Omega} U_{\mathbf{x}}$, where each bounded open set $U_{\mathbf{x}}$ comes from the definition of segment property. We have that $Q \subset \Omega$ because $\partial\Omega \subset \bigcup_{\mathbf{x} \in \partial\Omega} U_{\mathbf{x}}$. Hence, Q is a compact subset of Ω because it is bounded and closed.

Choose an open set U such that $Q \subset U \subset \overline{U} \subset \Omega$. We get an open cover $\{U\} \cup \left\{ \bigcup_{\mathbf{x} \in \partial\Omega} U_{\mathbf{x}} \right\}$ of K . So, there exists a finite subcover of K that we denote $\{U_i\}_{i=1}^I$. At most one of the U_i may have an empty intersection with $\partial\Omega$; namely, the open set U introduced before.

Let $\{\phi_i\}_{i=1}^I$ be a partition of unity of K subordinate to $\{U_i\}_{i=1}^I$

Suppose that $\underline{f} \in W^{k,p}(\mathbb{R}^n)$. Let $h_i \equiv \phi_i \underline{f}$. Then $\underline{h}_i = \phi_i \underline{f} \in W^{k,p}(\mathbb{R}^n)$ according to Proposition 11.1.17. We can write $\underline{f}$ as $\underline{f} = \sum_{i=1}^I \phi_i \underline{f} = \sum_{i=1}^I \underline{h}_i$ where either $\text{supp } \underline{h}_i \subset U_i \subset \Omega$ (if $U_i = U$ mentioned above) or $\text{supp } \underline{h}_i \subset U_i \cap \overline{\Omega}$. If we can prove that $h_i \in W_0^{k,p}(\Omega)$ for all i , then we will have that $f = \sum_{i=1}^I h_i \in W_0^{k,p}(\Omega)$.

If $\text{supp } h_i \subset \Omega$, we have that $h_i \in W_0^{k,p}(\Omega)$ according to Lemma 11.2.3.

if $\text{supp } \underline{h}_i \subset U_i \cap \overline{\Omega}$, we consider $h_{i,t} \in W^{k,p}(\mathbb{R}^n)$ for $0 < t < 1$ defined by $h_{i,t}(\mathbf{x}) = \underline{h}_i(\mathbf{x} - t\mathbf{y}_i)$. We have that $\text{supp } h_{i,t} \subset \Omega$ for $0 < t < 1$ because $\text{supp } h_{i,t} \subset t\mathbf{y}_i + U_i \cap \overline{\Omega} \subset \Omega$ for $0 < t < 1$. It follows from Lemma 11.2.3 that $h_{i,t} \in W_0^{k,p}(\Omega)$ for all t .

Since $h_{i,t} \rightarrow h_i$ in $W_0^{k,p}(\Omega)$ as $t \rightarrow 0^+$ according to Lemma 11.1.10, and since $W_0^{k,p}(\Omega)$ is closed, we get that $h_i \in W_0^{k,p}(\Omega)$.

ii) If $\text{supp } f$ is not bounded, we construct a sequence of functions $\{f_i\}_{i=1}^\infty \subset W^{k,p}(\Omega)$ as we did in (i) of the proof of Theorem 11.1.11¹⁰ such that $\text{supp } f_i$ is bounded and $f_i \rightarrow f$ in $W^{k,p}(\Omega)$ as $i \rightarrow \infty$.

From (i), we have that $f_i \in W_0^{k,p}(\Omega)$. Since $f_i \rightarrow f$ in $W^{k,p}(\Omega)$ as $i \rightarrow \infty$, and since $W_0^{k,p}(\Omega)$ is closed, we get that $f \in W_0^{k,p}(\Omega)$. ■

Example 11.2.6

As the following example shows, $W_0^{k,2}(\Omega) \subsetneq W^{k,2}(\Omega)$ when Ω is a bounded open subset of $\mathbb{R}^n$ and $k \geq 1$.

If Ω is bounded, we obviously have that $f : \Omega \rightarrow \mathbb{R}$ defined by $f(\mathbf{x}) = 1$ for all $\mathbf{x} \in \Omega$ is in $W^{k,2}(\Omega)$. We now show that $f \notin W_0^{k,2}(\Omega)$. Since $W_0^{k,2}(\Omega) \subset W_0^{1,1}(\Omega)$ because Ω is bounded, it is in fact enough to show that $f \notin W_0^{1,1}(\Omega)$.

Suppose that $f \in W_0^{1,1}(\Omega)$. Then, there exists a sequence $\{\phi_i\}_{i=1}^\infty \subset \mathcal{D}(\Omega)$ such that $\phi_i \rightarrow f$ with respect to the norm $\|\cdot\|_{1,1}$. Since $\int_\Omega |\phi_i| d\mathbf{x} < \infty$ and $\sum_{j=1}^n \int_\Omega \left| \frac{\partial \phi_i}{\partial x_j} \right| d\mathbf{x} < \infty$ because

¹⁰Take $\delta = 1/i$ in (i) of the proof of Theorem 11.1.11.

$\phi_i \in \mathcal{D}(\Omega)$, there exists a positive constant C_i such that $\int_{\Omega} |\phi_i| d\mathbf{x} < C_i \sum_{j=1}^n \int_{\Omega} \left| \frac{\partial \phi_i}{\partial x_j} \right| d\mathbf{x}$. We may assume that $C_i > 1$. Hence,

$$\begin{aligned} \|f - \phi_i\|_{1,1} &= \int_{\Omega} |f - \phi_i| d\mathbf{x} + \sum_{j=1}^n \int_{\Omega} \left| \frac{\partial \phi_i}{\partial x_j} \right| d\mathbf{x} \geq \int_{\Omega} |f| d\mathbf{x} - \int_{\Omega} |\phi_i| d\mathbf{x} + \sum_{j=1}^n \int_{\Omega} \left| \frac{\partial \phi_i}{\partial x_j} \right| d\mathbf{x} \\ &\geq \int_{\Omega} d\mathbf{x} + (1 - C_i) \sum_{j=1}^n \int_{\Omega} \left| \frac{\partial \phi_i}{\partial x_j} \right| d\mathbf{x} \geq \int_{\Omega} d\mathbf{x} + (1 - C_i) \|f - \phi_i\|_{1,1} \end{aligned}$$

for all i . Thus,

$$\|f - \phi_i\|_{1,1} \geq \frac{1}{C_i} \int_{\Omega} d\mathbf{x} \geq \int_{\Omega} d\mathbf{x} > 0$$

for all i . This is a contradiction that $\phi_i \rightarrow f$ with respect to the norm $\|\cdot\|_{1,1}$.

The previous reasoning can be easily modified to prove that $W_0^{k,p}(\Omega) \subsetneq W^{k,p}(\Omega)$ when Ω is a bounded open subset of $\mathbb{R}^n$ and $p, k \geq 1$. $\clubsuit$

11.3 The $H^{k,p}$ spaces

We introduce in this section the spaces $H^{k,p}(\Omega)$ and $H_0^{k,p}(\Omega)$ for $k \in \mathbb{N}$ and $1 \leq p \leq \infty$, and the spaces $H^{s,2}(\Omega)$ and $H_0^{s,2}(\Omega)$ for $s \in \mathbb{R}$. We also prove that, except for $k = \infty$, the spaces $H^{k,p}(\Omega)$ and $H_0^{k,p}(\Omega)$ are none other than the spaces $W^{k,p}(\Omega)$ and $W_0^{k,p}(\Omega)$ respectively.

Definition 11.3.1

Assume that $k \in \mathbb{N}$ and p is a positive integer or ∞ . Let Ω be an open set in $\mathbb{R}^n$ and

$$S = \left\{ f \in C^k(\Omega) : \|f\|_{k,p} < \infty \right\},$$

where $\|\cdot\|_{k,p}$ is the norm defined in (11.1.1).

The **Sobolev space** $H^{k,p}(\Omega)$ is the closure of S in $L^p(\Omega)$ with respect to the norm $\|\cdot\|_{k,p}$.

The **Sobolev space** $H_0^{k,p}(\Omega)$ is the closure of $\mathcal{D}(\Omega)$ in $L^p(\Omega)$ with respect to the norm $\|\cdot\|_{k,p}$.

Since the case $p = 2$ is predominant in the study of partial differential equations, we define $H^k(\Omega) \equiv H^{k,2}(\Omega)$ and $H_0^k(\Omega) \equiv H_0^{k,2}(\Omega)$.

Remark 11.3.2

1. Let Ω be a bounded open subset of $\mathbb{R}^n$. An equivalent definition of $H^{k,p}(\Omega)$ for p a positive integer is that $H^{k,p}(\Omega)$ is the closure of $C^\infty(\overline{\Omega})$ in $L^p(\Omega)$ with respect to the norm $\|\cdot\|_{k,p}$ defined in (11.1.1).

Let H be the closure of $C^\infty(\overline{\Omega})$ in $L^p(\Omega)$ with respect to the norm $\|\cdot\|_{k,p}$. If $f \in C^\infty(\overline{\Omega})$, then $D^\alpha f$ is a continuous function on the compact set $\overline{\Omega}$ for all multi-indices α . It

follows that $\|D^\alpha f|_\Omega\|_p \leq \|D^\alpha f|_{\overline{\Omega}}\|_p < \infty$ for all multi-indices α . Thus, $\|f|_\Omega\|_{k,p} < \infty$ for all $f \in C^\infty(\overline{\Omega})$. Hence, $C^\infty(\overline{\Omega}) \subset S$, where S is given in Definition 11.3.1.

However, it follows from Theorem 11.1.11 that $C^\infty(\overline{\Omega})$ is dense in S with respect to the norm $\|\cdot\|_{k,p}$ ¹¹.

Thus, the closure of $C^\infty(\overline{\Omega})$ in $L^p(\Omega)$ with respect to $\|\cdot\|_{k,p}$ is the closure of S in $L^p(\Omega)$ with respect to $\|\cdot\|_{k,p}$.

2. As well, an equivalent definition of $H_0^{k,p}(\Omega)$ is that $H_0^{k,p}(\Omega)$ is the closure of $C_c^k(\Omega)$ in $L^p(\Omega)$ with respect to the norm $\|\cdot\|_{k,p}$ defined in (11.1.1).

Given $f \in C_c^k(\Omega)$, it follows from Lemma 11.2.3 that there exists a sequence $\{\psi_i\}_{i=1}^\infty \subset \mathcal{D}(\Omega)$ such that $\psi_i \rightarrow f$ in $W^{k,p}(\Omega)$ as $i \rightarrow \infty$. Thus, $C_c^k(\Omega) \subset \overline{\mathcal{D}(\Omega)}$, where the closure is with respect to the norm $\|\cdot\|_{k,p}$. It follows that the closure of $C_c^k(\Omega)$ in $L^p(\Omega)$ with respect to $\|\cdot\|_{k,p}$ is included in $H_0^{k,p}(\Omega)$, the closure of $\mathcal{D}(\Omega)$ in $L^p(\Omega)$ with respect to $\|\cdot\|_{k,p}$.

But since $\mathcal{D}(\Omega) \subset C_c^k(\Omega)$, we also have that $H_0^{k,p}(\Omega)$, the closure of $\mathcal{D}(\Omega)$ in $L^p(\Omega)$ with respect to $\|\cdot\|_{k,p}$, is included in the closure of $C_c^k(\Omega)$ in $L^p(\Omega)$ with respect to $\|\cdot\|_{k,p}$. Thus, the closure of $C_c^k(\Omega)$ in $L^p(\Omega)$ with respect to $\|\cdot\|_{k,p}$ is $H_0^{k,p}(\Omega)$.

♠

Proposition 11.3.3

Let Ω be an open subset of $\mathbb{R}^n$. If $k \in \mathbb{N}$ and p is a positive integer or ∞ , then $H^{k,p}(\Omega) \subset W^{k,p}(\Omega)$ and $H_0^{k,p}(\Omega) = W_0^{k,p}(\Omega)$.

Proof.

That $H_0^{k,p}(\Omega) = W_0^{k,p}(\Omega)$ is simple to prove because the closure of $\mathcal{D}(\Omega)$ in $L^p(\Omega)$ or $W^{k,p}(\Omega) \subset L^p(\Omega)$ with respect to the norm $\|\cdot\|_{k,p}$ give the same set since $W^{k,p}(\Omega)$ is closed with respect to this norm. Recall that $W^{k,p}(\Omega)$ is a Banach space with respect to the norm $\|\cdot\|_{k,p}$, so $W^{k,p}(\Omega)$ is complete and thus closed with respect to this norm.

To prove that $H^{k,p}(\Omega) \subset W^{k,p}(\Omega)$, we first note that S in the definition of $H^{k,p}(\Omega)$ is a subset of $W^{k,p}(\Omega)$. Again, Since $W^{k,p}(\Omega)$ is closed, we have that $H^{k,p}(\Omega) = \overline{S} \subset W^{k,p}(\Omega)$, where the closure is with respect to the norm $\|\cdot\|_{k,p}$. ■

The first question that comes to mind is whether $W^{k,p}(\Omega) = H^{k,p}(\Omega)$ whatever the open set $\Omega \subset \mathbb{R}^n$. The answer is positive for $1 \leq p < \infty$. However, it is not always true for $p = \infty$.

Theorem 11.3.4

If $\Omega \subset \mathbb{R}^n$ is an open set, $k \in \mathbb{N}$ and p is a positive integer, then $H^{k,p}(\Omega) = W^{k,p}(\Omega)$.

¹¹In fact $C^\infty(\overline{\Omega})$ is dense in $W^{k,p}(\Omega) \supset S$ with respect to $\|\cdot\|_{k,p}$.

Proof.

$H^{k,p}(\Omega) \subset W^{k,p}(\Omega)$ follows from Proposition 11.3.3. The converse inclusion comes from Theorem 11.1.19 because S in the definition of $H^{k,p}(\Omega)$ contains $C^\infty(\Omega) \cap W^{k,p}(\Omega)$. ■

Example 11.3.5

We have that $H^{1,\infty}(]-1, 1[) \subsetneq W^{1,\infty}(]-1, 1[)$.

The function $f(x) = |x|$ is in $W^{1,\infty}(]-1, 1[)$. In particular,

$$Df(x) = \begin{cases} -1 & \text{for } x < 0 \\ c & \text{for } x = 0 \\ 1 & \text{for } x > 0 \end{cases}$$

where c can be any constant of your choice. However, it is impossible to find

$$\phi \in \{g \in C^1(]-1, 1[) : \|g\|_{1,\infty} < \infty\}$$

such that

$$\|f - \phi\|_{1,\infty} = \sup_{x \in]-1, 1[} |f(x) - \phi(x)| + \sup_{x \in]-1, 1[} |Df(x) - D\phi(x)| < 1.$$

If $D\phi(0) = b \in \mathbb{R}$, then either $\lim_{x \rightarrow 0^-} |Df(x) - D\phi(x)| = |-1 - b| \geq 1$ or $\lim_{x \rightarrow 0^+} |Df(x) - D\phi(x)| = |1 - b| \geq 1$. Recall also that the uniform limit of continuous functions on $] -1, 1[$ is a continuous function. Thus, if $\{\phi_j\}_{j=0}^\infty$ was to converge in $W^{1,\infty}(]-1, 1[)$ to f , then $\{D\phi_j\}_{j=1}^\infty$ would be converging uniformly to Df in $] -1, 1[$. It would follow that the limit Df will be continuous at the origin. This is not so for Df . ♣

Definition 11.3.6

To obtain a consistent notation (with many textbook on Sobolev spaces), we set $H^{s,2}(\Omega) = W^{s,2}(\Omega)$ for $s \in \mathbb{R}$. Moreover, the **Sobolev space** $H_0^{s,2}(\Omega)$ is the completion of $\mathcal{D}(\Omega)$ in $H^{s,2}(\Omega)$.

Because of Theorem 11.1.3, the definition of $H_0^{k,2}(\Omega)$ for $k \in \mathbb{N}$ above is equivalent to the definition of $H_0^{k,2}(\Omega)$ given in Definition 11.3.1.

11.4 Sobolev Spaces on Manifolds

We can define Sobolev spaces on subset of $\mathbb{R}^n$ other than open sets.

Definition 11.4.1

Let $M \subset \mathbb{R}^n$ be a C^k -manifold of dimension m (without a boundary). Let $\mathcal{B} = \{(U_j, \psi_j)\}_{j=1}^\infty$ be an atlas for M as defined in Section 9.9.3. Let $\{\phi_i\}_{i=1}^\infty$ be a partition of unity for M subordinate to $\mathcal{B}$. The space $W^{k,2}(M)$ is defined by

$$W^{k,2}(M) = \{f \in L^2(M) : (\phi_i f) \circ \psi_i^{-1} \in W_{loc}^{k,2}(\psi_i(U_i)), 0 < i < \infty\}.$$

For this definition to make sense, it has to be independent of the atlas used for the manifold M . This is a consequence of the following proposition.

Proposition 11.4.2

Suppose that Ω_1 and Ω_2 are two bounded open subsets of $\mathbb{R}^n$, and that χ is a C^k -diffeomorphism of Ω_1 onto Ω_2 ¹².

1. Then, $T : W_{loc}^{s,2}(\Omega_1) \rightarrow W_{loc}^{s,2}(\Omega_2)$ defined by $T(u) = u \circ \chi^{-1}$ for $u \in W_{loc}^{s,2}(\Omega_1)$ is an isomorphism with inverse defined by $T^{-1}(v) = v \circ \chi$ for all $v \in W_{loc}^{s,2}(\Omega_2)$.
2. If in addition all the partial derivatives of χ and χ^{-1} of order less or equal to k are bounded, then $T : W^{k,p}(\Omega_1) \rightarrow W^{k,p}(\Omega_2)$ defined by $T(u) = u \circ \chi^{-1}$ for $u \in W^{k,p}(\Omega_1)$ is an isomorphism with inverse defined by $T^{-1}(v) = v \circ \chi$ for all $v \in W^{k,p}(\Omega_2)$. This result is also true if $W^{k,p}(\Omega_1)$ and $v \in W^{k,p}(\Omega_2)$ are respectively replaced by $v \in W_0^{k,p}(\Omega_1)$ and $v \in W_0^{k,p}(\Omega_2)$.

Proof.

A proof of the first item is given in Proposition 25.8 of [39].

A proof of the second item is given in Propositions 24.6 and 24.7 of [39], and in Theorem 3.41 of [2]. This proof uses the result of Theorem 11.1.19 and the classical formula for the change of variables in integrals. ■

The concept of Sobolev spaces on manifolds is really important for some problems in applied mathematics, particularly in physics. However, we will make very little use of Sobolev spaces on manifolds in this book. For more details about Sobolev spaces on Manifold, the reader should consult Chapter 25 of [39].

11.5 Dual Spaces

Definition 11.5.1

Let Ω be an open subset of $\mathbb{R}^n$. The **dual** of the space $H_0^{s,2}(\Omega)$ is the space $H^{-s,2}(\Omega)$ of all bounded linear functionals on $H_0^{s,2}(\Omega)$.

Remark 11.5.2

If $\Omega = \mathbb{R}^n$ or Ω is a compact manifold (without boundary), we then have that $H^{k,2}(\Omega) = H_0^{k,2}(\Omega)$ and $H^{-k,2}(\Omega)$ is the dual space of $H^{k,2}(\Omega)$; namely, $H^{-k,2}(\Omega)$ is the set of all bounded linear functionals on $H^{k,2}(\Omega)$ ♠

It may seem from the definition above that we have two definitions of $H^{s,2}(\mathbb{R}^n)$ for $s < 0$: the definition above and Definition 11.1.4. We now show that it is not so and, in this process, show that the dual space of $H^{s,2}(\mathbb{R}^n)$ for $s \geq 0$ is the space $H^{-s,2}(\mathbb{R}^n)$.

¹²Namely, $\chi : \Omega_1 \rightarrow \Omega_2$ is an homeomorphism, and $\chi : \Omega_1 \rightarrow \Omega_2$ and $\chi^{-1} : \Omega_2 \rightarrow \Omega_1$ are of class C^k .

Proposition 11.5.3

For $s \in \mathbb{R}$, the dual of $H^{s,2}(\mathbb{R}^n)$ is canonically isomorphic to $H^{-s,2}(\mathbb{R}^n)$.

Proof.

As we did in the proof of Theorem 11.1.16, we consider the operator

$$\begin{aligned} T^{[s]} : H^{s,2}(\mathbb{R}^n) &\rightarrow L^2(\mathbb{R}^n) \\ f &\mapsto \mathcal{F}^{-1}\left((1 + \|\mathbf{y}\|_2^2)^{s/2} \mathcal{F}(f)\right) \end{aligned}$$

We saw that it is an isometry. We also consider the injection

$$\begin{aligned} i : \mathcal{D}(\mathbb{R}^n) &\rightarrow H^{s,2}(\mathbb{R}^n) \\ \phi &\mapsto \phi \end{aligned}$$

It is easy to see that i is a continuous mapping¹³. It is also easy to verify that the adjoint of i is the mapping

$$\begin{aligned} i^* : (H^{s,2}(\mathbb{R}^n))^* &\rightarrow \mathcal{D}'(\mathbb{R}^n) \\ f &\mapsto f \end{aligned}$$

The transpose of the sequence

$$\mathcal{D}(\mathbb{R}^n) \xrightarrow{i} H^{s,2}(\mathbb{R}^n) \xrightarrow{T^{[s]}} L^2(\mathbb{R}^n)$$

is the sequence

$$L^2(\mathbb{R}^n) \xrightarrow{(T^{[s]})^*} (H^{s,2}(\mathbb{R}^n))^* \xrightarrow{i^*} \mathcal{D}'(\mathbb{R}^n),$$

because the dual of $L^2(\mathbb{R}^n)$ is $L^2(\mathbb{R}^n)$ and $\mathcal{D}'(\mathbb{R}^n)$ is the dual of $\mathcal{D}(\mathbb{R}^n)$ by definition; namely, $\mathcal{D}'(\mathbb{R}^n)$ is the space of continuous linear functionals on $\mathcal{D}(\mathbb{R}^n)$.

For $\phi \in \mathcal{S}(\mathbb{R}^n) \subset H^{s,2}(\mathbb{R}^n)$ and $\psi \in \mathcal{S}(\mathbb{R}^n) \subset (L^2(\mathbb{R}^n))^* = L^2(\mathbb{R}^n)$, we have that

$$\begin{aligned} \psi(T^{[s]}(\phi)) &= \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} e^{i\mathbf{z} \cdot \mathbf{y}} \left((1 + \|\mathbf{y}\|_2^2)^{s/2} \int_{\mathbb{R}^n} e^{-i\mathbf{y} \cdot \mathbf{x}} \phi(\mathbf{x}) d\mu(\mathbf{x}) \right) d\mu(\mathbf{y}) \right) \psi(\mathbf{z}) d\mathbf{z} \\ &= \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} e^{-i\mathbf{x} \cdot \mathbf{y}} \left((1 + \|\mathbf{y}\|_2^2)^{s/2} \int_{\mathbb{R}^n} e^{i\mathbf{y} \cdot \mathbf{z}} \psi(\mathbf{z}) d\mu(\mathbf{z}) \right) d\mu(\mathbf{y}) \right) \phi(\mathbf{x}) d\mathbf{x} \\ &= \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} e^{i\mathbf{x} \cdot \mathbf{y}} \left((1 + \|\mathbf{y}\|_2^2)^{s/2} \int_{\mathbb{R}^n} e^{-i\mathbf{y} \cdot \mathbf{z}} \psi(\mathbf{z}) d\mu(\mathbf{z}) \right) d\mu(\mathbf{y}) \right) \phi(\mathbf{x}) d\mathbf{x} \\ &= (T^{[s]}(\psi))(\phi), \end{aligned}$$

where we have used the substitution $\mathbf{y} \rightarrow -\mathbf{y}$ to obtain the second to last equality. Since $\mathcal{S}(\mathbb{R}^n)$ is dense in $H^{s,2}(\mathbb{R}^n)$ and in $L^2(\mathbb{R}^n)$, we get that $(T^{[s]})^* = T^{[s]}$.

¹³It is a good exercise to test the reader understanding of the topology on $\mathcal{D}(\mathbb{R}^n)$.

Since $T^{[s]}$ is an isometry of $H^{s,2}(\mathbb{R}^n)$ onto $L^2(\mathbb{R}^n)$, we have that $(T^{[s]})^*$ is an isometry of $L^2(\mathbb{R}^n)$ onto $((H^{s,2}(\mathbb{R}^n))^*)^{14}$. So, $T^{[s]} = (T^{[s]})^*$ is an isometry of $L^2(\mathbb{R}^n)$ onto $((H^{s,2}(\mathbb{R}^n))^*)^*$.

For $\phi \in \mathcal{S}(\mathbb{R}^n)$, we have that

$$\begin{aligned} \|T^{[s]}(\phi)\|_{-s,\rho}^2 &= \int_{\mathbb{R}^n} (1 + \|\mathbf{z}\|_2^2)^{-s} |\mathcal{F}(\mathcal{F}^{-1}((1 + \|\mathbf{y}\|_2^2)^{s/2} \mathcal{F}(\phi)))(\mathbf{z})|^2 d\mathbf{z} \\ &= \int_{\mathbb{R}^n} |\mathcal{F}(\phi)(\mathbf{z})|^2 d\mathbf{z} = \|\mathcal{F}(\phi)\|_2^2 = \|\phi\|_2^2, \end{aligned}$$

where we have used Plancherel theorem, Theorem 6.6.6, to obtain the last equality. Hence, $T^{[s]}(L^2(\mathbb{R}^n)) \subset H^{-s,2}(\mathbb{R}^n) \subset (H^{s,2}(\mathbb{R}^n))^*$ since $\mathcal{S}(\mathbb{R}^n)$ is dense in $L^2(\mathbb{R}^n)$. It follows that $T^{[s]}$ is an isometry of $L^2(\mathbb{R}^n)$ onto $H^{-s,2}(\mathbb{R}^n)$.

To be thorough, we shall probably prove our previous claim that $H^{-s,2}(\mathbb{R}^n) \subset (H^{s,2}(\mathbb{R}^n))^*$. This is a consequence of the following observation. Given $f \in H^{-s,2}(\mathbb{R}^n)$, we have that

$$\begin{aligned} |f(g)| &= \left| \int_{\mathbb{R}^n} f(\mathbf{x})g(\mathbf{x}) d\mathbf{x} \right| = \left| \int_{\mathbb{R}^n} \hat{f}(\mathbf{y})\hat{g}(\mathbf{x}) d\mathbf{x} \right| \\ &= \left| \int_{\mathbb{R}^n} ((1 + \|\mathbf{y}\|_2^2)^{-s/2} \hat{f}(\mathbf{y}))((1 + \|\mathbf{y}\|_2^2)^{s/2} \hat{g}(\mathbf{y})) d\mathbf{y} \right| \\ &\leq \left(\int_{\mathbb{R}^n} (1 + \|\mathbf{y}\|_2^2)^{-s} |\hat{f}(\mathbf{y})|^2 d\mathbf{y} \right)^{1/2} \left(\int_{\mathbb{R}^n} (1 + \|\mathbf{y}\|_2^2)^s |\hat{g}(\mathbf{y})|^2 d\mathbf{y} \right)^{1/2} = \|f\|_{-s,\rho} \|g\|_{s,\rho} \end{aligned}$$

for all $g \in H^{s,2}(\mathbb{R}^n)$, where the second equality comes from Plancherel theorem, Theorem 6.6.6, and the inequality comes from Schwarz inequality. ■

It is possible to describe the elements of the dual space $H^{-k,2}(\Omega)$.

Theorem 11.5.4

Let Ω be an open subset of $\mathbb{R}^n$. Every $F \in H^{-k,2}(\Omega)$ can be expressed as

$$\langle F, f \rangle = F(f) = \sum_{|\alpha| \leq k} \int_{\Omega} F_{\alpha} D^{\alpha} f \quad (11.5.1)$$

for all $f \in H_0^{k,2}(\Omega)$, where the F_{α} are some functions in $L^2(\Omega)$.

Remark 11.5.5

Since $\mathcal{D}(\Omega)$ is dense in $H_0^{k,2}(\Omega)$, (11.5.1) is equivalent to $F = \sum_{|\alpha| \leq k} (-1)^{|\alpha|} D^{\alpha} F_{\alpha}$ in the sense of distribution. ♠

Proof.

Let $\{\alpha_1, \alpha_2, \dots, \alpha_N\}$, where $N = \sum_{|\alpha| \leq k} 1$, be an ordering of the multi-indices $\alpha \in \mathbb{N}^n$ such that $|\alpha| \leq k$. Consider the mapping

$$T : H_0^{k,2}(\Omega) \rightarrow (L^2(\Omega))^N$$

¹⁴We leave it to the reader to verify that the adjoint of a surjective isometry is a surjective isometry.

$$f \mapsto (D^{\alpha_1} f, D^{\alpha_2} f, \dots, D^{\alpha_N} f)$$

where $(L^2(\Omega))^N$ is an Hilbert space with the scalar product defined by $\langle \mathbf{f}, \mathbf{g} \rangle_2 = \sum_{j=0}^N \int_{\Omega} f_{\alpha_j} g_{\alpha_j}$ for all $\mathbf{f}, \mathbf{g} \in (L^2(\Omega))^N$. T is a linear mapping preserving the norm if the norm on $(L^2(\Omega))^N$ is the norm induced by the scalar product; namely, $\|\mathbf{f}\|_2 = \left(\sum_{j=0}^N \|f_{\alpha_j}\|_2^2 \right)^{1/2}$ for all $\mathbf{f} \in (L^2(\Omega))^N$.

Let $R = T(H_0^{k,2}(\Omega))$. Since T preserves the norm (so T is a closed mapping), R is a closed subspace of $(L^2(\Omega))^N$. Moreover, $T : H_0^{k,2}(\Omega) \rightarrow R$ is invertible. Let $S : R \rightarrow H_0^{k,2}(\Omega)$ be the inverse of T . S is an isometry and so a bounded linear mapping.

Consider the bounded linear functional

$$\begin{aligned} P : R &\rightarrow \mathbb{R} \\ g &\mapsto F(S(g)) = \langle F, S(g) \rangle \end{aligned}$$

Using Hahn-Banach theorem, we can obtain a bounded linear functional $Q : (L^2(\Omega))^N \rightarrow \mathbb{R}$ such that $Q|_R = P$ and $\|Q\| = \|P\|$; the norm of the operators is induced by the norm $\|\cdot\|_2$ on $(L^2(\Omega))^N$.

Using Riesz representation theorem, there exist $F_{\alpha} \in L^2(\Omega)$ for $|\alpha| \leq k$ such that

$$\langle F, f \rangle = \langle Q, T(f) \rangle = Q(T(f)) = \sum_{j=0}^N \int_{\Omega} F_{\alpha_j} D^{\alpha} f$$

for all $f \in H_0^{k,2}(\Omega)$. ■

11.6 Extension

Let Ω be an open subset of $\mathbb{R}^n$. The space $W_0^{k,2}(\Omega)$ can always be extended to $W^{k,2}(\mathbb{R}^n)$. If $f \in W_0^{k,2}(\Omega)$, we define $f(\mathbf{x}) = 0$ for $\mathbf{x} \in \mathbb{R}^n \setminus \Omega$ to get $f \in W^{k,2}(\mathbb{R}^n)$. The elements of $W_0^{k,2}(\Omega)$ are the elements of $W^{k,2}(\mathbb{R}^n)$ which are the limit with respect to the norm $\|\cdot\|_{k,2}$ of test functions with compact support in Ω .

Since it is generally simpler to work with $W^{k,2}(\mathbb{R}^n)$ than with $W^{k,2}(\Omega)$, where Ω is a bounded domain in $\mathbb{R}^n$, it should be useful to determine under which conditions $W^{k,2}(\Omega)$ may be extended to a subspace of $W^{k,2}(\mathbb{R}^n)$.

The concept of extension is also important when we have boundary conditions. It may be easier to deal with boundary conditions imposed on functions defined on $\mathbb{R}^n$ instead of functions defined on a bounded open set Ω .

Definition 11.6.1

Let Ω be an open subset of $\mathbb{R}^n$. Let E be a linear operator from the space of measurable functions on Ω to the space of measurable function on $\mathbb{R}^n$.

1. E is called a **(k, p)-extension operator** on Ω if the restriction of E to $W^{k,p}(\Omega)$ is a bounded linear mapping $E : W^{k,p}(\Omega) \rightarrow W^{k,p}(\mathbb{R}^n)$ such that $E(f)|_{\Omega} = f$ for all $f \in W^{k,p}(\Omega)$. If Ω has an (k, p) -extension operator, we say that Ω has the **(k, p)-extension property**.
2. E is called a **strong k-extension operator** on Ω if, for all $0 \leq m \leq k$ and $0 < p < \infty$, the restriction of E to $W^{m,p}(\Omega)$ is a (m, p) -extension operator on Ω . If Ω has a strong k -extension operator, we say that Ω has the **strong k-extension property**.
3. E is called a **total extension operator** on Ω if it is a strong k -extension operator for every $k \geq 0$.

The next proposition gives a simple example of how the concept of extension can be used. It is Theorem 11.1.11 for $p = 2$.

Theorem 11.6.2

Let Ω be an open subset of $\mathbb{R}^n$ with the $(k, 2)$ -extension property. Then $\mathcal{D}(\mathbb{R}^n)|_{\Omega}$ is dense in $W^{k,2}(\Omega)$ with respect to the norm $\|\cdot\|_{k,2}$.

Proof.

We have from Theorem 11.1.16 that $\mathcal{D}(\mathbb{R}^n)$ is dense in $W^{k,2}(\mathbb{R}^n) \supset E(W^{k,2}(\Omega))$.

Given $f \in W^{k,2}(\Omega)$, suppose that $\{\phi_j\}_{j=0}^{\infty} \subset \mathcal{D}(\mathbb{R}^n)$ converges to $E(f)$ with respect to the norm $\|\cdot\|_{k,2,\mathbb{R}^n}$. Then, $\phi_j|_{\Omega} \rightarrow f$ in $W^{k,2}(\Omega)$ with respect to the norm $\|\cdot\|_{k,2,\Omega}$ as $j \rightarrow \infty$ according to Lemma 11.1.18. ■

A first example of an open subset of $\mathbb{R}^n$ with a total extension operator is given by the following theorem.

Theorem 11.6.3

There exists a strong k -extension operator E on $\mathbb{R}_+^n \equiv \{\mathbf{x} \in \mathbb{R}^n : x_n > 0\}$. Given $f \in W^{k,p}(\mathbb{R}_+^n)$,

$$\text{supp } E(f) \subset \text{supp } f \cup \left(\bigcup_{j=1}^{k+1} \{(x_1, x_2, \dots, x_{n-1}, -(1/j)x_n) : \mathbf{x} \in \text{supp } f\} \right).$$

Moreover, for all multi-indices α such that $|\alpha| \leq k$, there exists a strong $(k - |\alpha|)$ -extension operator E_{α} on $\mathbb{R}_+^n$ such that $D^{\alpha}(E(f)) = E_{\alpha}(D^{\alpha} f)$ for all $f \in W^{k,p}(\mathbb{R}_+^n)$.

Remark 11.6.4

It is also useful to define for later use the sets $\mathbb{R}_0^n = \{\mathbf{x} \in \mathbb{R}^n : x_n = 0\}$ and $\mathbb{R}_-^n = \{\mathbf{x} \in \mathbb{R}^n : x_n < 0\}$. ♠

Proof.

Let $\mathbf{a} \in \mathbb{R}^{k+1}$ be the unique solution of the linear equation $A\mathbf{a} = \mathbf{u}$, where A is the Vandermonde matrix defined by $A_{i,j} = (-j)^{i-1}$ for $1 \leq i, j \leq k+1$ and $u_j = 1$ for $1 \leq j \leq k+1$. The a_j will be used in the definition of the extension operators E and E_α to ensure that $E(\phi)$ and $E_\alpha(\phi)$ are sufficiently differentiable functions on $\mathbb{R}^n$.

Let

$$E(\phi)(\mathbf{x}) = \begin{cases} \phi(\mathbf{x}) & \text{if } \mathbf{x} \in \mathbb{R}_+^n \cup \mathbb{R}_0^n \\ \sum_{j=1}^{k+1} a_j \phi(\check{\mathbf{x}}, -jx_n) & \text{if } \mathbf{x} \in \mathbb{R}_-^n \end{cases}$$

and

$$E_\alpha(\phi)(\mathbf{x}) = \begin{cases} \phi(\mathbf{x}) & \text{if } \mathbf{x} \in \mathbb{R}_+^n \cup \mathbb{R}_0^n \\ \sum_{j=1}^{k+1} (-j)^{\alpha_n} a_j \phi(\check{\mathbf{x}}, -jx_n) & \text{if } \mathbf{x} \in \mathbb{R}_-^n \end{cases}$$

for $\phi \in S = \{\phi|_{\mathbb{R}_+^n \cup \mathbb{R}_0^n} : \phi \in \mathcal{D}(\mathbb{R}^n)\}$, where $\check{\mathbf{x}} = (x_1, x_2, \dots, x_{n-1})$ for all $\mathbf{x} \in \mathbb{R}^n$ and $m \leq k$ is arbitrary but fixed. We prove that these operators satisfy that statement of the theorem.

We have that $E(\phi)$ is of class C^m on $\mathbb{R}_+^n \cup \mathbb{R}_-^n$ for all $\phi \in S$. Moreover,

$$D^\alpha(E(\phi))(\mathbf{x}) = \sum_{j=1}^{k+1} (-j)^{\alpha_n} a_j (D^\alpha \phi)(\check{\mathbf{x}}, -jx_n) = E_\alpha(D^\alpha \phi)(\mathbf{x})$$

for $|\alpha| \leq m$ and $\mathbf{x} \in \mathbb{R}_-^n$ and obviously

$$D^\alpha(E(\phi))(\mathbf{x}) = E_\alpha(D^\alpha \phi)(\mathbf{x}) = D^\alpha \phi(\mathbf{x})$$

for $|\alpha| \leq m$ and $\mathbf{x} \in \mathbb{R}_+^n$. Moreover, we have that

$$\lim_{\substack{\mathbf{x} \rightarrow (\check{\mathbf{z}}, 0) \\ x_n < 0}} D^\alpha(E(\phi))(\mathbf{x}) = \sum_{j=1}^{k+1} (-j)^{\alpha_n} a_j \left(\lim_{\substack{\mathbf{x} \rightarrow (\check{\mathbf{z}}, 0) \\ x_n > 0}} (D^\alpha \phi)(\mathbf{x}) \right) = \lim_{\substack{\mathbf{x} \rightarrow (\check{\mathbf{z}}, 0) \\ x_n > 0}} D^\alpha(E(\phi))(\mathbf{x})$$

for $|\alpha| \leq m$ because $\sum_{j=1}^{k+1} (-j)^{\alpha_n} a_j = 1$ for $0 \leq \alpha_n \leq m \leq k$. Thus, $E(\phi) \in C^m(\mathbb{R}^n)$ for $\phi \in S$.

To prove that E is a bounded linear operator, we note that

$$\begin{aligned} \|E(\phi)\|_{m,p,\mathbb{R}^n} &= \left(\sum_{|\alpha| \leq m} \|D^\alpha E(\phi)\|_p^p \right)^{1/p} = \left(\sum_{|\alpha| \leq m} \int_{\mathbb{R}^n} |D^\alpha(E(\phi))(\mathbf{x})|^p d\mathbf{x} \right)^{1/p} \\ &= \left(\sum_{|\alpha| \leq m} \left(\int_{\mathbb{R}_+^n} |D^\alpha \phi(\mathbf{x})|^p d\mathbf{x} + \int_{\mathbb{R}_-^n} \left| \sum_{j=1}^{k+1} (-j)^{\alpha_n} a_j D^\alpha \phi(\check{\mathbf{x}}, -jx_n) \right|^p d\mathbf{x} \right) \right)^{1/p} \\ &\leq \left(\sum_{|\alpha| \leq m} \left(\int_{\mathbb{R}_+^n} |D^\alpha \phi(\mathbf{x})|^p d\mathbf{x} \right. \right. \end{aligned}$$

$$+ \sum_{\substack{0 \leq s_1 \leq s_2 \leq \dots \leq s_p \leq p \\ s_1 + s_2 + \dots + s_p = p \\ 1 \leq j_1 < j_2 < \dots < j_p \leq k+1}} c_{s_1, s_2, \dots, s_p} \int_{\mathbb{R}_+^n} \prod_{i=1}^p |(-j_i)^{\alpha_n} D^\alpha \phi(\check{\mathbf{x}}, -j_i x_n)|^{s_i} d\mathbf{x} \Bigg)^{1/p},$$

where the $c_{s_1, s_2, \dots, s_p} \geq 0$ are constants. In the previous relation and those that follow below, we assume that factors with $j_i = 0$ are removed from the products.

Using the generalized Hölder's inequality with $\sum_{i=1}^p \frac{1}{p/s_i} = 1$, we get

$$\begin{aligned} \int_{\mathbb{R}_+^n} \prod_{i=1}^p |(-j_i)^{\alpha_n} D^\alpha \phi(\check{\mathbf{x}}, -j_i x_n)|^{s_i} d\mathbf{x} &\leq \prod_{i=1}^p \left(\int_{\mathbb{R}_+^n} |(-j_i)^{\alpha_n} D^\alpha \phi(\check{\mathbf{x}}, -j_i x_n)|^p d\mathbf{x} \right)^{s_i/p} \\ &= \prod_{i=1}^p \left(\int_{\mathbb{R}_+^n} j_i^{\alpha_n p - 1} |D^\alpha \phi(\mathbf{x})|^p d\mathbf{x} \right)^{s_i/p} = \left(\prod_{i=1}^p j_i^{(\alpha_n p - 1)s_i/p} \right) \int_{\mathbb{R}_+^n} |D^\alpha \phi(\mathbf{x})|^p d\mathbf{x}, \end{aligned}$$

where we have used the substitution $x_n \rightarrow -x_n/j_i$ for each of the integrals in the product to obtain the first equality. Therefore,

$$\begin{aligned} \|E(\phi)\|_{m,p,\mathbb{R}^n} &\leq \left(\sum_{|\alpha| \leq m} \left(1 + \sum_{\substack{0 \leq s_1 \leq s_2 \leq \dots \leq s_p \leq p \\ s_1 + s_2 + \dots + s_p = p \\ 1 \leq j_1 < j_2 < \dots < j_p \leq k+1}} c_{s_1, s_2, \dots, s_p} \left(\prod_{i=1}^p j_i^{(\alpha_n p - 1)s_i/p} \right) \right) \int_{\mathbb{R}_+^n} |D^\alpha \phi(\mathbf{x})|^p d\mathbf{x} \right)^{1/p} \\ &\leq C \|\phi\|_{m,p,\mathbb{R}_+^n} \end{aligned} \quad (11.6.1)$$

for all $\phi \in S$, $0 \leq m \leq k$ and $0 < p < \infty$, where

$$C = N \left(1 + \sum_{\substack{0 \leq s_1 \leq s_2 \leq \dots \leq s_p \leq p \\ s_1 + s_2 + \dots + s_p = p \\ 1 \leq j_1 < j_2 < \dots < j_p \leq k+1}} c_{s_1, s_2, \dots, s_p} \left(\prod_{i=1}^p j_i^{(mp-1)s_i/p} \right) \right)^p,$$

where $N = \sum_{|\alpha| \leq m} 1$.

Since E is defined by extension from the dense set S in $W^{m,p}(\mathbb{R}_+^n)$ according to Theorem 11.1.11, we have that (11.6.1) is true for all $\phi \in W^{m,p}(\mathbb{R}_+^n)$. Thus, proving that $E : W^{m,p}(\mathbb{R}_+^n) \rightarrow W^{m,p}(\mathbb{R}^n)$ is a (m,p) -extension operator on $\mathbb{R}_+^n$ for all $0 \leq m \leq k$ and $0 < p < \infty$.

To prove that E_α is a strong $(k-|\alpha|)$ -extension operator on $\mathbb{R}_+^n$, we note that $D^\beta(E_\alpha(\phi)) = E_\alpha(D^\beta \phi) = E_{\alpha+\beta}(\phi)$ for all $\phi \in S$ and $0 \leq |\beta| \leq k-|\alpha|$. Then, we repeat the previous arguments with E and α replaced by E_α and β respectively, and $W^{m,p}(\mathbb{R}_+^n)$ for $0 \leq m \leq k-|\alpha|$. ■

There is a version of the previous theorem for total extension operators.

Theorem 11.6.5

There exists a total extension operator on $\mathbb{R}_+^n = \{\mathbf{x} : x_n > 0\}$.

Proof (Sketch).

The proof resembles the proof of the previous theorem. However, there are some major technical details. We only sketch the proof. Consult Lemma 5.20 and Theorem 5.21 in [2] for all the details of the proof.

Since $S = \left\{ \phi|_{\mathbb{R}_+^n} : \phi \in \mathcal{D}(\mathbb{R}^n) \right\}$ is dense in $W^{k,p}(\mathbb{R}_+^n)$ according to Theorem 11.1.11, it is enough to define the extension operators on S .

Let $\psi : \mathbb{R} \rightarrow [0, 1]$ be a function of class C^∞ such that $\psi(x) = 1$ for $x \leq 1/2$ and $\psi(x) = 0$ for $x \geq 1$. We consider the operator E defined on S by

$$E(\phi)(\mathbf{x}) = \begin{cases} \phi(\mathbf{x}) & \text{if } \mathbf{x} \in \mathbb{R}_+^n \cup \mathbb{R}_0^n \\ \sum_{j=1}^{\infty} a_j \psi(-2^j x_n) \phi(\check{\mathbf{x}}, -2^j x_n) & \text{if } \mathbf{x} \in \mathbb{R}_-^n \end{cases}$$

where $\check{\mathbf{x}} = (x_1, x_2, \dots, x_{n-1})$ for all $\mathbf{x} \in \mathbb{R}^n$ and $\{a_j\}_{j=1}^{\infty} \subset \mathbb{R}$ is a sequence satisfying

$$\sum_{j=0}^{\infty} 2^{ij} a_j = (-1)^i \quad \text{and} \quad \sum_{j=0}^{\infty} 2^{ij} |a_j| < \infty \quad (11.6.2)$$

for $i \in \mathbb{N}$.

The function $E(\phi)$ is well defined because the series used to define $E(\phi)(\mathbf{x})$ for $\mathbf{x} \in \mathbb{R}_-^n$ is finite on any compact set $K \subset \mathbb{R}_-^n$. In fact, the summation is on j such that $2^j \inf_{\mathbf{x} \in K} |x_n| \leq 1$. Moreover, $E(\phi)$ has a compact support and $E(\phi) \in C(\overline{\mathbb{R}_+^n}) \cap C^\infty(\mathbb{R}_+^n)$.

The convergence of the last series in (11.6.2) is to ensure that the series used to define $E(\phi)(\mathbf{x})$ for $\mathbf{x} \in \mathbb{R}_-^n$ is absolutely and uniformly convergent as $x_n \rightarrow 0^-$. It is also used to show that E is a bounded operator from $W^{k,p}(\mathbb{R}_+^n)$ to $W^{k,p}(\mathbb{R}^n)$ for all $k \geq 0$ and $0 < p < \infty$. The first series in (11.6.2) is to ensure that $E(\phi) \in C^\infty(\mathbb{R}^n)$ for all $\phi \in S$ similarly to what we had in the proof of Theorem 11.6.3. ■

The next theorem gives conditions on open sets $\Omega \subset \mathbb{R}^n$ that ensure the existence of a strong k -extension operator on Ω . But first, we need some definitions.

Definition 11.6.6

A collection $\{U_j\}_{j=1}^{\infty}$ of open subsets of $\mathbb{R}^n$ is a **locally finite open cover** of a set $\Omega \subset \mathbb{R}^n$ if $\Omega \subset \bigcup_{j=1}^{\infty} U_j$ and, for any compact set $K \subset \mathbb{R}^n$, the set K intersects at most a finite number of U_j .

Definition 11.6.7

A open set $\Omega \subset \mathbb{R}^n$ satisfies the **uniform C^k -regularity condition** if there exists a locally finite open cover $\{U_j\}_{j=1}^{\infty}$ for $\partial\Omega$ and a collection $\{\psi_j\}_{j=1}^{\infty}$ of C^k -diffeomorphisms $\psi_i : U_i \rightarrow B_1(\mathbf{0}) \subset \mathbb{R}^n$ with the following properties:

1. There exists a positive integer J such that the intersection of any $J + 1$ sets U_j

is empty.

2. There exists $\delta > 0$ such that

$$\{\mathbf{x} \in \mathbb{R}^n : \text{dist}(\mathbf{x}, \partial\Omega) < \delta\} \subset \bigcup_{j=0}^{\infty} \psi_j^{-1}(B_{1/2}(\mathbf{0})) .$$

3. For each j , we have $\psi_j(U_j \cap \Omega) = B_1(\mathbf{0}) \cap \mathbb{R}_+^n$.

4. There exists $C > 0$ such that

$$\sup_{\mathbf{x} \in U_j} |D^{\alpha} \psi_{j,s}(\mathbf{x})| < C \quad \text{and} \quad \sup_{\mathbf{x} \in B_1(\mathbf{0})} |D^{\alpha} \psi_{j,s}^{-1}(\mathbf{x})| < C$$

for all multi-indices α with $|\alpha| \leq k$, all $j > 0$, and all $s \in \{1, 2, \dots, n\}$. Recall that $\psi_{j,s}$ denoted the s component of ψ_j with a similar meaning for $\psi_{j,s}^{-1}$.

If $\partial\Omega$ is bounded (so compact), we may assume that the locally finite cover of $\partial\Omega$ is a finite collection $\{U_j\}_{j=1}^J$ for some positive integer J .

Theorem 11.6.8

Let Ω be an open subset of $\mathbb{R}^n$ with a bounded boundary ¹⁵ and suppose that Ω satisfies the uniform C^k -regularity condition, then there exists a strong k -extension operator F on Ω .

Proof (Sketch).

Let $\{U_j\}_{j=1}^J$ be the finite cover of $\partial\Omega$ from the definition of uniform C^k -regularity, and $\{\psi_j\}_{j=1}^J$ be the associated collection of C^k -diffeomorphisms with the properties listed in the definition of uniform C^k -regularity.

Let $V_j = \psi_j^{-1}(B_{1/2}(\mathbf{0}))$ and $W_j = \psi_j^{-1}(B)$ for $1 \leq j \leq J$, where

$$B = \{\mathbf{y} \in B_1(\mathbf{0}) : \|\check{\mathbf{y}}\| < 1/2 \text{ and } |y_n| < \sqrt{3}/2\} .$$

Since $B_{1/2}(\mathbf{0}) \subset B \subset B_1(\mathbf{0})$, we have that $V_j \subset W_j \subset U_j$ for $1 \leq j \leq J$ (Figure 11.4).

As stated in the definition of uniform C^k -regularity, there exists $\delta > 0$ such that

$$\{\mathbf{x} \in \mathbb{R}^n : \text{dist}(\mathbf{x}, \partial\Omega) < \delta\} \subset \bigcup_{j=1}^J V_j .$$

Let $W_0 = \{\mathbf{x} \in \Omega : \text{dist}(\mathbf{x}, \partial\Omega) > \delta/2\}$. We then have that $\{W_j\}_{j=0}^J$ is an open cover of Ω . Let $\{\phi_j\}_{j=0}^J$ be a partition of unity for Ω subordinated to $\{W_j\}_{j=0}^J$.

Since Ω satisfies the uniform C^k -regularity condition, we have that $\partial\Omega$ is a boundary of class C^k . Thus, $\partial\Omega$ satisfies the segment property. It then follows from Theorem 11.1.11

¹⁵The open set Ω does not have to be bounded. This result may be used for problems on an unbounded domain like we had for the exterior Neumann or Dirichlet problems in a previous chapter.

that $S = \{\phi|_{\Omega} : \phi \in \mathcal{D}(\mathbb{R}^n)\}$ is dense in $W^{m,p}(\Omega)$ for $m \leq k$ arbitrary but fixed. Hence, as we have done before, it is enough to define the extension operators on S .

Given $\phi \in S$, we have that $\phi = \sum_{j=0}^J \omega_j$, where $\omega_j = \phi \phi_j \in C^\infty(\mathbb{R}^n)$ has its support in W_j .

For $j = 0$, let

$$\rho_0(\mathbf{x}) = \begin{cases} \omega_0(\mathbf{x}) & \text{if } \mathbf{x} \in W_0 \\ 0 & \text{if } \mathbf{x} \in \mathbb{R}^n \setminus W_0 \end{cases}$$

So, ρ_0 is an extension of ω_0 to $\mathbb{R}^n$ and, according to Lemma 11.1.7, $\rho_0 \in W^{m,p}(\mathbb{R}^n)$ since $\text{supp } \omega_0 \subset \Omega$.

For $0 < j \leq J$, we consider $\tilde{\omega}_j = \omega_j \circ \psi_j^{-1}$ on B . Since $\text{supp } \tilde{\omega}_j \subset B$, we may again use Lemma 11.1.7 to extend $\tilde{\omega}_j$ to $\mathbb{R}^n$ by setting $\tilde{\omega}_j(\mathbf{x}) = 0$ for $\mathbf{x} \in \mathbb{R}^n \setminus B$. We have that $\tilde{\omega}_j \in W^{m,p}(\mathbb{R}_+^n)$ according to Lemma 11.1.18. Using the strong k -extension operator E provided by Theorem 11.6.3, we have that $E(\tilde{\omega}_j) \in W^{m,p}(\mathbb{R}^n)$, $E(\tilde{\omega}_j)|_{\mathbb{R}_+^n} = \tilde{\omega}_j$ and $\|E(\tilde{\omega}_j)\|_{m,p,\mathbb{R}^n} \leq C_j \|\tilde{\omega}_j\|_{m,p,\mathbb{R}_+^n}$ for a constant C_j that depends on m and p . Since $\text{supp } \tilde{\omega}_j \subset B$ and $\text{supp } E(\tilde{\omega}_j) \subset B$, we have that $\|E(\tilde{\omega}_j)\|_{m,p,B} \leq C_j \|\tilde{\omega}_j\|_{m,p,B_+}$ where $B_+ = \{\mathbf{x} \in B : x_n > 0\}$.

Let $\rho_j = E(\tilde{\omega}_j) \circ \psi_j$ on W_j . We have that $\rho_j \in C_c^m(W_j)$ and

$$\rho_j(\mathbf{x}) = E(\tilde{\omega}_j)(\psi_j(\mathbf{x})) = \tilde{\omega}_j(\psi_j(\mathbf{x})) = \omega_j(\mathbf{x})$$

for $\mathbf{x} \in W_j \cap \Omega$ because $\psi_j(\mathbf{x}) \in B_+$. Since $\text{supp } \rho_j \subset W_j$, we may again use Lemma 11.1.7 to extend ρ_j to a function in $W^{m,p}(\mathbb{R}^n)$ by setting $\rho_j(\mathbf{x}) = 0$ for $\mathbf{x} \in \mathbb{R}^n \setminus W_j$.

A lengthy computation shows that

$$\|\rho_j\|_{m,p,\mathbb{R}^n} = \|\rho_j\|_{m,p,W_j} \leq K \|\omega_j\|_{m,p,W_j \cap \Omega} = K \|\omega_j\|_{m,p,\Omega}$$

for some constant K that depends only of m and p . To justify this claim, one has to use item (4) in the definition of uniform C^k -regularity. We may also assume that $K \geq 1$. Consult Theorem 5.22 in [2] for the details.

The operator F is defined on S by $F(\phi) = \sum_{j=0}^J \rho_j$. We have that $F(\phi) \in W^{m,p}(\mathbb{R}^n)$ by construction,

$$F(\phi)(\mathbf{x}) = \sum_{j=0}^J \rho_j(\mathbf{x}) = \sum_{j=0}^J \omega_j(\mathbf{x}) = \phi(\mathbf{x})$$

for $\mathbf{x} \in \Omega$, and

$$\|F(\phi)\|_{m,p,\mathbb{R}^n} \leq \sum_{j=0}^J \|\rho_j\|_{m,p,\mathbb{R}^n} \leq K \sum_{j=0}^J \|\omega_j\|_{m,p,\Omega} \leq QK(1+J)\|\phi\|_{m,p,\Omega}, \quad (11.6.3)$$

where Q is a constant that depends on $\sup_{\mathbf{x} \in \mathbb{R}^n} |D^\alpha \phi_j(\mathbf{x})|$ for $0 \leq j \leq J$ and $|\alpha| \leq m$.

Note that F does not depend on $0 \leq m \leq k$ or $0 < p < \infty$ as it is the case for the strong k -extension operator E used above.

Since F is defined by extension from the dense set S in $W^{m,p}(\Omega)$, we have that (11.6.3) is true for all $\phi \in W^{m,p}(\Omega)$. Thus, proving that $F : W^{m,p}(\Omega) \rightarrow W^{m,p}(\mathbb{R}^n)$ is a (m, p) -extension operator on Ω for all $0 \leq m \leq k$ and $0 < p < \infty$. ■

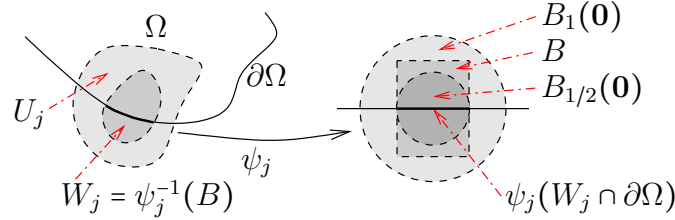


Figure 11.4: Representation of the local structure provided by the uniform C^k -regularity of Ω in the proof of Theorem 11.6.8.

If Ω is an open subset of $\mathbb{R}^n$ with a compact boundary of class C^k (so $\partial\Omega$ is a compact oriented manifold of class C^k), then Ω satisfies the uniform C^k -regularity condition. There exists therefore a strong k -extension operator on Ω .

11.7 Imbedding

Imbeddings will be used in the next chapter to show that solutions of linear elliptic partial differential equations in Sobolev spaces are in fact smoother than predicted.

There are a lot of results about imbedding of Sobolev spaces into others spaces of functions; in particular into spaces of differentiable functions. The interested reader should consult [2] for more results on this subject. We have included some of the most important results, at least a version of them because there may be many versions of these results. Our intention was to include only the results that we need for our study of partial differential equations. However, there are some fundamental results that we could not ignore because of their significance even if we may not use them later.

11.7.1 In $\mathbb{R}$

Before looking at results in higher dimensions, it is interesting to consider $H^{1,p}(I)$, where I is an open interval of $\mathbb{R}$. We first start by generalizing the Fundamental Theorem of Calculus to the theory of distributions.

Lemma 11.7.1

Suppose that $I =]a, b[\subset \mathbb{R}$ and that $f \in L^1_{loc}(I)$. For $x_0 \in I$, let

$$v(x) = \int_{x_0}^x f(t) dt .$$

Then $v \in C(\bar{I})$ and

$$\int_I v(x) \phi'(x) dx = - \int_I f(x) \phi(x) dx$$

for all $\phi \in \mathcal{D}(I)$; namely, $v' = f$ where v' is the derivative of v in the sense of distributions.

Proof.

That v is continuous on $\bar{I}$ can be easily deduced from Theorem 8.17 in [30].

For $\phi \in \mathcal{D}(I)$, we have

$$\begin{aligned} \int_I v(x) \phi'(x) dx &= \int_I \left(\int_{x_0}^x g(t) dt \right) \phi'(x) dx \\ &= - \int_a^{x_0} \left(\int_x^{x_0} g(t) \phi'(x) dt \right) dx + \int_{x_0}^b \left(\int_{x_0}^x g(t) \phi'(x) dt \right) dx \end{aligned}$$

Using Fubini's Theorem, we get

$$\begin{aligned} \int_I v(x) \phi'(x) dx &= - \int_a^{x_0} \left(\int_a^t g(t) \phi'(x) dx \right) dt + \int_{x_0}^b \left(\int_t^b g(t) \phi'(x) dx \right) dt \\ &= - \int_a^{x_0} g(t) \phi(t) dt - \int_{x_0}^b g(t) \phi(t) dt = - \int_I g(t) \phi(t) dt. \quad \blacksquare \end{aligned}$$

Theorem 11.7.2

Given $u \in H^{1,p}(I)$, where $I =]a, b[\subset \mathbb{R}$, there exists $w \in C(\bar{I})$ such that $w = u$ almost everywhere on I . Moreover,

$$w(x) - w(y) = \int_y^x u'(t) dt$$

for all x and y in I , where u' is the derivative of u in the sense of distributions.

Proof.

Given $x_0 \in I$, let

$$\tilde{w}(x) = \int_{x_0}^x u'(t) dt$$

for $x \in I$. The function $\tilde{w}$ is well defined because $u \in H^{1,p}(I)$, so $u' \in L^p(I) \subset L^p(J)$ where J is the bounded interval with endpoints x_0 and x . Moreover, $L^p(J) \subset L^1(J)$ because J is bounded.

From Lemma 11.7.1, we have that

$$\int_I \tilde{w}(x) \phi'(x) dx = - \int_I u'(x) \phi(x) dx = \int_I u(x) \phi'(x) dx$$

for all $\phi \in \mathcal{D}(I)$. Thus,

$$\int_I (\tilde{w}(x) - u(x)) \phi'(x) dx = 0$$

for all $\phi \in \mathcal{D}(I)$. It follows from Lemma 6.4.4 that there exists a constant C such that $\tilde{w} - u = C$ almost everywhere on I . The function w is defined by $w = \tilde{w} - C$.

Note that

$$w(x) - w(y) = \tilde{w}(x) - \tilde{w}(y) = \int_{x_0}^x u'(t) dt - \int_{x_0}^y u'(t) dt = \int_y^x u'(t) dt .$$

for all x and y in I . ■

It follows from the previous theorem that $H^{1,p}(I) \subset C(\bar{I})$. We prove more general and stronger results in the next section.

11.7.2 In $\mathbb{R}^n$

To illustrate the complexity of the results about imbeddings, we summarize some of the results listed in [3].

	$\frac{1}{p'} = \frac{1}{p} - \frac{1}{n} > 0$	$p \leq q < \infty$ or $p = n$	$p > n$
$\mathbb{R}^n$	$H^{1,p}(\mathbb{R}^n) \rightarrow L^q(\mathbb{R}^n)$ for $p \leq q \leq p'$ is continuous.	$H^{1,p}(\mathbb{R}^n) \rightarrow L^q(\mathbb{R}^n)$ is continuous.	$H^{1,p}(\mathbb{R}^n) \rightarrow L^\infty(\mathbb{R}^n)$ is continuous.
Ω open, and $\partial\Omega$ bounded and of class C^1	$H^{1,p}(\Omega) \rightarrow L^{p'}(\Omega)$ is continuous.	$H^{1,p}(\Omega) \rightarrow L^q(\Omega)$ is continuous.	$H^{1,p}(\Omega) \rightarrow L^\infty(\Omega)$ is continuous.
Ω open and bounded, and $\partial\Omega$ of class C^1	$H^{1,p}(\Omega) \rightarrow L^q(\Omega)$ for $1 \leq q < p'$ is compact.	$H^{1,p}(\Omega) \rightarrow L^q(\Omega)$ is compact.	$H^{1,p}(\Omega) \rightarrow C(\bar{\Omega})$ is compact.

There are similar results for $H^{m,p}$ with $m > 1$. This is a small sample of all the possible imbeddings. In [2], imbeddings are given based on the following criteria: Ω bounded or unbounded, $\partial\Omega$ satisfying only the segment property, $\partial\Omega$ of class C^k for $k \geq 1$, Ω satisfying the uniform C^k -regularity conditions, and several other criteria.

As mentioned previously, we will only give the results that will be useful to us later. This is sufficient to get an idea of what is involved in proving results about imbeddings.

Theorem 11.7.3 (Sobolev's Lemma)

If $s > k + n/2$, then $H^{s,2}(\mathbb{R}^n) \subset C^k(\mathbb{R}^n)$ and there exists a constant C depending on s and k only such that

$$\|f\|_{k,\infty} \leq C \|f\|_{s,\rho} \quad (11.7.1)$$

for $f \in H^{s,2}(\mathbb{R}^n)$. In particular, $H^{s,2}(\mathbb{R}^n) \subset C_b^k(\mathbb{R}^n)$. Moreover, the imbedding $i : H^{s,2}(\mathbb{R}^n) \rightarrow C_b^k(\mathbb{R}^n)$ is continuous.

Proof.

Let f be an element of $\mathcal{S}(\mathbb{R}^n) \subset H^{s,2}(\mathbb{R}^n)$. Note that $\mathcal{D}(\mathbb{R}^n) \subset H^{s,2}(\mathbb{R}^n)$ could also have been used.

Assume that $|\alpha| \leq k$. Since $(D^\alpha f)^\wedge(\mathbf{y}) = (i\mathbf{y})^\alpha \hat{f}(\mathbf{y})$, we get from Schwarz inequality that

$$\begin{aligned} \|(D^\alpha f)^\wedge\|_1 &= \int_{\mathbb{R}^n} |(-i\mathbf{y})^\alpha \hat{f}(\mathbf{y})| \, d\mathbf{y} \leq \int_{\mathbb{R}^n} (1 + \|\mathbf{y}\|^2)^{k/2} |\hat{f}(\mathbf{y})| \, d\mathbf{y} \\ &= \int_{\mathbb{R}^n} (1 + \|\mathbf{y}\|^2)^{(k-s)/2} (1 + \|\mathbf{y}\|^2)^{s/2} |\hat{f}(\mathbf{y})| \, d\mathbf{y} \\ &\leq \left(\int_{\mathbb{R}^n} (1 + \|\mathbf{y}\|^2)^{k-s} \, d\mathbf{y} \right)^{1/2} \left(\int_{\mathbb{R}^n} (1 + \|\mathbf{y}\|^2)^s |\hat{f}(\mathbf{y})|^2 \, d\mathbf{y} \right)^{1/2} = C_1 \|f\|_{s,\rho}, \end{aligned}$$

where

$$C = \left(\int_{\mathbb{R}^n} (1 + \|\mathbf{y}\|^2)^{k-s} \, d\mathbf{y} \right)^{1/2} < \infty$$

because $k - s < -n/2$. Note that the constant C does not depend on f but only on s and k . Since $D^\alpha f$, and $(D^\alpha f)^\wedge$ are in $L^1(\mathbb{R}^n)$, we get from Proposition 6.6.5 that $((D^\alpha f)^\wedge)^\vee \in C_0(\mathbb{R}^n)$, $((D^\alpha f)^\wedge)^\vee = D^\alpha f$ almost everywhere, and

$$\|D^\alpha f\|_\infty \leq \|(D^\alpha f)^\wedge\|_1 \leq C \|f\|_{s,\rho}.$$

Since this is true for all $|\alpha| \leq k$, we have that

$$\max_{|\alpha| \leq k} \|D^\alpha f\|_\infty \leq C \|f\|_{s,\rho} \quad (11.7.2)$$

for all $f \in \mathcal{S}(\mathbb{R}^n)$. Since $\mathcal{S}(\mathbb{R}^n)$ is dense in $H^{s,2}(\mathbb{R}^n)$, (11.7.2) is true for all $f \in H^{s,2}(\mathbb{R}^n)$. More precisely, if $\{f_i\}_{i=1}^\infty \subset \mathcal{S}(\mathbb{R}^n)$ converges in $H^{s,2}(\mathbb{R}^n)$ to $f \in H^{s,2}(\mathbb{R}^n)$, then $\{f_i\}_{i=1}^\infty$ is a Cauchy sequence in $H^{s,2}(\mathbb{R}^n)$. From (11.7.2), $\{f_i\}_{i=1}^\infty$ is a Cauchy sequence in the complete space $C_b^k(\mathbb{R}^n)$. Thus, there exists $g \in C_b^k(\mathbb{R}^n)$ such that $\{f_i\}_{i=1}^\infty$ converges in $C_b^k(\mathbb{R}^n)$ to g . Moreover, since $\{f_i\}_{i=1}^\infty$ converges in $H^{s,2}(\mathbb{R}^n)$ to f , we have that $\{f_i\}_{i=1}^\infty$ converges pointwise to f almost everywhere. Thus $f = g$ almost everywhere.

The continuity of the injection $i : H^{m,2}(\mathbb{R}^n) \rightarrow C_b^k(\mathbb{R}^n)$ is a direct consequence of (11.7.1). ■

Remark 11.7.4

In the previous theorem, we may have proved directly that $H^{s,2}(\mathbb{R}^n) \subset C^k(\mathbb{R}^n)$.

Let f be any element of $H^{s,2}(\mathbb{R}^n)$. Let V be a bounded open subset of $\mathbb{R}^n$. Choose $\psi \in \mathcal{D}(\mathbb{R}^n)$ such that $\psi = 1$ on $\overline{V}$, $|\psi| \leq 1$ and define $F = \psi f$. It follows from Proposition 11.1.22 that $F \in H^{s,2}(\mathbb{R}^n)$.

For $|\alpha| \leq k$, since $D^\alpha F \in L^2(\mathbb{R}^n)$ has a compact support, we also have that $D^\alpha F \in L^1(\mathbb{R}^n)$. Since $(D^\alpha F)^\wedge(\mathbf{y}) = (i\mathbf{y})^\alpha \hat{F}(\mathbf{y})$, as we did for f in the proof of the previous theorem, we get from Schwarz inequality that

$$\begin{aligned} \|(D^\alpha F)^\wedge\|_1 &= \int_{\mathbb{R}^n} |(-i\mathbf{y})^\alpha \hat{F}(\mathbf{y})| \, d\mathbf{y} \leq \int_{\mathbb{R}^n} (1 + \|\mathbf{y}\|^2)^{k/2} |\hat{F}(\mathbf{y})| \, d\mathbf{y} \\ &= \int_{\mathbb{R}^n} (1 + \|\mathbf{y}\|^2)^{(k-s)/2} (1 + \|\mathbf{y}\|^2)^{s/2} |\hat{F}(\mathbf{y})| \, d\mathbf{y} \\ &\leq \left(\int_{\mathbb{R}^n} (1 + \|\mathbf{y}\|^2)^{k-s} \, d\mathbf{y} \right)^{1/2} \left(\int_{\mathbb{R}^n} (1 + \|\mathbf{y}\|^2)^s |\hat{F}(\mathbf{y})|^2 \, d\mathbf{y} \right)^{1/2} = C_1 \|F\|_{s,\rho} \end{aligned}$$

for $|\alpha| \leq k$, where $C = \left(\int_{\mathbb{R}^n} (1 + \|\mathbf{y}\|^2)^{k-s} \, d\mathbf{y} \right)^{1/2} < \infty$ because $k - s < -n/2$.

Since $D^\alpha F$ and $(D^\alpha F)^\wedge$ are in $L^1(\mathbb{R}^n)$, we get from Proposition 6.6.5 that the inverse Fourier transform of $(D^\alpha F)^\wedge$ is in $C_0(\mathbb{R}^n)$ and is equal to $D^\alpha F$ almost everywhere for $|\alpha| \leq k$. From Proposition 6.4.5, we get that $F \in C^k(\mathbb{R}^n)$. Since this is true for any bounded open subset V of $\mathbb{R}^n$ and $f = F$ on V , then $f \in C^k(\mathbb{R}^n)$. ♠

Corollary 11.7.5

If $f \in H^{s,2}(\mathbb{R}^n)$ for all $s > n/2$, then $f \in C^\infty(\mathbb{R}^n)$.

Corollary 11.7.6

Every distribution on $\mathbb{R}^n$ with a compact support belongs to $H^{s,2}(\mathbb{R}^n)$ for some $s \in \mathbb{R}$.

Proof.

If u is a distribution and $K = \text{supp } u$ is compact, it follows from Proposition 6.3.1 that there exist a constant C_1 and an integer N such that $|u(\phi)| \leq C_1 \max_{|\alpha| \leq N} \|D^\alpha \phi\|_{K,\infty}$ for all $\phi \in \mathcal{D}_K$.

Thus, $|u(\phi)| \leq C_1 \max_{|\alpha| \leq N} \|D^\alpha \phi\|_\infty$ for all $\phi \in \mathcal{D}(\mathbb{R}^n)$ because $K = \text{supp } u$.

From Theorem 11.7.3, Sobolev's Lemma, there exists a constant C such that $\max_{|\alpha| \leq N} \|D^\alpha \phi\|_\infty \leq C \|\phi\|_{s,2}$ for all $\phi \in \mathcal{D}(\mathbb{R}^n)$ if $s > N + n/2$. Thus, $|u(\phi)| \leq CC_1 \|\phi\|_{s,2}$ for all $\phi \in \mathcal{D}(\mathbb{R}^n)$.

Since $\mathcal{D}(\mathbb{R}^n)$ is dense in $H^{s,2}(\mathbb{R})$, we may therefore extend u to a bounded linear functional on $H^{s,2}(\mathbb{R})$. Thus, $u \in (H^{s,2}(\mathbb{R}^n))^* = H^{-s,2}(\mathbb{R}^n)$. ■

Corollary 11.7.7

Let $\Omega \subset \mathbb{R}^n$ be an open set with the $(m, 2)$ -extension property. If $m > k + n/2$, then $H^{m,2}(\Omega) \subset C_b^k(\bar{\Omega})$. Moreover, the imbedding $i : H^{m,2}(\Omega) \rightarrow C_b^k(\bar{\Omega})$ is continuous.

Proof.

Let $E : H^{m,2}(\Omega) \rightarrow H^{m,2}(\mathbb{R}^n)$ be a bounded linear mapping provided by the $(m, 2)$ -extension

property. Given $f \in H^{m,2}(\Omega)$, we have from the Sobolev's Lemma that $E(f) \in H^{m,2}(\mathbb{R}^n) \subset C_b^k(\mathbb{R}^n)$ and the injection $i_0 : H^{m,2}(\mathbb{R}^n) \rightarrow C_b^k(\mathbb{R}^n)$ is continuous.

Let $R : C_b^k(\mathbb{R}^n) \rightarrow C_b^k(\overline{\Omega})$ be the mapping defined by $R(f) = f|_{\overline{\Omega}}$. This is a continuous mapping. Since $i = R \circ i_0 \circ E$, the composition of continuous mappings, we have that i is continuous.

Note that $f = i(f)|_{\Omega}$, so $i(f) \in C_b^k(\overline{\Omega})$ is the continuous extension of f from Ω to $\overline{\Omega}$. ■

To prove the next corollary, we need to introduce a new space of functions.

Definition 11.7.8

Given $0 < \mu \leq 1$, the space of **Hölder continuous functions of exponent μ** on the open set $\Omega \subset \mathbb{R}^n$, denoted $C^{k,\mu}(\overline{\Omega})$, is the set of all functions $f \in C_{bu}^k(\overline{\Omega})$ such that, for each multi-index α satisfying $|\alpha| \leq k$, there exists a constant K such that

$$|D^\alpha f(\mathbf{x}) - D^\alpha f(\mathbf{y})| \leq K \|\mathbf{x} - \mathbf{y}\|^\mu$$

for all $\mathbf{x}, \mathbf{y} \in \Omega$.

The spaces of Hölder continuous functions have several properties:

1. $C^{k,\mu}(\overline{\Omega})$ is a Banach space with respect to the norm

$$\|f\| = \|f\|_{k,\infty,\Omega} + \max_{0 \leq |\alpha| \leq k} \sup_{\substack{\mathbf{x}, \mathbf{y} \in \Omega \\ \mathbf{x} \neq \mathbf{y}}} \frac{|D^\alpha f(\mathbf{x}) - D^\alpha f(\mathbf{y})|}{\|\mathbf{x} - \mathbf{y}\|^\mu}.$$

2. $C^{k,\mu_2}(\overline{\Omega}) \subsetneq C^{k,\mu_1}(\overline{\Omega})$ for $0 < \mu_1 < \mu_2 < 1$.
3. Given $0 < \mu_1 < \mu_2 < 1$, the imbeddings $i_1 : C^{k,\mu_1}(\overline{\Omega}) \rightarrow C_{bu}^k(\overline{\Omega})$ and $i_2 : C^{k,\mu_2}(\overline{\Omega}) \rightarrow C^{k,\mu_1}(\overline{\Omega})$ are continuous. If Ω is bounded, the imbeddings i_1 and i_2 are compact operators.
4. If $m - k > \mu + n/2 > n/2 > m - k - 1$, then $H^{m,2}(\Omega) \subset C^{k,\mu}(\overline{\Omega})$ and the imbedding $i_3 : H^{m,2}(\Omega) \rightarrow C^{k,\mu}(\overline{\Omega})$ is continuous. If $n/2 = m - k - 1$, then $H^{m,2}(\Omega) \subset C^{k,\mu}(\overline{\Omega})$ and the imbedding $i_3 : H^{m,2}(\Omega) \rightarrow C^{k,\mu}(\overline{\Omega})$ is continuous for all $0 < \mu < 1$. These properties are often listed as parts of the Sobolev Imbedding Theorem.

All these results and many more are proved in [2] (i.e. Theorems 1.34 and 4.12). In this reference, the authors refer to a bounded open set that satisfies the “strong local Lipschitz condition.” This condition is satisfied by the bounded open sets Ω that we consider because we assume that Ω has the $(m, 2)$ -extension property. There are many approaches to prove the next corollary. However, using spaces of Hölder continuous functions seem to be the approach more suitable to generalization. This is the reason why we have adopted it.

Corollary 11.7.9

Let $\Omega \subset \mathbb{R}^n$ be a bounded open set with the $(m, 2)$ -extension property. If $m > k + n/2$, then the imbedding $i : H^{m,2}(\Omega) \rightarrow C_{bu}^k(\overline{\Omega})$ is compact.

Proof.

i) To prove that $I : H^{m,2}(\Omega) \rightarrow C_{bu}^k(\overline{\Omega})$ is compact when Ω is bounded, we first show that it is enough to prove that the imbedding $i : H^{m-k,2}(\Omega) \rightarrow C_{bu}^0(\overline{\Omega})$ is compact. Suppose that $i : H^{m-k,2}(\Omega) \rightarrow C_{bu}^0(\overline{\Omega})$ is compact and that $\{\alpha_1, \alpha_2, \dots, \alpha_N\}$ is an ordering of the multi-indices α such that $|\alpha| \leq k$. Let $\{f_i\}_{i=1}^\infty$ be a bounded sequence in $H^{m,2}(\Omega)$.

- Then, $\{D^{\alpha_1} f_i\}_{i=1}^\infty$ is a bounded sequence in $H^{m-|\alpha_1|,2}(\Omega) \subset H^{m-k,2}(\Omega)$. Since we assume that $i : H^{m-k,2}(\Omega) \rightarrow C_{bu}^0(\overline{\Omega})$ is compact, there exists a subsequence $\{f_{\alpha_1,i}\}_{i=1}^\infty$ of $\{f_i\}_{i=1}^\infty$ such that $\{D^{\alpha_1} f_{\alpha_1,i}\}_{i=1}^\infty$ converges in $C_{bu}^0(\overline{\Omega})$.
- We now have that $\{D^{\alpha_2} f_{\alpha_1,i}\}_{i=1}^\infty$ is a bounded sequence in $H^{m-|\alpha_1|-|\alpha_2|,2}(\Omega) \subset H^{m-k,2}(\Omega)$. Since we assume that $i : H^{m-k,2}(\Omega) \rightarrow C_{bu}^0(\overline{\Omega})$ is compact, there exists a subsequence $\{f_{\alpha_2,i}\}_{i=1}^\infty$ of $\{f_{\alpha_1,i}\}_{i=1}^\infty$ such that $\{D^{\alpha_2} f_{\alpha_2,i}\}_{i=1}^\infty$ converges in $C_{bu}^0(\overline{\Omega})$.
- and so on until we reach α_N .

The sequence $\{f_{\alpha_N,i}\}_{i=1}^\infty$ is such that $\{D^\alpha f_{\alpha_N,i}\}_{i=1}^\infty$ converges in $C_{bu}^0(\overline{\Omega})$ for all multi-indices α such that $|\alpha| \leq k$. Thus, $\{f_{\alpha_N,i}\}_{i=1}^\infty$ converges in $C_{bu}^k(\overline{\Omega})$.

ii) Choose $j \geq 0$ such that $m - k - j > n/2 \geq m - k - j - 1$ and $0 < \mu < 1$ such that $m - k - j > \mu + n/2 > n/2 \leq m - k - j - 1$. We obviously have that the imbedding $i_1 : H^{m-k,2}(\Omega) \rightarrow H^{m-k-j,2}(\Omega)$ is continuous. We have from the fourth item listed before the statement of the corollary that the imbedding $i_2 : H^{m-k-j,2}(\Omega) \rightarrow C^{k+j,\mu}(\overline{\Omega}) \subset C^{0,\mu}(\overline{\Omega})$ is also continuous. Finally, we have from the third item listed before the statement of the corollary that the imbedding $i_3 : C^{0,\mu}(\overline{\Omega}) \rightarrow C_{bu}^0(\overline{\Omega})$ is compact. It follows that Proposition 4.4.22 that $i = i_3 \circ i_2 \circ i_1 : H^{m-k,2}(\Omega) \rightarrow C_{bu}^0(\overline{\Omega})$ is compact. ■

Theorem 11.7.10 (Rellich's Theorem)

Consider two integers k_1 and k_2 such that $0 \leq k_1 \leq k_2$. Let $\Omega \subset \mathbb{R}^n$ be a bounded open set with the $(k_2, 2)$ -extension property. Then, the inclusion $i : H^{k_2,2}(\Omega) \rightarrow H^{k_1,2}(\Omega)$ is a compact mapping. In fact, every bounded sequence in $H^{k_2,2}(\Omega)$ has a subsequence which converges in $H^{k_1,2}(\Omega)$ for all $k_1 < k_2$.

Proof.

Suppose that $\{f_i\}_{i=1}^\infty$ is a bounded sequence in $H^{k_2,2}(\Omega)$.

Choose $\psi \in \mathcal{D}(\mathbb{R}^n)$ such that $\psi(\mathbf{x}) = 1$ for all $\mathbf{x} \in \Omega$ and $0 \leq \psi(\mathbf{x}) \leq 1$ for all $\mathbf{x} \in \mathbb{R}^n$. If $E : H^{k_2,2}(\Omega) \rightarrow H^{k_2,2}(\mathbb{R}^n)$ is the bounded linear operator satisfying Definition 11.6.1, then

$F = \psi E : H^{k_2,2}(\Omega) \rightarrow H^{k_2,2}(\mathbb{R}^n)$ is a bounded linear mapping such that $F(f)|_{\Omega} = f$ for all $f \in H^{k_2,2}(\Omega)$ and $\text{supp } F(f) \subset K \equiv \text{supp } \psi$ for all $f \in H^{k_2,2}(\Omega)$.

Consider the sequence $\{g_i\}_{i=1}^{\infty}$, where $g_i = F(f_i)$ for all i . We have that $g_i \in H^{k_2,2}(\mathbb{R}^n)$, $\text{supp } g_i \subset K$ for all i , $g_i|_{\Omega} = f_i$ for all i , and there exists $C \in \mathbb{R}$ such that $\|g_i\|_{k_2,\rho} < C$ for all i because the sequence $\{f_i\}_{i=1}^{\infty}$ is bounded and F is a bounded linear mapping.

i) Choose $\phi \in \mathcal{D}(\mathbb{R}^n)$ such that $\phi(\mathbf{x}) = 1$ for all $\mathbf{x} \in K$. For all i , we have $\hat{g}_i = \hat{g}_i * \hat{\phi}$ because $g_i = g_i \phi$. Note that $\hat{g}_i * \hat{\phi}$ is defined by the standard integration formula because $g_i \in H^{k_2,2}(\mathbb{R}^n) \subset L^2(\mathbb{R}^n)$ implies $\hat{g}_i \in L^2(\mathbb{R}^n)$, and $\phi \in \mathcal{S}(\mathbb{R}^n)$ implies $\hat{\phi} \in \mathcal{S}(\mathbb{R}^n) \subset L^2(\mathbb{R}^n)$ by Proposition 6.6.11. Moreover, it follows from item 3 of Proposition 6.6.12 that $\hat{g}_i = \hat{g}_i * \hat{\phi} \in C^{\infty}(\mathbb{R}^n)$ because $\hat{\phi} \in \mathcal{S}(\mathbb{R}^n)$ and $\hat{g}_i \in L^2(\mathbb{R}^n) \subset \mathcal{S}'(\mathbb{R}^n)$.

It follows from Lemma 11.1.21 and Schwarz inequality that

$$\begin{aligned} |\hat{g}_i(\mathbf{y})| &\leq \int_{\mathbb{R}^n} |\hat{\phi}(\mathbf{y} - \mathbf{x})| |\hat{g}_i(\mathbf{x})| d\mathbf{x} \\ &\leq \frac{2^{k_2/2}}{(1 + \|\mathbf{y}\|^2)^{k_2/2}} \int_{\mathbb{R}^n} (1 + \|\mathbf{y} - \mathbf{x}\|^2)^{k_2/2} |\hat{\phi}(\mathbf{y} - \mathbf{x})| (1 + \|\mathbf{x}\|^2)^{k_2/2} |\hat{g}_i(\mathbf{x})| d\mathbf{x} \\ &\leq \frac{2^{k_2/2}}{(1 + \|\mathbf{y}\|^2)^{k_2/2}} \|\phi\|_{k_2,\rho} \|g_i\|_{k_2,\rho} \leq 2^{k_2/2} \|\phi\|_{k_2,\rho} \|g_i\|_{k_2,\rho} \leq 2^{k_2/2} C \|\phi\|_{k_2,\rho} < \infty \end{aligned}$$

for all $\mathbf{y} \in \mathbb{R}^n$ and all i . Similarly, since $D^{e_j} \hat{g}_i = D^{e_j} \hat{g}_i * \hat{\phi}$ for $1 \leq j \leq n$, we get

$$\begin{aligned} \left| \frac{\partial \hat{g}_i}{\partial y_j}(\mathbf{y}) \right| &= \left| \int_{\mathbb{R}^n} \hat{\phi}(\mathbf{y} - \mathbf{x}) \frac{\partial \hat{g}_i}{\partial x_j}(\mathbf{x}) d\mathbf{x} \right| = \left| \int_{\mathbb{R}^n} \frac{\partial \hat{\phi}}{\partial x_j}(\mathbf{y} - \mathbf{x}) \hat{g}_i(\mathbf{x}) d\mathbf{x} \right| \\ &\leq \int_{\mathbb{R}^n} \left| \frac{\partial \hat{\phi}}{\partial x_j}(\mathbf{y} - \mathbf{x}) \right| |\hat{g}_i(\mathbf{x})| d\mathbf{x} \leq \dots \leq 2^{k_2/2} C \left\| \frac{\partial \phi}{\partial y_j} \right\|_{k_2,\rho} < \infty \end{aligned}$$

for all $\mathbf{y} \in \mathbb{R}^n$ and all i , and $1 \leq j \leq n$.

Hence, the sequence of functions $\{\hat{g}_i\}_{i=1}^{\infty}$ is uniformly bounded and equicontinuous on $\mathbb{R}^n$ because

$$\|\nabla_{\mathbf{y}} \hat{g}_i(\mathbf{y})\| \leq \sqrt{n} 2^{k_2/2} C \max_{1 \leq j \leq n} \left\| \frac{\partial \phi}{\partial y_j} \right\|_{k_2,\rho}$$

for $\mathbf{y} \in \mathbb{R}^n$ and $i > 0$, where $\|\cdot\|$ is the Euclidean norm in $\mathbb{R}^n$. The equicontinuity comes from the Mean Value Theorem.

It follows from Arzelà-Ascoli Theorem that there exists a subsequence $\{\hat{g}_i^{[1]}\}_{i=1}^{\infty}$ of $\{\hat{g}_i\}_{i=1}^{\infty}$ that converges uniformly on the compact set of $\overline{B_1(\mathbf{0})}$. Again, using Arzelà-Ascoli Theorem, there exists a subsequence $\{\hat{g}_i^{[2]}\}_{i=1}^{\infty}$ of $\{\hat{g}_i^{[1]}\}_{i=1}^{\infty}$ that converges uniformly on the compact set of $\overline{B_2(\mathbf{0})}$. In general, having the $\{\hat{g}_i^{[N]}\}_{i=1}^{\infty}$ that converges uniformly on the compact set of $\overline{B_N(\mathbf{0})}$ for $N \in \mathbb{N}$, we use Arzelà-Ascoli Theorem, to obtain a subsequence $\{\hat{g}_i^{[N+1]}\}_{i=1}^{\infty}$ of $\{\hat{g}_i^{[N]}\}_{i=1}^{\infty}$ that converges uniformly on the compact set of $\overline{B_{N+1}(\mathbf{0})}$. The sequence $\{\hat{g}_i^{[i]}\}_{i=1}^{\infty}$ is a subsequence of $\{\hat{g}_i\}_{i=1}^{\infty}$ that converges uniformly on all sets $\overline{B_N(\mathbf{0})}$ for $N > 0$.

ii) To prove that the inclusion $i : H^{k_2,2}(\Omega) \rightarrow H^{k_1,2}(\Omega)$ is a compact mapping, it suffices to prove that $\{g_i^{[i]}\}_{i=1}^\infty$ converges in $H^{k_1,2}(\mathbb{R}^n)$ since $g_i^{[i]}|_\Omega = f_i^{[i]}$ for all i .

For $N > 0$, we have

$$\begin{aligned} \|g_i^{[i]} - g_j^{[j]}\|_{k_1, \rho}^2 &= \int_{\|\mathbf{y}\| \leq N} (1 + \|\mathbf{y}\|^2)^{k_1} |\hat{g}_i^{[i]}(\mathbf{y}) - \hat{g}_j^{[j]}(\mathbf{y})|^2 d\mathbf{y} \\ &\quad + \int_{\|\mathbf{y}\| > N} (1 + \|\mathbf{y}\|^2)^{k_1-k_2} (1 + \|\mathbf{y}\|^2)^{k_2} |\hat{g}_i^{[i]}(\mathbf{y}) - \hat{g}_j^{[j]}(\mathbf{y})|^2 d\mathbf{y} \\ &\leq \sup_{\|\mathbf{x}\| \leq N} |\hat{g}_i^{[i]}(\mathbf{x}) - \hat{g}_j^{[j]}(\mathbf{x})|^2 \int_{\|\mathbf{y}\| \leq N} (1 + \|\mathbf{y}\|^2)^{k_1} d\mathbf{y} \\ &\quad + (1 + N^2)^{k_1-k_2} \int_{\|\mathbf{y}\| > N} (1 + \|\mathbf{y}\|^2)^{k_2} |\hat{g}_i^{[i]}(\mathbf{y}) - \hat{g}_j^{[j]}(\mathbf{y})|^2 d\mathbf{y} \\ &\leq \sup_{\|\mathbf{x}\| \leq N} |\hat{g}_i^{[i]}(\mathbf{x}) - \hat{g}_j^{[j]}(\mathbf{x})|^2 \int_{\|\mathbf{y}\| \leq N} (1 + \|\mathbf{y}\|^2)^{k_1} d\mathbf{y} + (1 + N^2)^{k_1-k_2} \|g_i^{[i]} - g_j^{[j]}\|_{k_2, \rho}^2 \\ &\leq \sup_{\|\mathbf{x}\| \leq N} |\hat{g}_i^{[i]}(\mathbf{x}) - \hat{g}_j^{[j]}(\mathbf{x})|^2 \int_{\|\mathbf{y}\| \leq N} (1 + \|\mathbf{y}\|^2)^{k_1} d\mathbf{y} + 2C(1 + N^2)^{k_1-k_2}. \end{aligned}$$

Given $\epsilon > 0$, there exists $N > 0$ such that $2C(1 + N^2)^{k_1-k_2} < \epsilon/2$ because $k_1 - k_2 < 0$. For this N , since $\{\hat{g}_i^{[i]}\}_{i=1}^\infty$ converges uniformly on the compact set $\overline{B_N(\mathbf{0})}$, there exists $J > 0$ such that

$$\sup_{\|\mathbf{x}\| \leq N} |\hat{g}_i^{[i]}(\mathbf{x}) - \hat{g}_j^{[j]}(\mathbf{x})|^2 \int_{\|\mathbf{y}\| \leq N} (1 + \|\mathbf{y}\|^2)^{k_1} d\mathbf{y} < \frac{\epsilon}{2}$$

for $j, k > J$. Thus, $\|g_i^{[i]} - g_j^{[j]}\|_{k_1, \rho}^2 < \epsilon$ for $j, k > J$. This prove that $\{g_i^{[i]}\}_{i=1}^\infty$ is a Cauchy sequence in $H^{k_1,2}(\mathbb{R}^n)$. Since $H^{k_1,2}(\mathbb{R}^n)$ is a Banach space, this series converges in $H^{k_1,2}(\mathbb{R}^n)$. $\blacksquare$

Theorem 11.7.11 (Poincaré's Inequality)

Let $\Omega \subset \mathbb{R}^n$ be an open subset contained in the region $R = \{\mathbf{x} \in \mathbb{R}^n : a < x_n < b\}$. Let

$$\|f\|_{k,2,*} = \left(\sum_{|\alpha|=k} \|D^\alpha f\|_{2,\Omega}^2 \right)^{1/2} \quad (11.7.3)$$

for $f \in H_0^{k,2}(\Omega)$. Then, there exists a constant C such that $\|f\|_{k,2,\Omega} \leq C\|f\|_{k,2,*}$ for all $f \in H_0^{k,2}(\Omega)$.

Remark 11.7.12

It follows from Poincaré's Inequality that, if $\Omega \subset \mathbb{R}^n$ is an open subset contained between two parallel hyperplanes, then $\|\cdot\|_{k,2,*}$ is a norm on $H_0^{k,2}(\Omega)$ equivalent to the norm $\|\cdot\|_{k,2,\Omega}$. $\spadesuit$

Proof.

We only need to prove that there exists a constant C such that

$$\|D^\alpha f\|_{2,\Omega} \leq C\|f\|_{k,2,*} \quad (11.7.4)$$

for $f \in H_0^{k,2}(\Omega)$ and $|\alpha| < k$.

Recall that the D^α are derivatives in the sense of distributions. Moreover, it is enough to prove (11.7.4) for $f \in \mathcal{D}(\Omega)$ since $\mathcal{D}(\Omega)$ is dense in $H_0^{1,2}(\Omega)$ by definition. We have that $\|D^\alpha g\|_{2,\Omega} \leq \|g\|_{k,2}$ and $\|g\|_{k,2,*} \leq \|g\|_{k,2,\Omega}$ for $g \in H_0^{k,2}(\Omega)$ and $0 \leq |\alpha| < k$. Hence, if $\{\phi_i\}_{i=1}^\infty$ converges to $f \in H_0^{k,2}(\Omega)$ according to the norm $\|\cdot\|_{k,2,\Omega}$, then $\|D^\alpha \phi_i\|_{2,\Omega} \rightarrow \|D^\alpha f\|_{2,\Omega}$ and $\|\phi_i\|_{k,2,*} \rightarrow \|f\|_{k,2,*}$ as $i \rightarrow \infty$. So, $\|D^\alpha \phi_i\|_{2,\Omega} \leq C\|\phi_i\|_{k,2,*}$ for all i implies that $\|D^\alpha f\|_{2,\Omega} \leq C\|f\|_{k,2,*}$.

The proof of (11.7.4) is by induction on k . Consider $k = 1$.

We write $\mathbf{x} \in \mathbb{R}^n$ as $\mathbf{x} = (\check{\mathbf{x}}, x_n)$ with $\check{\mathbf{x}} \in \mathbb{R}^{n-1}$ and $x_n \in \mathbb{R}$. Since $\phi(\mathbf{x}) = \int_a^{x_n} \frac{\partial \phi}{\partial x_n}(\check{\mathbf{x}}, t) dt$ for $\phi \in \mathcal{D}(\Omega)$, we get from Schwarz inequality that

$$\begin{aligned} |\phi(\mathbf{x})|^2 &= \left(\int_a^{x_n} \frac{\partial \phi}{\partial x_n}(\mathbf{x}', t) dt \right)^2 \leq \left(\int_a^{x_n} \left| \frac{\partial \phi}{\partial x_n}(\mathbf{x}', t) \right|^2 dt \right) \left(\int_a^{x_n} dt \right) \\ &\leq (b-a) \int_a^{x_n} \left| \frac{\partial \phi}{\partial x_n}(\mathbf{x}', t) \right|^2 dt \end{aligned}$$

for $\mathbf{x} \in \Omega$ and $\phi \in \mathcal{D}(\Omega)$. Hence,

$$\begin{aligned} \|\phi\|_{2,\Omega}^2 &= \int_\Omega |\phi(\mathbf{x})|^2 d\mathbf{x} = \int_R |\phi(\mathbf{x})|^2 d\mathbf{x} \leq (b-a) \int_a^b \int_{\mathbb{R}^{n-1}} \left(\int_a^{x_n} \left| \frac{\partial \phi}{\partial x_n}(\check{\mathbf{x}}, t) \right|^2 dt \right) d\check{\mathbf{x}} dx_n \\ &\leq (b-a) \int_a^b \left(\int_{\mathbb{R}^{n-1}} \int_a^b \left| \frac{\partial \phi}{\partial x_n}(\check{\mathbf{x}}, t) \right|^2 dt d\check{\mathbf{x}} \right) dx_n = (b-a)^2 \left\| \frac{\partial \phi}{\partial x_n} \right\|_{2,\Omega}^2 \leq (b-a)^2 \|f\|_{1,2,*}^2 \end{aligned}$$

for $\phi \in \mathcal{D}(\Omega)$. As usual $\phi \in \mathcal{D}(\Omega)$ is extended to R by $\phi(\mathbf{x}) = 0$ for $\mathbf{x} \in R \setminus \Omega$. This proves (11.7.4) for $k = 1$ with $C = b - a$.

We assume that (11.7.4) is true for $k = j$ and consider the claim for $k = j + 1$. Note that $H_0^{j+1,2}(\Omega) \subset H_0^{j,2}(\Omega)$. By induction, there exists a constant C_1 such that $\|D^\alpha \phi\|_{2,\Omega} \leq C_1 \|\phi\|_{j,2,*}$ for all $\phi \in \mathcal{D}(\Omega)$ and $|\alpha| < j$. So we only have to prove that there exists a constant C_2 such that $\|D^\alpha \phi\|_{2,\Omega} \leq C_2 \|\phi\|_{j+1,2,*}$ for $|\alpha| = j$ and $\phi \in \mathcal{D}(\Omega)$.

The reasoning for $k = 1$ applied to $D^\alpha \phi$ instead of ϕ shows that $\|D^\alpha \phi\|_{2,\Omega} \leq (b-a) \|D^\alpha \phi\|_{1,2,*}$ for $|\alpha| = j$ and $\phi \in \mathcal{D}(\Omega)$. But

$$\|D^\alpha \phi\|_{1,2,*}^2 = \sum_{j=1}^n \|D^{e_j} D^\alpha \phi\|_{2,\Omega}^2 \leq \sum_{|\alpha|=j+1} \|D^\alpha \phi\|_{2,\Omega}^2 = \|\phi\|_{j+1,2,*}^2$$

for all $\phi \in \mathcal{D}(\Omega)$. Thus, with $C_2 = b - a$, we have $\|D^\alpha \phi\|_{2,\Omega} \leq C_2 \|\phi\|_{j+1,2,*}$ for $|\alpha| = j$ and $\phi \in \mathcal{D}(\Omega)$. This proves (11.7.4) for $k = j + 1$ with $C = C_1 C_2$, and complete the proof by induction. $\blacksquare$

11.8 Trace Operators

The trace theorems allow us to generalize boundary conditions to functions defined almost everywhere. We will make an extensive use of the fractional Sobolev spaces $H^{s,2}(\mathbb{R}^n)$ to define the trace operators.

A detailed study of trace operators is done in [20].

Before studying the general trace operator, we consider a simple trace theorem for $u \in H^{1,2}(\Omega)$, where Ω is a bounded open subset of $\mathbb{R}^n$ with a nice boundary.

Proposition 11.8.1

Suppose that Ω is a bounded open subset of $\mathbb{R}^n$ and that $\partial\Omega$ is a manifold of class C^2 . There exists $C > 0$ such that

$$\|u\|_{2,\partial\Omega} = \left(\int_{\partial\Omega} |u(\mathbf{x})|^2 dS_{\mathbf{x}} \right)^{1/2} \leq C \|u\|_{1,2,\Omega}$$

for all $u \in C^1(\overline{\Omega})$.

Proof.

Let $\nu(\mathbf{z})$ be the outward unit normal to $\partial\Omega$ at $\mathbf{z} \in \partial\Omega$ and N_η be a tubular neighbourhood of $\partial\Omega$; in particular, $N_\eta = \{\mathbf{x} + t\nu(\mathbf{x}) : \mathbf{x} \in \partial\Omega, |t| < \eta\}$.

Choose $\phi \in \mathcal{D}(\mathbb{R})$ such that $|\phi| \leq 1$, $\phi(0) = 1$ and $\text{supp } \phi \subset]-\eta, \eta[$. We extend $\nu : \partial\Omega \rightarrow \mathbb{R}^n$ to $\mathbb{R}^n$ as follows.

$$\nu(\mathbf{z}) = \begin{cases} \phi(t)\nu(\mathbf{x}) & \text{if } \mathbf{z} = \mathbf{x} + t\nu(\mathbf{x}) \text{ with } \mathbf{x} \in \partial\Omega \text{ and } |t| < \eta. \\ 0 & \text{if } \mathbf{z} \notin N_\eta \end{cases}$$

Note that $\nu \in C^1(\mathbb{R}^n)$ has a compact support inside N_η . Using the divergence theorem, we find that

$$\begin{aligned} \left| \int_{\partial\Omega} |u(\mathbf{x})|^2 dS_{\mathbf{x}} \right| &= \left| \int_{\partial\Omega} (|u(\mathbf{x})|^2 \nu(\mathbf{x})) \cdot \nu(\mathbf{x}) dS_{\mathbf{x}} \right| = \left| \int_{\Omega} \text{div}(|u(\mathbf{x})|^2 \nu(\mathbf{x})) d\mathbf{x} \right| \\ &= \left| \int_{\Omega} \sum_{j=1}^n \left(u(\mathbf{x}) \frac{\partial u}{\partial x_j}(\mathbf{x}) \nu_j(\mathbf{x}) + \overline{u}(\mathbf{x}) \frac{\partial u}{\partial x_j}(\mathbf{x}) \nu_j(\mathbf{x}) + |u(\mathbf{x})|^2 \frac{\partial \nu_j}{\partial x_j}(\mathbf{x}) \right) d\mathbf{x} \right| \\ &\leq C_1 \int_{\Omega} \left(2 \sum_{j=1}^n \left| u(\mathbf{x}) \frac{\partial u}{\partial x_j}(\mathbf{x}) \right| + n |u(\mathbf{x})|^2 \right) d\mathbf{x}, \end{aligned}$$

where

$$C_1 = \max \left\{ \max_{\substack{\mathbf{x} \in N_\eta \\ 1 \leq j \leq n}} |\nu_j(\mathbf{x})|, \max_{\substack{\mathbf{x} \in N_\eta \\ 1 \leq j \leq n}} \left| \frac{\partial \nu_j}{\partial x_j}(\mathbf{x}) \right| \right\}.$$

Using Schwarz inequality, we get

$$\begin{aligned} \|u\|_{2,\Omega}^2 &= \left| \int_{\partial\Omega} |u(\mathbf{x})|^2 dS_{\mathbf{x}} \right| \leq C_1 \left(2 \sum_{j=1}^n \int_{\Omega} \left| u(\mathbf{x}) \frac{\partial u}{\partial x_j}(\mathbf{x}) \right| d\mathbf{x} + \int_{\Omega} n |u(\mathbf{x})|^2 d\mathbf{x} \right) \\ &\leq C_1 \left(2 \sum_{j=1}^n \left(\int_{\Omega} |u(\mathbf{x})|^2 d\mathbf{x} \right)^{1/2} \left(\int_{\Omega} \left| \frac{\partial u}{\partial x_j}(\mathbf{x}) \right|^2 d\mathbf{x} \right)^{1/2} + \int_{\Omega} n |u(\mathbf{x})|^2 d\mathbf{x} \right) \end{aligned}$$

$$\begin{aligned}
&\leq C_1 \left(\sum_{j=1}^n \left(\int_{\Omega} |u(\mathbf{x})|^2 d\mathbf{x} + \int_{\Omega} \left| \frac{\partial u}{\partial x_j}(\mathbf{x}) \right|^2 d\mathbf{x} \right) + \int_{\Omega} n |u(\mathbf{x})|^2 d\mathbf{x} \right) \\
&= 2C_1(n+1) \int_{\Omega} |u(\mathbf{x})|^2 d\mathbf{x} + C_1 \sum_{j=1}^n \int_{\Omega} \left| \frac{\partial u}{\partial x_j}(\mathbf{x}) \right|^2 d\mathbf{x} \leq C^2 \|u\|_{1,2,\Omega}^2
\end{aligned}$$

with $C^2 = 2(n+1)C_1$. Note that we have used the inequality $2ab \leq a^2 + b^2$ to get the third inequality above. ■

Corollary 11.8.2 (Elementary trace operator)

Suppose that Ω is a bounded open subset of $\mathbb{R}^n$ and that $\partial\Omega$ is a manifold of class C^2 . The restriction map

$$\begin{aligned}
T : C^1(\overline{\Omega}) &\rightarrow C^1(\partial\Omega) \\
u &\mapsto u|_{\partial\Omega}
\end{aligned}$$

can be continuously extended to a map $T : H^{1,2}(\Omega) \rightarrow L^2(\partial\Omega)$.

Proof.

Since

$$\|u\|_{2,\partial\Omega} \leq C \|u\|_{1,2,\Omega} \quad (11.8.1)$$

for all $u \in C^1(\overline{\Omega})$, and $C^1(\overline{\Omega}) \supset C^\infty(\overline{\Omega})$ is dense in $H^{1,2}(\Omega)$ with respect to the norm $\|\cdot\|_{1,2,\Omega}$, we may extend the restriction map T to $H^{1,2}(\Omega)$ as follows.

Given $u \in H^{1,2}(\Omega)$, choose $\{u_j\}_{j=0}^\infty \subset C^1(\overline{\Omega})$ such that $\|u_j - u\|_{1,2,\Omega} \rightarrow 0$ as $j \rightarrow \infty$. From (11.8.1), $\{u_j|_{\partial\Omega}\}_{j=0}^\infty$ is a Cauchy sequence in $L^2(\partial\Omega)$. Thus there exists $f \in L^2(\partial\Omega)$ such that $\{u_j|_{\partial\Omega}\}_{j=0}^\infty$ converges to f in $L^2(\partial\Omega)$. We set $T(u) = f$.

The procedure is independent of the chosen sequence $\{u_j\}_{j=0}^\infty \subset C^1(\overline{\Omega})$. If $\{v_j\}_{j=0}^\infty \subset C^1(\overline{\Omega})$ is another sequence such that $\|v_j - u\|_{1,2,\Omega} \rightarrow 0$ as $j \rightarrow \infty$ and $g \in L^2(\partial\Omega)$ is the limit of $\{v_j|_{\partial\Omega}\}_{j=0}^\infty$ in $L^2(\partial\Omega)$, then (11.8.1) yields

$$\|u_j - v_j\|_{2,\partial\Omega} \leq C \|u_j - v_j\|_{1,2,\Omega} \rightarrow 0 \quad \text{as } j \rightarrow \infty.$$

Hence, $f = g$ almost everywhere. ■

The first domain with a boundary that we consider for the general space $H^k(\Omega)$ is $\Omega = \mathbb{R}_+^n$. Its boundary is the hyperplane $\mathbb{R}_0^n \equiv \{\mathbf{x} \in \mathbb{R}^n : x_n = 0\} \cong \mathbb{R}^{n-1}$.

Proposition 11.8.3

There exists a bounded linear mapping $T : H^k(\mathbb{R}^n) \rightarrow H^{k-1/2}(\mathbb{R}^{n-1})$ such that $T(\phi)(\mathbf{z}) = \phi(\mathbf{z}, 0)$ for all $\phi \in \mathcal{S}(\mathbb{R}^n)$ and all $\mathbf{z} \in \mathbb{R}^{n-1}$. The operator T is called a **trace operator**,

Moreover, if $f \in H^k(\mathbb{R}^n)$ and $W \subset \mathbb{R}^n$ is an open set such that $\text{supp } f \subset W$, then we may assume that $\text{supp } T(f) \subset W \cap \mathbb{R}_0^n$.

Remark 11.8.4

Obviously, when we write $\text{supp } T(f) \subset W \cap \mathbb{R}_0^n$, we view $W \cap \mathbb{R}_0^n$ has a subset of $\mathbb{R}^{n-1}$. To be more precise, we should have written that $\text{supp } T(f) \subset \{z \in \mathbb{R}^{n-1} : (z, 0) \in W\}$. However, we would have lost the visual aspect of the statement; namely, that $\text{supp } T(f)$ is "in" $\partial\mathbb{R}_+^n = \mathbb{R}_0^n$. We will keep using this abuse of language in the rest of this section. ♠

Proof.

In this proof, we sometime write $\mathbf{x} \in \mathbb{R}^n$ as $\mathbf{x} = (\check{\mathbf{x}}, x)$ for $\check{\mathbf{x}} \in \mathbb{R}^{n-1}$ and $x \in \mathbb{R}$. Moreover, $\phi_{\check{\mathbf{x}}}(x) = \phi(\check{\mathbf{x}}, x) = \phi(\mathbf{x})$ for $\phi \in \mathcal{S}(\mathbb{R}^n)$. We use a similar notation for $\mathbf{y} \in \mathbb{R}^n$.

Given $\phi \in \mathcal{S}(\mathbb{R}^n)$, we have that

$$(\phi_{\check{\mathbf{x}}})^\wedge(y) = \int_{-\infty}^{\infty} \phi(\check{\mathbf{x}}, x) e^{-ixy} d\mu(x)$$

for $y \in \mathbb{R}$. Since $\phi_{\check{\mathbf{x}}} \in \mathcal{S}(\mathbb{R}) \subset L^1(\mathbb{R})$ and $(\phi_{\check{\mathbf{x}}})^\wedge \in \mathcal{S}(\mathbb{R}) \in L^1(\mathbb{R})$ by Proposition 6.6.11, it follows from Proposition 6.6.5 that the inverse Fourier transform of $(\phi_{\check{\mathbf{x}}})^\wedge$ is $\phi_{\check{\mathbf{x}}}$. If we assume that $T(\phi)$ for $\phi \in \mathcal{S}(\mathbb{R}^n)$ satisfies $T(\phi)(z) = \phi(z, 0)$ for all $z \in \mathbb{R}^{n-1}$, we have that

$$g(z) \equiv T(\phi)(z) = \phi_z(0) = \mathcal{F}^{-1}((\phi_z)^\wedge)(0) = \int_{-\infty}^{\infty} (\phi_z)^\wedge(y) d\mu(y)$$

for $z \in \mathbb{R}^{n-1}$. Moreover,

$$\begin{aligned} \hat{g}(\mathbf{w}) &= \int_{\mathbb{R}^{n-1}} \left(\int_{-\infty}^{\infty} (\phi_z)^\wedge(y) d\mu(y) \right) e^{-i\mathbf{w} \cdot \mathbf{z}} d\mu(z) = \int_{-\infty}^{\infty} \left(\int_{\mathbb{R}^{n-1}} (\phi_z)^\wedge(y) e^{-i\mathbf{w} \cdot \mathbf{z}} d\mu(z) \right) d\mu(y) \\ &= \int_{-\infty}^{\infty} \left(\int_{\mathbb{R}^{n-1}} \left(\int_{-\infty}^{\infty} \phi(z, x) e^{-ixy} d\mu(x) \right) e^{-i\mathbf{w} \cdot \mathbf{z}} d\mu(z) \right) d\mu(y) = \int_{-\infty}^{\infty} \hat{\phi}(\mathbf{w}, y) d\mu(y) \end{aligned}$$

for all $\mathbf{w} \in \mathbb{R}^{n-1}$. Hence, using Schwarz inequality, we get

$$\begin{aligned} \|g\|_{k-1/2, \rho}^2 &= \int_{\mathbb{R}^{n-1}} |\hat{g}(\mathbf{w})|^2 (1 + \|\mathbf{w}\|^2)^{k-1/2} d\mathbf{w} = \int_{\mathbb{R}^{n-1}} \left| \int_{-\infty}^{\infty} \hat{\phi}(\mathbf{w}, y) d\mu(y) \right|^2 (1 + \|\mathbf{w}\|^2)^{k-1/2} d\mathbf{w} \\ &= \int_{\mathbb{R}^{n-1}} \left| \int_{-\infty}^{\infty} \hat{\phi}(\mathbf{w}, y) (1 + \|\mathbf{w}\|^2 + y^2)^{k/2} (1 + \|\mathbf{w}\|^2 + y^2)^{-k/2} d\mu(y) \right|^2 (1 + \|\mathbf{w}\|^2)^{k-1/2} d\mathbf{w} \\ &\leq \int_{\mathbb{R}^{n-1}} \left[\left(\int_{-\infty}^{\infty} |\hat{\phi}(\mathbf{w}, y)|^2 (1 + \|\mathbf{w}\|^2 + y^2)^k d\mu(y) \right) \left(\int_{-\infty}^{\infty} (1 + \|\mathbf{y}\|^2 + y^2)^{-k} d\mu(y) \right) \right. \\ &\quad \left. (1 + \|\mathbf{w}\|^2)^{k-1/2} \right] d\mathbf{w} \end{aligned}$$

However, with the substitution $y = \sqrt{1 + \|\mathbf{w}\|^2} z$, we get

$$\int_{-\infty}^{\infty} (1 + \|\mathbf{w}\|^2 + y^2)^{-k} d\mu(y) = \frac{1}{(1 + \|\mathbf{w}\|^2)^{k-1/2}} \int_{-\infty}^{\infty} (1 + z^2)^{-k} d\mu(z) .$$

Therefore,

$$\begin{aligned} \|g\|_{k-1/2,\rho}^2 &\leq \int_{\mathbb{R}^{n-1}} \left[\left(\int_{-\infty}^{\infty} |\hat{\phi}(\mathbf{w}, y)|^2 (1 + \|\mathbf{w}\|^2 + y^2)^k d\mu(y) \right) \right. \\ &\quad \left. \left(\frac{1}{(1 + \|\mathbf{w}\|^2)^{k-1/2}} \int_{-\infty}^{\infty} (1 + z^2)^{-k} d\mu(z) \right) (1 + \|\mathbf{w}\|^2)^{k-1/2} \right] d\mathbf{w} \\ &= \int_{-\infty}^{\infty} (1 + z^2)^{-k} d\mu(z) \int_{\mathbb{R}^{n-1}} \int_{-\infty}^{\infty} |\hat{\phi}(\mathbf{w}, y)|^2 (1 + \|\mathbf{w}\|^2 + y^2)^k d\mu(y) d\mathbf{w} = C \|\phi\|_{k,\rho}^2 \end{aligned}$$

with $C = (2\pi)^{-1/2} \int_{-\infty}^{\infty} (1 + z^2)^{-k} d\mu(z) < \infty$.

We have shown that $\|T(\phi)\|_{k-1/2,\rho} \leq \sqrt{C} \|\phi\|_{k,\rho}$ for all $\phi \in \mathcal{S}(\mathbb{R}^n)$. Since $\mathcal{S}(\mathbb{R}^n)$ is dense in $H^{k,2}(\mathbb{R}^n)$, we can extend T to $H^{k,2}(\mathbb{R}^n)$ and still have $\|T(f)\|_{k-1/2,\rho} \leq \sqrt{C} \|f\|_{k,\rho}$ for all $f \in H^{k,2}(\mathbb{R}^n)$.

For the last statement of the theorem, choose $\phi \in \mathcal{D}(\mathbb{R}^n)$ such that $\phi(\mathbf{x}) = 1$ for $\mathbf{x} \in \text{supp } f$ and $\phi(\mathbf{x}) = 0$ for $\mathbf{x} \in \mathbb{R}^n \setminus W$. Then $f = \phi f$ and $T(f) = T(\phi)T(f)$ ¹⁶. Since $T(\phi)(\mathbf{z}) = \phi(\mathbf{z}, 0)$, we have that $T(f)(\mathbf{z}) = 0$ for $\mathbf{z} \notin \{\mathbf{z} \in \mathbb{R}^{n-1} : (\mathbf{z}, 0) \in W\}$. ■

Theorem 11.8.5

Let T be the trace operator defined in the previous proposition and $E : H^k(\mathbb{R}_+^n) \rightarrow H^k(\mathbb{R}^n)$ be a strong k -extension operator. Such an operator exists according to Theorem 11.6.3. Then, $T_0 = T \circ E : H^k(\mathbb{R}_+^n) \rightarrow H^{k-1/2}(\mathbb{R}^{n-1})$ is continuous and onto. Moreover, if $f = \phi|_{\mathbb{R}_+^n}$ for some $\phi \in \mathcal{S}(\mathbb{R}^n)$, then $T_0(f)(\mathbf{z}) = f(\mathbf{z}, 0)$ for $\mathbf{z} \in \mathbb{R}^{n-1}$. The operator T_0 is also called a **trace operator**,

Proof.

Since T and E are continuous, we obviously have that the composition $T_0 = T \circ E$ is continuous.

To prove that T_0 is onto, we prove that there exists a right inverse to T_0 ; namely a continuous mapping $S_0 : H^{k-1/2}(\mathbb{R}^{n-1}) \rightarrow H^k(\mathbb{R}_+^n)$. Since $\mathcal{S}(\mathbb{R}^{n-1})$ is dense in $H^{k-1/2}(\mathbb{R}^{n-1})$, we only have to define the mapping S_0 on $\mathcal{S}(\mathbb{R}^{n-1})$.

We consider the mapping $S : \mathcal{S}(\mathbb{R}^{n-1}) \rightarrow \mathcal{S}(\mathbb{R}^n)$ defined by

$$S(\phi)(\mathbf{x}) = \mathcal{F}^{-1} \left(\frac{(1 + \|\check{\mathbf{y}}\|^2)^{k-1/2}}{C(1 + \|\mathbf{y}\|^2)^k} \mathcal{F}(\phi)(\check{\mathbf{y}}) \right)(\mathbf{x})$$

for $\phi \in \mathcal{S}(\mathbb{R}^{n-1})$ and $\mathbf{x} \in \mathbb{R}^n$, where $C = (2\pi)^{-1/2} \int_{\mathbb{R}} (1 + x^2)^{-k} dx$ and $\mathbf{y} = (\check{\mathbf{y}}, y_n) \in \mathbb{R}^n$. Since the Fourier transform is a bijection from the space of rapidly decreasing function to itself, we have that all the functions involved in the definition of S are rapidly decreasing functions.

¹⁶We have by construction that $T(\phi\phi_i) = T(\phi)T(\phi_i)$ for $\phi \in \mathcal{D}(\mathbb{R}^n)$ and $\phi_i \in \mathcal{S}(\mathbb{R}^n)$. If $\phi_i \rightarrow f$ in $H^k(\mathbb{R}^n)$ as $i \rightarrow \infty$, then $\phi\phi_i \rightarrow \phi f$ in $H^k(\mathbb{R}^n)$ according to Proposition 11.1.22, and $T(\phi f) = \lim_{i \rightarrow \infty} T(\phi\phi_i) = \lim_{i \rightarrow \infty} T(\phi)T(\phi_i) = T(\phi)T(f)$.

We prove that S is a continuous mapping from $H^{k-1/2}(\mathbb{R}^{n-1})$ to $H^k(\mathbb{R}^n)$. We have that

$$\begin{aligned}
\|S(\phi)\|_{k,\rho}^2 &= \int_{\mathbb{R}^n} |\mathcal{F}(S(\phi))(\mathbf{y})|^2 (1 + \|\mathbf{y}\|^2)^k d\mathbf{y} \\
&= \int_{\mathbb{R}^n} \left| \frac{(1 + \|\check{\mathbf{y}}\|^2)^{k-1/2}}{C(1 + \|\mathbf{y}\|^2)^k} \mathcal{F}(\phi)(\check{\mathbf{y}}) \right|^2 (1 + \|\mathbf{y}\|^2)^k d\mathbf{y} \\
&= \int_{\mathbb{R}^{n-1}} \frac{1}{C^2} \left(\int_{\mathbb{R}} \frac{1}{(1 + \|\mathbf{y}\|^2)^k} dy_n \right) (1 + \|\check{\mathbf{y}}\|^2)^{2k-1} |\mathcal{F}(\phi)(\check{\mathbf{y}})|^2 d\check{\mathbf{y}} \\
&= \frac{(2\pi)^{1/2}}{C} \int_{\mathbb{R}^{n-1}} (1 + \|\check{\mathbf{y}}\|^2)^{k-1/2} |\mathcal{F}(\phi)(\check{\mathbf{y}})|^2 d\check{\mathbf{y}} = \frac{(2\pi)^{1/2}}{C} \|\phi\|_{k-1/2,\rho}^2
\end{aligned}$$

because

$$\begin{aligned}
\int_{\mathbb{R}} \frac{1}{(1 + \|\mathbf{y}\|^2)^k} dy_n &= \int_{\mathbb{R}} \frac{1}{(1 + \|\check{\mathbf{y}}\|^2 + y_n^2)^k} dy_n = \frac{1}{(1 + \|\check{\mathbf{y}}\|^2)^{k-1/2}} \int_{\mathbb{R}} \frac{1}{(1 + u^2)^k} du \\
&= \frac{C(2\pi)^{1/2}}{(1 + \|\check{\mathbf{y}}\|^2)^{k-1/2}}
\end{aligned} \tag{11.8.2}$$

after the substitution $y_n = (1 + \|\check{\mathbf{y}}\|^2)^{1/2} u$.

We have shown that $\|S(\phi)\|_{k,\rho} = (2\pi)^{1/4} C^{-1/2} \|\phi\|_{k-1/2,\rho}$ for all $\phi \in \mathcal{S}(\mathbb{R}^{n-1})$. Since $\mathcal{S}(\mathbb{R}^{n-1})$ is dense in $H^{k-1/2}(\mathbb{R}^{n-1})$. We can extend S to $H^{k-1/2}(\mathbb{R}^{n-1})$ and still have $\|S(f)\|_{k,\rho} = (2\pi)^{1/4} C^{-1/2} \|f\|_{k-1/2,\rho}$ for all $f \in H^{k-1/2}(\mathbb{R}^{n-1})$.

The operator S_0 is defined by $S_0 = i \circ S$ where $i : H^k(\mathbb{R}^n) \rightarrow H^k(\mathbb{R}_+^n)$ is the continuous mapping defined by $i(f) = f|_{\mathbb{R}_+^n}$ for all $f \in H^k(\mathbb{R}^n)$. We obviously have that $\|i(f)\|_{k,2,\mathbb{R}_+^n} = \|f\|_{k,2,\mathbb{R}_+^n} \leq \|f\|_{k,2,\mathbb{R}^n}$. Because of the equivalence between the norms $\|\cdot\|_{k,2,\mathbb{R}^n}$ and $\|\cdot\|_{k,\rho}$, there exists a constant Q such that $\|S_0(f)\|_{k,2,\mathbb{R}_+^n} \leq Q \|f\|_{k-1/2,\rho}$ for all $f \in H^{k-1/2}(\mathbb{R}^{n-1})$.

The only piece missing to complete the proof of the theorem is to show that S_0 is really the right inverse of T_0 . Using (11.8.2), we get that

$$\begin{aligned}
S(\phi)(\mathbf{z}, 0) &= \mathcal{F}^{-1} \left(\frac{(1 + \|\check{\mathbf{y}}\|^2)^{k-1/2}}{C(1 + \|\mathbf{y}\|^2)^k} \mathcal{F}(\phi)(\check{\mathbf{y}}) \right) (\mathbf{z}, 0) \\
&= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^{n-1}} \int_{\mathbb{R}} \frac{(1 + \|\check{\mathbf{y}}\|^2)^{k-1/2}}{C(1 + \|\mathbf{y}\|^2)^k} \mathcal{F}(\phi)(\check{\mathbf{y}}) e^{i\check{\mathbf{y}} \cdot \mathbf{z}} dy_n d\check{\mathbf{y}} \\
&= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^{n-1}} \left(\int_{\mathbb{R}} \frac{1}{(1 + \|\mathbf{y}\|^2)^k} dy_n \right) \frac{(1 + \|\check{\mathbf{y}}\|^2)^{k-1/2}}{C} \mathcal{F}(\phi)(\check{\mathbf{y}}) e^{i\check{\mathbf{y}} \cdot \mathbf{z}} d\check{\mathbf{y}} \\
&= \frac{1}{(2\pi)^{(n-1)/2}} \int_{\mathbb{R}^{n-1}} \mathcal{F}(\phi)(\check{\mathbf{y}}) e^{i\check{\mathbf{y}} \cdot \mathbf{z}} d\check{\mathbf{y}} = \mathcal{F}^{-1}(\mathcal{F}(\phi))(\mathbf{z}) = \phi(\mathbf{z})
\end{aligned}$$

for all $\mathbf{z} \in \mathbb{R}^{n-1}$. Therefore, $T_0(S_0(\phi))(\mathbf{z}) = T(S(\phi))(\mathbf{z}) = S(\phi)(\mathbf{z}, 0) = \phi(\mathbf{z})$ for all $\mathbf{z} \in \mathbb{R}^{n-1}$ and all $\phi \in \mathcal{S}(\mathbb{R}^{n-1})$. Hence, $T_0(S_0(\phi)) = \phi$ for all $\phi \in \mathcal{S}(\mathbb{R}^{n-1})$. We get that $T_0(S_0(f)) = f$ for all $f \in H^{k-1/2}(\mathbb{R}^{n-1})$ because $\mathcal{S}(\mathbb{R}^{n-1})$ is dense in $H^{k-1/2}(\mathbb{R}^{n-1})$. ■

Since $\frac{\partial^j}{\partial x_n^j}$ is a bounded operator from $H^k(\mathbb{R}^n)$ to $H^{k-j}(\mathbb{R}^n)$ for $j \leq k$, we may define $T_j : H^k(\mathbb{R}_+^n) \rightarrow H^{k-j-1/2}(\mathbb{R}^{n-1})$ as $T_j = T \circ \frac{\partial^j}{\partial x_n^j} \circ E$, where T and E were defined in Theorem 11.8.5. We get the following theorem.

Theorem 11.8.6

Let T be the trace operator defined in the previous proposition and $E : H^k(\mathbb{R}_+^n) \rightarrow H^k(\mathbb{R}^n)$ be a strong k -extension operator. If $k > 1/2 + j$ with j a non-negative integer, then $T_j = T \circ \frac{\partial^j}{\partial x_n^j} \circ E : H^k(\mathbb{R}_+^n) \rightarrow H^{k-j-1/2}(\mathbb{R}^{n-1})$ is continuous and onto. Moreover, if $f = \phi|_{\mathbb{R}_+^n}$ for $\phi \in \mathcal{S}(\mathbb{R}^n)$, then $T_j(f)(z) = \frac{\partial^j f}{\partial x_n^j}(z, 0)$ for $z \in \mathbb{R}^{n-1}$. The operator T_j is also called a **trace operator**,

Proof.

Since T , $\frac{\partial^j}{\partial x_n^j}$ and E are continuous, we obviously have that their composition $T_j = T \circ \frac{\partial^j}{\partial x_n^j} \circ E$ is continuous.

To prove that T_j is onto, we do as we did for T_0 in the proof of Theorem 11.8.5; namely, we build a right inverse $S_j : H^{k-j-1/2}(\mathbb{R}^{n-1}) \rightarrow H^k(\mathbb{R}_+^n)$ for T_j . Since $\mathcal{S}(\mathbb{R}^{n-1})$ is dense in $H^{k-j-1/2}(\mathbb{R}^{n-1})$, we only have to define S_j on $\mathcal{S}(\mathbb{R}^{n-1})$.

We first consider the mapping $S : \mathcal{S}(\mathbb{R}^{n-1}) \rightarrow \mathcal{S}(\mathbb{R}^n)$ defined by

$$S(\phi)(x) = x_n^j \mathcal{F}^{-1} \left(\frac{(1 + \|\check{y}\|^2)^{k-j-1/2}}{C(1 + \|\mathbf{y}\|^2)^{k-j}} \mathcal{F}(\phi)(\check{y}) \right)(x)$$

for $\phi \in \mathcal{S}(\mathbb{R}^{n-1})$ and $x \in \mathbb{R}^n$, where $C = (2\pi)^{-1/2} j! \int_{\mathbb{R}} (1 + x^2)^{-k+j} dx$ and $\mathbf{y} = (\check{y}, y_n) \in \mathbb{R}^n$.

To prove that S is a continuous mapping from $H^{k-j-1/2}(\mathbb{R}^{n-1})$ to $H^k(\mathbb{R}^n)$, we note that

$$\begin{aligned} \|S(\phi)\|_{k,\rho}^2 &= \int_{\mathbb{R}^n} |\mathcal{F}(S(\phi))(\mathbf{y})|^2 (1 + \|\mathbf{y}\|^2)^k d\mathbf{y} \\ &= \int_{\mathbb{R}^n} \left| \frac{\partial^j}{\partial y_n^j} \left(\frac{(1 + \|\check{y}\|^2)^{k-j-1/2}}{C(1 + \|\mathbf{y}\|^2)^{k-j}} \mathcal{F}(\phi)(\check{y}) \right) \right|^2 (1 + \|\mathbf{y}\|^2)^k d\mathbf{y} \\ &= \int_{\mathbb{R}^{n-1}} \frac{1}{C^2} \left(\int_{\mathbb{R}} \left| \frac{\partial^j}{\partial y_n^j} \left(\frac{1}{(1 + \|\mathbf{y}\|^2)^{k-j}} \right) \right|^2 (1 + \|\mathbf{y}\|^2)^k dy_n \right) (1 + \|\check{y}\|^2)^{2k-2j-1} |\mathcal{F}(\phi)(\check{y})|^2 d\check{y} \\ &= \int_{\mathbb{R}^{n-1}} \frac{1}{C^2} \left(\frac{1}{(1 + \|\check{y}\|_2^2)^{k-j-1/2}} \int_{\mathbb{R}} \left| \sum_{0 \leq s \leq [j/2]} C_{j,s} \frac{u^{j-2s}}{(1 + u^2)^{k-s}} \right|^2 (1 + u^2)^k du \right) \\ &\quad (1 + \|\check{y}\|^2)^{2k-2j-1} |\mathcal{F}(\phi)(\check{y})|^2 d\check{y} \\ &= Q_0^2 \int_{\mathbb{R}^{n-1}} (1 + \|\check{y}\|^2)^{k-j-1/2} |\mathcal{F}(\phi)(\check{y})|^2 d\check{y} = Q_0^2 \|\phi\|_{k-j-1/2,\rho}^2, \end{aligned}$$

where

$$Q_0^2 = \frac{1}{C^2} \int_{\mathbb{R}} \left| \sum_{0 \leq s \leq j/2} C_{j,s} \frac{u^{j-2s}}{(1+u^2)^{k-s}} \right|^2 (1+u^2)^k du$$

and the $C_{j,s}$ are constants depending on k and j . The substitution $y_n = (1 + \|\check{\mathbf{y}}\|^2)^{1/2} u$ was used to obtain the fourth equality in the expression above. We also note that $\frac{\partial}{\partial y_n} = (1 + \|\check{\mathbf{y}}\|^2)^{-1/2} \frac{\partial}{\partial u}$.

We have shown that $\|S(\phi)\|_{k,\rho} = Q_0 \|\phi\|_{k-1/2,\rho}$ for all $\phi \in \mathcal{S}(\mathbb{R}^{n-1})$ and some constant Q_0 . Since $\mathcal{S}(\mathbb{R}^{n-1})$ is dense in $H^{k-j-1/2}(\mathbb{R}^{n-1})$. We can extend S to $H^{k-j-1/2}(\mathbb{R}^{n-1})$ and still have $\|S(f)\|_{k,\rho} = Q_0 \|f\|_{k-j-1/2,\rho}$ for all $f \in H^{k-j-1/2}(\mathbb{R}^{n-1})$.

The operator S_j is defined by $S_j = i \circ S$ where $i : H^k(\mathbb{R}^n) \rightarrow H^k(\mathbb{R}_+^n)$ is the continuous mapping defined by $i(f) = f|_{\mathbb{R}_+^n}$ for all $f \in H^k(\mathbb{R}^n)$. Because of the equivalence between the norms $\|\cdot\|_{k,2,\mathbb{R}^n}$ and $\|\cdot\|_{k,\rho}$, there exists a constant Q_1 such that $\|S_j(f)\|_{k,2,\mathbb{R}_+^n} \leq Q_1 \|f\|_{k-j-1/2,\rho}$ for all $f \in H^{k-j-1/2}(\mathbb{R}^{n-1})$.

To complete the proof of the theorem, we now show that S_j is the right inverse of T_j . Using the substitution $y_n = (1 + \|\check{\mathbf{y}}\|^2)^{1/2} u$, we obtain

$$\begin{aligned} \int_{\mathbb{R}} \frac{1}{(1 + \|\mathbf{y}\|^2)^{k-j}} dy_n &= \int_{\mathbb{R}} \frac{1}{(1 + \|\check{\mathbf{y}}\|^2 + y_n^2)^{k-j}} dy_n = \frac{1}{(1 + \|\check{\mathbf{y}}\|^2)^{k-j-1/2}} \int_{\mathbb{R}} \frac{1}{(1 + u^2)^{k-j}} du \\ &= \frac{C(2\pi)^{1/2}}{(1 + \|\check{\mathbf{y}}\|^2)^{k-j-1/2} j!} . \end{aligned}$$

Hence,

$$\begin{aligned} \frac{\partial^j}{\partial x_n^j} S(\phi)(\mathbf{z}, 0) &= j! \mathcal{F}^{-1} \left(\frac{(1 + \|\check{\mathbf{y}}\|^2)^{k-j-1/2}}{C(1 + \|\mathbf{y}\|^2)^{k-j}} \mathcal{F}(\phi)(\check{\mathbf{y}}) \right) (\mathbf{z}, 0) \\ &= \frac{j!}{(2\pi)^{n/2}} \int_{\mathbb{R}^{n-1}} \int_{\mathbb{R}} \frac{(1 + \|\check{\mathbf{y}}\|^2)^{k-j-1/2}}{C(1 + \|\mathbf{y}\|^2)^{k-j}} \mathcal{F}(\phi)(\check{\mathbf{y}}) e^{i\check{\mathbf{y}} \cdot \mathbf{z}} dy_n d\check{\mathbf{y}} \\ &= \frac{j!}{(2\pi)^{n/2}} \int_{\mathbb{R}^{n-1}} \left(\int_{\mathbb{R}} \frac{1}{(1 + \|\mathbf{y}\|^2)^{k-j}} dy_n \right) \frac{(1 + \|\check{\mathbf{y}}\|^2)^{k-j-1/2}}{C} \mathcal{F}(\phi)(\check{\mathbf{y}}) e^{i\check{\mathbf{y}} \cdot \mathbf{z}} d\check{\mathbf{y}} \\ &= \frac{1}{(2\pi)^{(n-1)/2}} \int_{\mathbb{R}^{n-1}} \mathcal{F}(\phi)(\check{\mathbf{y}}) e^{i\check{\mathbf{y}} \cdot \mathbf{z}} d\check{\mathbf{y}} = \mathcal{F}^{-1}(\mathcal{F}(\phi))(\mathbf{z}) = \phi(\mathbf{z}) \end{aligned}$$

for all $\mathbf{z} \in \mathbb{R}^{n-1}$, we get that

$$T_j(S_j(\phi))(\mathbf{z}) = T \left(\frac{\partial^j}{\partial x_n^j} S(\phi) \right) = \frac{\partial^j}{\partial x_n^j} S(\phi)(\mathbf{z}, 0) = \phi(\mathbf{z})$$

for all $\mathbf{z} \in \mathbb{R}^{n-1}$ and all $\phi \in \mathcal{S}(\mathbb{R}^{n-1})$.

Hence, $T_j(S_j(\phi)) = \phi$ for all $\phi \in \mathcal{S}(\mathbb{R}^{n-1})$. We get that $T_j(S_j(f)) = f$ for all $f \in H^{k-j-1/2}(\mathbb{R}^{n-1})$ because $\mathcal{S}(\mathbb{R}^{n-1})$ is dense in $H^{k-j-1/2}(\mathbb{R}^{n-1})$. ■

Proposition 11.8.7

Given $f \in H^k(\mathbb{R}_+^n)$ and $T_j : H^k(\mathbb{R}_+^n) \rightarrow H^{k-j-1/2}(\mathbb{R}^{n-1})$ defined in Theorem 11.8.5 for $j = 0$ and Theorem 11.8.6 for $j > 0$. If $\text{supp } f \subset \mathbb{R}_+^n$ is compact, then $T_j(f) = 0$.

Proof.

Since $\text{supp } f \subset \mathbb{R}_+^n$ is compact, we have that $f \in H_0^k(\mathbb{R}_+^n)$ according to Lemma 11.2.3. So, there exists a sequence $\{\xi_i\}_{i=1}^\infty \subset \mathcal{D}(\mathbb{R}_+^n)$ that converges to f in $H_0^k(\mathbb{R}_+^n)$.

Since $T_j(\xi_i)(z) = \frac{\partial^j \xi_i}{\partial x_n^j}(z, 0) = 0$ for $z \in \mathbb{R}^{n-1}$, we get from the continuity of T_j on $H^k(\mathbb{R}_+^n)$ that $T_j(f) = \lim_{i \rightarrow \infty} T_j(\xi_i) = 0$. $\blacksquare$

Proposition 11.8.8

Given $f \in H^k(\mathbb{R}_+^n)$ and $T_j : H^k(\mathbb{R}_+^n) \rightarrow H^{k-j-1/2}(\mathbb{R}^{n-1})$ defined in Theorem 11.8.5 for $j = 0$ and Theorem 11.8.6 for $j > 0$. If $V \subset \mathbb{R}^n$ is an open set such that $V \supset K = \{(z, 0) : z \in \text{supp } T_j(f)\}$, then there exists $\tilde{f} \in H^k(\mathbb{R}_+^n)$ such that $\text{supp } \tilde{f} \subset V$ and $T_j(\tilde{f}) = T_j(f)$.

Proof.

Choose an open set W such that $K \subset W \subset \overline{W} \subset V$ and $\phi \in \mathcal{D}(V)$ such that $\phi(x) = 1$ for $x \in W$.

Since $\mathcal{D}(\mathbb{R}^n)|_{\mathbb{R}_+^n}$ is dense in $H^k(\mathbb{R}_+^n)$, there exists a sequence $\{\xi_i\}_{i=1}^\infty \subset \mathcal{D}(\mathbb{R}^n)$ such that $\xi_i \rightarrow f$ in $H^k(\mathbb{R}_+^n)$. Since $\|g\|_{j,2,\mathbb{R}_+^n} \leq \|g\|_{k,2,\mathbb{R}_+^n}$ for $g \in H^k(\mathbb{R}_+^n)$ and $0 \leq j \leq k$, we also have that $\xi_i \rightarrow f$ in $H^j(\mathbb{R}_+^n)$ as $i \rightarrow \infty$ for $j < k$.

Since $\phi \in \mathcal{D}(\mathbb{R}^n)$, we have from Lemma 11.1.17 that $\phi \xi_i \rightarrow \phi f$ in $H^j(\mathbb{R}_+^n)$ as $i \rightarrow \infty$ for $0 \leq j \leq k$. Moreover,

$$\begin{aligned} T_j(\phi \xi_i)(z) &= \frac{\partial^j}{\partial x_n^j}(\phi \xi_i)(z, 0) = \sum_{m=0}^j \binom{j}{m} \frac{\partial^{j-m} \phi}{\partial x_n^{j-m}}(z, 0) \frac{\partial^m \xi_i}{\partial x_n^m}(z, 0) \\ &= \sum_{m=0}^j \binom{j}{m} \frac{\partial^{j-m} \phi}{\partial x_n^{j-m}}(z, 0) T_m(\xi_i)(z) \end{aligned}$$

for all $z \in \mathbb{R}^{n-1}$. Since $\xi_i \rightarrow f$ in $H^m(\mathbb{R}_+^n)$ as $i \rightarrow \infty$ for $0 \leq m \leq j$, we get that $T_m(\xi_i) = \frac{\partial^m \xi_i}{\partial x_n^m} \rightarrow T_m(f)$ in $H^{k-m-1/2}(\mathbb{R}^{n-1})$ as $i \rightarrow \infty$ for $0 \leq m \leq j$. It follows from Lemma 11.1.22

that $\frac{\partial^{j-m} \phi}{\partial x_n^{j-m}} T_m(\xi_i) \rightarrow \frac{\partial^{j-m} \phi}{\partial x_n^{j-m}} T_m(f)$ in $H^{k-j-1/2}(\mathbb{R}^{n-1})$ as $i \rightarrow \infty$ for $0 \leq m \leq j$.

Hence,

$$T_j(\phi f) = \lim_{i \rightarrow \infty} T_j(\phi \xi_i) = \frac{\partial^j(\phi \xi_i)}{\partial x_n^j} = \sum_{m=0}^j \binom{j}{m} \frac{\partial^{j-m} \phi}{\partial x_n^{j-m}} T_m(f) .$$

Since $\frac{\partial^{j-m}\phi}{\partial x_n^{j-m}}(\mathbf{z}, 0) = 0$ for $\mathbf{z} \in W \supset K$ and $0 \leq m < j$, and $\phi(\mathbf{z}, 0) = 1$ for $\mathbf{z} \in W \supset K$, we get that $T_j(\phi f) = T_j(f)$. We can take $\tilde{f} = \phi f$. ■

Problems involving partial differential equations of degree greater than one do not (generally) come with only one condition on the boundary. For instance, the problem may have the conditions $T_m(f) = g_m \in H^{k-m-1/2}(\mathbb{R}^{n-1})$ for $0 \leq m \leq j$ if we assume that $k > 1/2 + j$.

Because T_m is onto, we know that there exists $f_m \in H^k(\mathbb{R}_+^n)$ such that $T_m(f_m) = g_m$ for each m . But, we want something stronger. We want f_m to be the same for all m . How can we find $f \in H^k(\mathbb{R}_+^n)$ such that $T_m(f) = g_m$ for $0 \leq m \leq j$? To develop a method to find f , we need to look at the operator S defined in the context of the proof of Theorem 11.8.6. We can easily see that $\frac{\partial^m}{\partial x_n^m} S(f)(\mathbf{z}, 0) = 0$ for $0 \leq m < j$ and $f \in H^{k-j-1/2}(\mathbb{R}^{n-1})$ because it is true for $\phi \in \mathcal{S}(\mathbb{R}^{n-1})$. Thus, $T_{m_1}(S_{m_2}(f)) = 0$ for $m_1 < m_2$. To find f , we proceed as follows. If $f_0 = S_0(g_0)$ and $f_m = S_m(g_m - T_m(f_{m-1})) + f_{m-1}$ for $0 < m \leq j$, then $f = f_j$. It is easy to verify that $T_m(f) = g_m$ for $0 \leq m \leq j$.

We now generalize the previous results about trace to bounded open subsets Ω of $\mathbb{R}^n$ with a smooth boundary.

Theorem 11.8.9

Let Ω be an open subset of $\mathbb{R}^n$ with a bounded boundary $\partial\Omega$. Assume that Ω satisfies the uniform C^k -regularity condition. If $k > 1/2 + j$ with $j \in \mathbb{N}$, then there exists a continuous and surjective linear operator $\mathcal{T}_j : H^k(\Omega) \rightarrow H^{k-j-1/2}(\partial\Omega)$ such that $\mathcal{T}_j(f)(\mathbf{z}) = \frac{\partial^j f}{\partial \nu^j}(\mathbf{z})$ for $\mathbf{z} \in \partial\Omega$ and $T(f) = \phi|_{\partial\Omega}$ for $\phi \in \mathcal{D}(\mathbb{R}^n)$. Again, the operator T_j is also called a **trace operator**,

Remark 11.8.10

As usual, $\nu(\mathbf{x})$ is the outward unit normal to $\partial\Omega$ at $\mathbf{x} \in \partial\Omega$. ♠

Proof.

Since $\partial\Omega$ is a compact manifold, we may assume that the uniform C^k -regularity condition given in Definition 11.6.7 is given for a finite atlas $\{(U_i, \psi_i)\}_{i=1}^I$ for $\partial\Omega$.

We have from item (3) in Definition 11.6.7 that $\psi_i(U_i \cap \partial\Omega) = B_0 \equiv B_1(\mathbf{0}) \cap \mathbb{R}_0^n$ for all i .

As stated in the definition of uniform C^k -regularity, there exists $\delta > 0$ such that $\{\mathbf{x} \in \mathbb{R}^n : \text{dist}(\mathbf{x}, \partial\Omega) < \delta\} \subset \bigcup_{i=1}^I U_i$. Let $U_0 = \{\mathbf{x} \in \Omega : \text{dist}(\mathbf{x}, \partial\Omega) > \delta/2\}$. We then have that $\{U_i\}_{i=0}^I$ is an open cover of Ω . Let $\{\phi_i\}_{i=0}^I$ be a partition of unity for Ω subordinated to $\{U_i\}_{i=0}^I$.

Definition of $\mathcal{T}_j$: Given $f \in H^k(\Omega)$, we have that $f = \sum_{i=0}^I f_i$ where $f_i \equiv \phi_i f \in H^k(\Omega)$ and $\text{supp } f_i \subset U_i \cap \overline{\Omega}$.

Since $\text{supp } f_0 \subset U_0 \subset \Omega$, we have that $\mathcal{T}_j(f_0) = 0 \in H^{k-j-1/2}(\partial\Omega)$ ¹⁷. We will assume that $0 < i \leq I$ from now on.

We have that $\text{supp } f_i \cap \partial U_i = \emptyset$ but we may have that $\text{supp } f_i \cap \partial\Omega \neq \emptyset$ (Figure 11.5). Since $U_i \cap \Omega \subset \Omega$, we have from Lemma 11.1.18 that $f_i \in H^k(U_i \cap \Omega)$.

Let $B_+ = B_1(\mathbf{0}) \cap \mathbb{R}_+^n$. We have from Proposition 11.4.2 that $h_i \equiv f_i \circ \psi_i^{-1} \in H^k(B_+)$ and $\text{supp } h_i \subset B_+ \cup B_0$. Therefore, we may extend h_i to get $h_i \in H^k(\mathbb{R}_+^n)$ by defining $h_i(\mathbf{y}) = 0$ for $\mathbf{y} \in \mathbb{R}_+^n \setminus B_+$. This can be justified by slightly modifying the proof of Lemma 11.1.7.

Let $q_i = T_j(h_i) \in H^{k-j-1/2}(\mathbb{R}^{n-1})$, where $T_j : H^k(\mathbb{R}_+^n) \rightarrow H^{k-j-1/2}(\mathbb{R}^{n-1})$ is the trace operator given in Theorem 11.8.5 for $j = 0$ or Theorem 11.8.6 for $j > 0$. We have that $q_i|_{B_1(\mathbf{0})} \in H^{k-j-1/2}(B_1(\mathbf{0}))$. It follows from the definitions E in Theorem 11.6.3 that $\text{supp } E(h_i) \subset B_1(\mathbf{0}) \subset \mathbb{R}^n$, and from the definition of T in Proposition 11.8.3 (both E and T were used to define T_j) that $\text{supp } q_i \subset B_1(\mathbf{0}) \subset \mathbb{R}^{n-1}$,

We get that $g_i \equiv q_i \circ \psi_i \in H^{k-j-1/2}(U_i \cap \partial\Omega)$. Since $\text{supp } g_i \subset U_i \cap \partial\Omega$, we may extend g_i to get $g_i \in H^{k-j-1/2}(\partial\Omega)$ by setting $g_i(\mathbf{x}) = 0$ for $\mathbf{x} \in \partial\Omega \setminus U_i$.

A trace operator $\mathcal{T}_j$ acting on f is defined by $\mathcal{T}_j(f) = \sum_{i=0}^I g_i$. Recall that $g_0 = \mathcal{T}_j(f_0) = 0$.

Surjectivity of $\mathcal{T}_j$: Given $g \in H^{k-j-1/2}(\partial\Omega)$, we have that $g = \sum_{i=0}^I (\phi_i g)$ where $\phi_i g \in H^{k-j-1/2}(\partial\Omega)$ and $\text{supp}(\phi_i g) \subset U_i \cap \partial\Omega \subset \partial\Omega$. Hence, $\phi_i g \in H^{k-j-1/2}(U_i \cap \partial\Omega)$.

It follows from Proposition 11.4.2 that $h_i \equiv (\phi_i g) \circ \psi_i^{-1} \in H^{k-j-1/2}(B_0)$. Since $\text{supp } h_i \subset B_0$, we may extend h_i to get $h_i \in H^{k-j-1/2}(\mathbb{R}_0^n)$ by defining $h_i(\mathbf{y}) = 0$ for $\mathbf{y} \in \mathbb{R}_0^n \setminus B_0$.

Since $T_j : H^k(\mathbb{R}_+^n) \rightarrow H^{k-j-1/2}(\mathbb{R}^{n-1})$ is onto, there exists $q_i \in H^k(\mathbb{R}_+^n)$ such that $T_j(q_i) = h_i$. According to Proposition 11.8.8, we may assume that $\text{supp } q_i \subset B_1(\mathbf{0}) \subset \mathbb{R}^n$.

Let $f_i = q_i \circ \psi_i$ on $U_i \cap \Omega$. We have that $\text{supp } f_i \subset U_i$ and $f_i \in H^k(U_i \cap \Omega)$. Note that we may have $\text{supp } f_i \cap \partial\Omega \neq \emptyset$ but that is fine. If we set $f_i(\mathbf{x}) = 0$ for $\mathbf{x} \in \Omega \setminus U_i$, we have that $f_i \in H^k(\Omega)$.

If $f = \sum_{i=0}^I f_i$, then $\mathcal{T}_j(f) = \sum_{i=0}^I \mathcal{T}_j(f_i) = \sum_{i=0}^I \phi_i g = g$ as required.

$\mathcal{T}_j$ for $\phi \in \mathcal{D}(\mathbb{R}^n)$: For this part of the proof, we assume that the ψ_i and ψ_j^{-1} are of class C^∞ . We basically go through the step to define $\mathcal{T}_j(\phi)$ and show what happen when $\phi \in \mathcal{D}(\mathbb{R}^n)$.

We have that $\phi = \sum_{i=0}^I (\phi_i \phi)$ where $\phi_i \phi \in \mathcal{D}(U_i \cap \overline{\Omega})$; in particular, $\text{supp}(\phi_i \phi) \cap \partial U_i = \emptyset$. We may have that $\text{supp}(\phi_i \phi) \cap \partial\Omega \neq \emptyset$.

We have that $h_i \equiv (\phi_i \phi) \circ \psi_i^{-1} \in \mathcal{D}(\overline{B_+})$; in particular, $\text{supp } h_i \cap \partial B_1(\mathbf{0}) = \emptyset$. We extend h_i to get $h_i \in \mathcal{D}(\overline{\mathbb{R}_+^n})$ by defining $h_i(\mathbf{y}) = 0$ for $\mathbf{y} \in \overline{\mathbb{R}_+^n} \setminus B_1(\mathbf{0})$.

¹⁷To be precise and consistent with our presentation, we may assume that there is a homeomorphism ψ_0 from U_0 onto an open set $W \subset \mathbb{R}_+^n$ such that ψ_0 and ψ_0^{-1} are of class C^k . Then $h_0 \equiv f_0 \circ \psi_0^{-1} \in H^k(W)$ can be extended to $H^k(\mathbb{R}_+^n)$ by $h_0(\mathbf{y}) = 0$ for $\mathbf{y} \in \mathbb{R}_+^n \setminus W$ because $\text{supp } h_0 \subset W$. It then follows from Proposition 11.8.7 that $q_0 = T_j(h_0) = 0$. We are now justified to set $\mathcal{T}_j(f_0) = 0$.

We have that $T_j(h_i) = \frac{\partial^j}{\partial y_n^j}((\phi_i \phi) \circ \psi_i^{-1})$ on $\mathbb{R}_0^n \cong \mathbb{R}^{n-1}$. Hence

$$\mathcal{T}_j(\phi)(z) = \sum_{i=0}^I \mathcal{T}_j(\phi_i \phi)(z) = \sum_{i=0}^I \frac{\partial^j}{\partial y_n^j}((\phi_i \phi) \circ \psi_i^{-1})(\psi_i(z)) = - \sum_{i=0}^I \frac{\partial^j}{\partial \nu^j} \phi_i \phi(z) = - \frac{\partial^j}{\partial \nu^j} \phi(z)$$

for $z \in \partial\Omega$.

We get the wrong sign because $\frac{\partial}{\partial y_n}$ refers to the normal to $\mathbb{R}_0^n$ pointing inside the region $\mathbb{R}_+^n$ while $\frac{\partial}{\partial \nu}$ refers to the normal to $\partial\Omega$ pointing outside Ω . To get the right sign, it suffices to change the orientation of the ψ_i in the atlas for $\partial\Omega$. This is left to the reader. We have ignored this fact until now to provide a cleaner presentation which is consistent with the notation used so far.

To prove the previous result when ψ_i is only of class C^k , one needs to prove that $T_j(f)(z) = \frac{\partial^j}{\partial x_n^j} f(z, 0)$ for $f \in C_c^k(\mathbb{R}_+^n)$. This can easily be proved because $\mathcal{D}(\mathbb{R}_+^n)$ is dense in $C_c^k(\mathbb{R}_+^n)$ according to the norm on $\|\cdot\|_{k,2,\mathbb{R}_+^n}$. ■

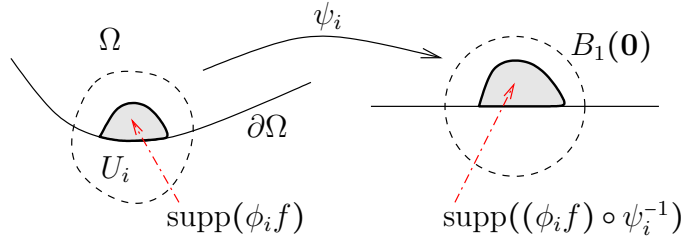


Figure 11.5: Graphical representation of a local chart used in the proof of Theorem 11.8.9.

We can use a trace operator to define $H_0^{k,2}(\Omega)$ where $\partial\Omega$ is a bounded C^k -manifold in $\mathbb{R}^n$. Before doing that, we need the following result.

Proposition 11.8.11

$$H_0^k(\mathbb{R}_+^n) = \{f \in H^k(\mathbb{R}_+^n) : T_j(f) = 0 \text{ for } 0 \leq j < k\},$$

where the operator T_j is defined in Theorem 11.8.5 for $j = 0$ and Theorem 11.8.6 for $j > 0$.

Proof.

⊂: Given $f \in H_0^k(\mathbb{R}_+^n)$, choose a sequence $\{\xi_i\}_{i=1}^\infty \subset \mathcal{D}(\mathbb{R}_+^n)$ that converges to f in $H^k(\mathbb{R}_+^n)$. Since $\|f\|_{j,2,\mathbb{R}_+^n} \leq \|f\|_{k,2,\mathbb{R}_+^n}$ for $0 \leq j < k$, we also have that $\{\xi_i\}_{i=1}^\infty$ converges to f in $H^j(\mathbb{R}_+^n)$

for $0 \leq j < k$. We have that $\mathcal{T}_j(\xi_i)(\mathbf{z}) = \frac{\partial^j \xi_i}{\partial x_n^j}(\mathbf{z}, 0) = 0$ for $\mathbf{z} \in \mathbb{R}^{n-1}$, $0 \leq j < k$ and all i because $\text{supp } \xi_i \subset \mathbb{R}_+^n$. Hence, we get by continuity of $\mathcal{T}_j$ on $H^j(\mathbb{R}_+^n)$ that $\mathcal{T}_j(f) = \lim_{i \rightarrow \infty} \mathcal{T}_j(\xi_i) = 0$ for $0 \leq j < k$.

$\supset$: Consider $f \in H^k(\mathbb{R}_+^n)$ such that $T_j(f) = 0$ for $0 \leq j < k$. According to Proposition 11.2.5, we will have that $f \in H_0^k(\mathbb{R}_+^n)$ if we show that

$$\underline{f}(\mathbf{x}) = \begin{cases} f(\mathbf{x}) & \text{for } \mathbf{x} \in \mathbb{R}_+^n \\ 0 & \text{for } \mathbf{x} \in \mathbb{R}^n \setminus \mathbb{R}_+^n \end{cases}$$

is in $H^k(\mathbb{R}^n)$. We have that $\underline{f}$ defines a distribution on $\mathbb{R}^n$ because

$$\int_{\mathbb{R}^n} |\underline{f}|^2 d\mathbf{x} = \int_{\mathbb{R}_+^n} |f|^2 d\mathbf{x} < \infty$$

since $f \in H^k(\mathbb{R}_+^n)$.

Since $\mathcal{D}(\mathbb{R}^n)|_{\mathbb{R}_+^n}$ is dense in $H^k(\mathbb{R}_+^n)$, there exists a sequence $\{\xi_i\}_{i=1}^\infty \subset \mathcal{D}(\mathbb{R}^n)$ such that $\tilde{\xi}_i \equiv \xi_i|_{\mathbb{R}_+^n} \rightarrow f$ in $H^k(\mathbb{R}_+^n)$ as $i \rightarrow \infty$.

Suppose that $\alpha \in \mathbb{N}^n$ is a multi-index such that $|\alpha| \leq k$ and $0 \leq \alpha_n = m \leq k$. Let $\alpha' = (\alpha_1, \dots, \alpha_{n-1}, 0)$. We get from integration by parts that

$$\begin{aligned} D^\alpha \tilde{\xi}_i(\phi) &= (-1)^{|\alpha|} \int_{\mathbb{R}^n} \tilde{\xi}_i D^\alpha \phi d\mathbf{x} = (-1)^{|\alpha|} \int_{\mathbb{R}_+^n} \tilde{\xi}_i D^\alpha \phi d\mathbf{x} \\ &= \sum_{j=0}^{m-1} (-1)^{m+j} \int_{\mathbb{R}^{n-1}} \left(\frac{\partial^j}{\partial x_n^j} D^{\alpha'} \tilde{\xi}_i(\mathbf{z}, 0) \right) \frac{\partial^{m-j-1} \phi}{\partial x_n^{m-j-1}}(\mathbf{z}, 0) d\mathbf{z} + \int_{\mathbb{R}_+^n} (D^\alpha \tilde{\xi}_i) \phi d\mathbf{x} \end{aligned} \quad (11.8.3)$$

for all $\phi \in \mathcal{D}(\mathbb{R}^n)$. If $m = 0$, the sum in the previous expression is not present.

We have for $0 \leq j < m$ that

$$\frac{\partial^j}{\partial x_n^j} D^{\alpha'} \tilde{\xi}_i(\mathbf{z}, 0) = D^{\alpha'} \frac{\partial^j}{\partial x_n^j} \tilde{\xi}_i(\mathbf{z}, 0) = D^{\alpha'} T_j(\tilde{\xi}_i) \rightarrow D^{\alpha'} T_j(f) = 0$$

in $H^{k-|\alpha'|-j-1/2}(\mathbb{R}^{n-1})$ as $i \rightarrow \infty$ because $T_j(\tilde{\xi}_i) \rightarrow T_j(f)$ in $H^{k-j-1/2}(\mathbb{R}^{n-1})$ and $D^{\alpha'}$ is a bounded operator from $H^{k-j-1/2}(\mathbb{R}^{n-1})$ to $H^{k-|\alpha'|-j-1/2}(\mathbb{R}^{n-1})$ according to Theorem 11.1.5.

Since $H^{k-j-1/2}(\mathbb{R}^{n-1}) \subset H^0(\mathbb{R}^{n-1})$ because $\|g\|_{0,\rho} \leq \|g\|_{k-j-1/2,\rho}$ for all $g \in H^{k-j-1/2}(\mathbb{R}^{n-1})$, we have that $\frac{\partial^j}{\partial x_n^j} D^{\alpha'} \tilde{\xi}_i(\mathbf{z}, 0) \rightarrow 0$ in $H^0(\mathbb{R}^{n-1})$ as $i \rightarrow \infty$ for $0 \leq j < m$. From Plancherel theorem, we get $\frac{\partial^j}{\partial x_n^j} D^{\alpha'} \tilde{\xi}_i(\mathbf{z}, 0) \rightarrow 0$ in $L^2(\mathbb{R}^{n-1})$ as $i \rightarrow \infty$ for $0 \leq j < m$. It then follows from Schwarz inequality that

$$\left| \int_{\mathbb{R}^{n-1}} \left(\frac{\partial^j}{\partial x_n^j} D^{\alpha'} \tilde{\xi}_i(\mathbf{z}, 0) \right) \frac{\partial^{m-j-1} \phi}{\partial x_n^{m-j-1}}(\mathbf{z}, 0) d\mathbf{z} \right|$$

$$\leq \left(\int_{\mathbb{R}^{n-1}} \left(\frac{\partial^j}{\partial x_n^j} D^{\alpha'} \tilde{\xi}_i(\mathbf{z}, 0) \right)^2 d\mathbf{z} \right)^{1/2} \left(\int_{\mathbb{R}^{n-1}} \left(\frac{\partial^{m-j-1} \phi}{\partial x_n^{m-j-1}}(\mathbf{z}, 0) \right)^2 d\mathbf{z} \right)^{1/2} \rightarrow 0$$

as $i \rightarrow \infty$ for $0 \leq j < m$. Since $\xi_i|_{\mathbb{R}_+^n} \rightarrow f$ in $H^k(\mathbb{R}_+^n)$ as $i \rightarrow \infty$, it also follows from Schwarz inequality that $\int_{\mathbb{R}_+^n} (D^\alpha \tilde{\xi}_i) \phi d\mathbf{x} \rightarrow \int_{\mathbb{R}_+^n} (D^\alpha f) \phi d\mathbf{x}$ as $i \rightarrow \infty$.

If $i \rightarrow \infty$ in (11.8.3), we then get that

$$D^\alpha \underline{f}(\phi) = (-1)^{|\alpha|} \int_{\mathbb{R}^n} \underline{f} D^\alpha \phi d\mathbf{x} = \int_{\mathbb{R}_+^n} (D^\alpha f) \phi d\mathbf{x} = \int_{\mathbb{R}^n} \underline{D^\alpha f} \phi d\mathbf{x}$$

for all $\phi \in \mathcal{D}(\mathbb{R}^n)$ ¹⁸. Thus, we have for all $|\alpha| \leq k$ that $D^\alpha \underline{f} = \underline{D^\alpha f}$ and $D^\alpha \underline{f} \in L^2(\mathbb{R}^n)$ because

$$\int_{\mathbb{R}^n} |\underline{D^\alpha f}|^2 d\mathbf{x} = \int_{\mathbb{R}_+^n} |D^\alpha f|^2 d\mathbf{x} < \infty$$

since $f \in H^k(\mathbb{R}_+^n)$. Therefore, $\underline{f} \in H^k(\mathbb{R}^n)$. ■

Theorem 11.8.12

Let $\Omega \subset \mathbb{R}^n$ be an open set that satisfies the uniform C^k -regularity condition. Then

$$H_0^k(\Omega) = \{f \in H^k(\Omega) : \mathcal{T}_j(f) = 0 \text{ for } 0 \leq j < k\}.$$

Proof (Sketch).

⊃: This part is basically identical to the part ⊃ in the proof of Proposition 11.8.11

Given $f \in H_0^k(\Omega)$, choose a sequence $\{\xi_i\}_{i=1}^\infty \subset \mathcal{D}(\Omega)$ that converges to f in $H^k(\Omega)$. Since $\|f\|_{j,2,\Omega} \leq \|f\|_{k,2,\Omega}$ for $0 \leq j < k$, we also have that $\{\xi_i\}_{i=1}^\infty$ converges to f in $H^j(\Omega)$ for $0 \leq j < k$. We know that $\mathcal{T}_j(\xi_i) = 0$ for $0 \leq j < k$ and all i . Hence, we get by continuity of $\mathcal{T}_j$ on $H^j(\Omega)$ that $\mathcal{T}_j(f) = 0$ for $0 \leq j < k$.

⊃: Suppose that $f \in H^k(\Omega)$ and $\mathcal{T}_j(f) = 0$ for $0 \leq j < k$.

As in the proof of Theorem 11.8.9, we may assume that $\{(U_i, \psi_i)\}_{i=1}^I$ is a finite atlas for $\partial\Omega$ with $\psi_i(U_i) = B_1(\mathbf{0})$, $\psi_i(U_i \cap \Omega) = B_+ = B_1(\mathbf{0}) \cap \mathbb{R}_+^n$ and $\psi_i(U_i \cap \partial\Omega) = B_0 = B_1(\mathbf{0}) \cap \mathbb{R}_0^n$ for all i . We also consider the open set $U_0 = \{\mathbf{x} \in \Omega : \text{dist}(\mathbf{x}, \partial\Omega) > \delta/2\}$ associated to the constant δ provided by the definition of uniform C^k -regularity. Let $\{\phi_i\}_{i=0}^I$ be a partition of unity for Ω subordinated to the open cover $\{U_i\}_{i=0}^I$ of Ω .

We have that $f = \sum_{i=0}^I f_i$ where $f_i \equiv \phi_i f \in H^k(\Omega)$. If we prove that $f_i \in H_0^k(\Omega)$ for all i , then we will have that $f \in H_0^k(\Omega)$. Since $\text{supp } f_0 \subset U_0 \subset \overline{U_0} \subset \Omega$, we have that $f_0 \in H_0^k(\Omega)$ according to Lemma 11.2.3. So, we only have to prove that $f_i \in H_0^k(\Omega)$ for $0 < i \leq I$.

It can be shown that $\mathcal{T}_j(f) = 0$ for $f \in H^k(\mathbb{R}_+^n)$ and $0 \leq j < k$ implies that $\mathcal{T}_j(\phi f) = 0$ for $\phi \in \mathcal{D}(\mathbb{R}^n)$ and $0 \leq j < k$. This can be generalized to show that $\mathcal{T}_j(f) = 0$ for $f \in H^k(\Omega)$ and

¹⁸Note that $\tilde{\xi}_i \rightarrow \underline{f}$ as distributions on $\mathbb{R}^n$, not necessarily in $H^k(\mathbb{R}^n)$.

$0 \leq j < k$ implies that $\mathcal{T}_j(\phi_i f) = \mathcal{T}_j(f_i) = 0$ for $0 \leq j < k$ ¹⁹. Basically, what we are saying is that the trace operator is determined locally.

If we refer to the definition of $\mathcal{T}_j$ given in the proof of Theorem 11.8.9, we get that $\mathcal{T}_j(h_i) = 0$ for $0 \leq j < k$, where $h_i \in H^k(\mathbb{R}_+^n)$ is defined by $h_i \equiv f_i \circ \psi_i^{-1}$ on $B_+ = B_1(\mathbf{0}) \cap \mathbb{R}_+^n$ and 0 on $\mathbb{R}_+^n \setminus B_1(\mathbf{0})$. We also have that $\text{supp } h_i \subset B_+ \cup B_0$.

It follows from Proposition 11.8.11 that $h_i \in H_0^k(\mathbb{R}_+^n)$. Therefore, there exists a sequence $\{\xi_m\}_{m=1}^\infty \subset \mathcal{D}(\mathbb{R}_+^n)$ converging to h_i in $H^k(\mathbb{R}_+^n)$. Since $\text{supp } h_i \subset B_1(\mathbf{0})$, we may assume that $\text{supp } \xi_m \subset B_+$. If it is not so, choose $\phi \in \mathcal{D}(\mathbb{R}^n)$ such that $\phi(\mathbf{x}) = 1$ for $\mathbf{x} \in \text{supp } h_i$ and $\phi(\mathbf{x}) = 0$ for $\mathbf{x} \in \mathbb{R}^n \setminus B_1(\mathbf{0})$. Then, $\text{supp } \phi \xi_m \subset B_1(\mathbf{0}) \cap \mathbb{R}_+^n = B_+$. It follows from Proposition 11.1.17 that $\phi \xi_m \rightarrow \phi h_i = h_i$ in $H^k(\mathbb{R}_+^n)$. Hence, $\tilde{\xi}_m \equiv (\phi \xi_m) \circ \psi_i \in C_c^k(\Omega \cap U_i)$ and $\tilde{\xi}_m \rightarrow h_i \circ \psi_i = f_i$ in $H^k(\Omega \cap U_i)$. Hence, $f_i \in H_0^k(U_i \cap \Omega)$ according to the second item of Remark 11.3.2 and thus $f_i \in H_0^k(\Omega)$ because $f_i = \phi_i f = 0$ on $\Omega \setminus U_i$. ■

11.8.1 Appendix

Theorem 11.8.12 applied to smooth functions can be proved without the use of trace operators. First, we need a lemma.

Lemma 11.8.13

Suppose $f \in W^{k,p}(\Omega)$ with $k \in \mathbb{N}$ and let U be an open subset of $\mathbb{R}^n$ such that $\overline{f(\Omega)} \subset U$. If $\psi \in C^k(U)$ is such that the derivatives $\psi^{(j)}$ for $0 \leq j \leq k$ are bounded on $\overline{f(\Omega)}$, then $\psi \circ f \in W^{k,p}(\Omega)$.

Proof (Sketch).

The proof that $\psi \circ f \in W^{k,p}$ requires the chain rule for the derivatives of composed functions and the formula for the change of variables in integrals. ■

Proposition 11.8.14

Let Ω be a bounded open subset of $\mathbb{R}^n$ that satisfies the uniform C^k -regularity condition. If $f \in C^{k-1}(\overline{\Omega}) \cap H^k(\Omega)$, then $f \in H_0^k(\Omega)$ if and only if $D^\alpha f = 0$ on $\partial\Omega$ for $|\alpha| < k$.

Proof.

The proof is basically a simpler version of the proof of Theorem 11.8.12 because $\mathcal{T}_j(f)(z) = \frac{\partial^j f}{\partial \nu^j}(z)$ for $z \in \partial\Omega$ and $0 \leq j < k$ when $f \in C^{k-1}(\overline{\Omega})$.

There is however an interesting and different proof due to [3] in the case $k = 1$ that we present below.

¹⁹This sketchy step of the proof is the reason behind the mention “Sketch”. The details require playing with the definition of manifold which we are trying to minimize in this book.

Since $\partial\Omega$ is bounded, it is a compact manifold. Therefore, we may assume that the uniform C^1 -regularity condition given in Definition 11.6.7 is given for a finite cover $\{U_i\}_{i=1}^I$ for $\partial\Omega$. As stated in the definition of uniform C^k -regularity, there exists $\delta > 0$ such that $\{\mathbf{x} \in \mathbb{R}^n : \text{dist}(\mathbf{x}, \partial\Omega) < \delta\} \subset \bigcup_{i=1}^I U_i$. Let $U_0 = \{\mathbf{x} \in \Omega : \text{dist}(\mathbf{x}, \partial\Omega) > \delta/2\}$. We then have that $\{U_i\}_{i=0}^I$ is an open cover of Ω . Let $\{\phi_i\}_{i=0}^I$ be a partition of unity for Ω subordinated to $\{U_i\}_{i=0}^I$.

Given $f \in C(\overline{\Omega})$, we can write $f = \sum_{j=0}^N f_j$ on $\overline{\Omega}$, where $f_j \equiv \phi_j f : \overline{\Omega} \rightarrow \mathbb{R}$ is such that $\text{supp } f_j \subset U_i \cap \overline{\Omega}$. We obviously have that $f_j \in C(\overline{\Omega})$ and it follows that $f_j \in H^1(\Omega)$.

$\Leftarrow$) Choose $\psi \in \mathcal{D}(\mathbb{R})$ such that $|\psi(x)| \leq 1$ for all x , $\psi(x) = 1$ for $|x| \geq 2$ and $\psi(x) = 0$ for $|x| \leq 1$. Let $\xi(x) = \psi(x)x$ for $x \in \mathbb{R}$.

For any positive integer m , the function $g_m : \overline{\Omega} \rightarrow \mathbb{R}$ defined by $g_m(\mathbf{x}) = \frac{1}{m} \xi(m f_j(\mathbf{x}))$ for $\mathbf{x} \in \overline{\Omega}$ is therefore in $H^1(\Omega)$ according to Lemma 11.8.13. Moreover,

$$\text{supp } g_m = \left\{ \mathbf{x} \in \overline{\Omega} : |f_j(\mathbf{x})| \geq \frac{1}{m} \right\} \subset \Omega$$

because f_j is continuous on $\overline{\Omega}$ and $f_j = 0$ on $\partial\Omega$. From Lemma 11.2.3, $g_m \in H_0^1(\Omega)$ for all positive integer m .

We now show that $g_m \rightarrow f_j$ in $H_0^1(\Omega)$. Since $H_0^1(\Omega)$ is closed, this will prove that $f_j \in h_0^1(\Omega)$.

Let α be a multi-index such that $|\alpha| \leq 1$. We first prove that $D^\alpha g_m(\mathbf{x}) \rightarrow D^\alpha f_j(\mathbf{x})$ as $m \rightarrow \infty$ for almost all $\mathbf{x} \in \Omega$. Since $D^\alpha g_m(\mathbf{x}) = 0$ for all m and $\mathbf{x} \in \Omega \setminus \text{supp } f_j$ because $D^\alpha f_j(\mathbf{x}) = 0$ for all $\mathbf{x} \in \Omega \setminus \text{supp } f_j$, we trivially have that $D^\alpha g_m(\mathbf{x}) \rightarrow D^\alpha f_j(\mathbf{x})$ as $m \rightarrow \infty$ for all $\mathbf{x} \in \Omega \setminus \text{supp } f_j$. Given $\mathbf{x} \in \Omega$ such that $f_j(\mathbf{x}) \neq 0$. Choose $C > 2$ and let M be a positive integer such that $M|f_j(\mathbf{x})| \geq C$. By continuity of the norm, there exists a neighbourhood $U \subset \Omega$ of $\mathbf{x}$ such that $M|f_j(\mathbf{y})| \geq 2$ for $\mathbf{y} \in U$. Hence, $g_m(\mathbf{y}) = f_j(\mathbf{y})$ for $m \geq M$ and $\mathbf{y} \in U$. We deduce that $D^\alpha g_m(\mathbf{x}) = D^\alpha f_j(\mathbf{x})$ for $m \geq M$. Thus $D^\alpha g_m(\mathbf{x}) \rightarrow D^\alpha f_j(\mathbf{x})$ as $m \rightarrow \infty$. Since this is true for all $\mathbf{x} \in \Omega$ such that $f_j(\mathbf{x}) \neq 0$, we have that $D^\alpha g_m \rightarrow D^\alpha f_j$ pointwise on Ω almost everywhere as $m \rightarrow \infty$. The only issue may be for $\mathbf{x} \in \Omega \cap \partial(\text{supp } f_j)$ where $f_j(\mathbf{x}) = 0$ but this is a set of measure zero since f_j is continuous on $\overline{\Omega}$.

We now prove that $|D^\alpha g_m|^2$ is bounded by the integrable function $K|D^\alpha f_j|^2$ for all m , where K is a constant. For all $m \geq 1$ and $\mathbf{x} \in \Omega$, we have

$$|g_m(\mathbf{x})|^2 = \left| \frac{1}{m} \xi(m f_j(\mathbf{x})) \right|^2 \leq |f_j(\mathbf{x})|^2$$

because $|\xi(x)| \leq |x|$ for all $x \in \mathbb{R}$. For $|\alpha| = 1$, $m \geq 1$ and $\mathbf{x} \in \Omega$, we have that

$$|D^\alpha g_m(\mathbf{x})|^2 = |\xi'(m f_j(\mathbf{x})) D^\alpha f_j(\mathbf{x})|^2 \leq K |D^\alpha f_j(\mathbf{x})|^2,$$

where

$$\sqrt{K} = \max \left\{ 1, \sup_{1 \leq |x| \leq 2} \{|\xi'(x)|\} \right\}.$$

Recall that $\xi(x) = 0$ for $|x| < 1$ and $\xi(x) = x$ for $x > 2$.

Since the sequence $\{(D^\alpha g_m)^2\}_{m=1}^\infty$ converges pointwise almost everywhere to $(D^\alpha f_j)^2$ and is bounded by the integrable function $K(D^\alpha f_j)^2$ on Ω , we get from the Lebesgue Dominated Convergence Theorem that $\{D^\alpha g_m\}_{m=1}^\infty$ converges to $D^\alpha f_j$ in $L^2(\Omega)$. Since α such that $|\alpha| \leq 1$ is arbitrary, we have $\|g_m - f_j\|_{1,2,\Omega} \rightarrow 0$ as $m \rightarrow \infty$.

$\Rightarrow$) Using the expansion $f = \sum_{j=0}^N f_j$ given at the beginning of the proof, it is enough to prove that $f_j|_{\partial\Omega} = 0$. We have that $f_j \in W_0^{1,2}(\Omega)$ because $f_j = f\phi_j$ with $f \in W_0^{1,2}(\Omega)$ and $\phi_j \in \mathcal{D}(\mathbb{R}^n)$. In fact, for the previous statement, the readers should convince themselves that if $\{\eta_i\}_{i=1}^\infty \subset \mathcal{D}(\Omega)$ satisfies $\lim_{i \rightarrow \infty} \|\eta_i - f\|_{1,2} = 0$ then $\{\phi_j \eta_i\}_{i=1}^\infty \subset \mathcal{D}(\Omega)$ and $\lim_{i \rightarrow \infty} \|\phi_j \eta_i - f_j\|_{1,2} = 0$. Thus $f_j \in W_0^{1,2}(\Omega)$.

Recall that we have from Definition 11.6.7 that the C^1 -diffeomorphism $\psi_j : U_j \rightarrow B_1(\mathbf{0}) \subset \mathbb{R}^n$ satisfies $\psi_j(U_j \cap \Omega) = B_+ = \{\mathbf{x} \in B_1(\mathbf{0}) : x_n > 0\}$, $\psi_j(U_j \setminus \overline{\Omega}) = B_- = \{\mathbf{x} \in B_1(\mathbf{0}) : x_n < 0\}$ and $\psi_j(U_j \cap \partial\Omega) = B_0 = \{\mathbf{x} \in B_1(\mathbf{0}) : x_n = 0\}$. We may also assume that Lemma 11.4.2 is satisfied by ψ_j .

Let $B = B_+ \cup B_0$ and

$$\tilde{f}_j(\mathbf{x}) = \begin{cases} f_j(\psi_j^{-1}(\mathbf{x})) & \text{if } \mathbf{x} \in B \\ 0 & \text{if } \mathbf{x} \in \overline{\mathbb{R}_+^n} \setminus B \end{cases}$$

We have that $\tilde{f}_j \in C(\overline{\mathbb{R}_+^n})$ because $f_j \in C(U_j \cap \overline{\Omega})$ and $\text{supp } f_j \subset U_j \cap \overline{\Omega}$ (Figure 11.5).

The problem is reduced to proving that $\tilde{f}_j|_{\mathbb{R}_+^n} = 0$. Let $\{\eta_i\}_{i=1}^\infty \subset \mathcal{D}(\Omega)$ be a sequence such that $\lim_{i \rightarrow \infty} \|\eta_i - f_j\|_{1,2,\Omega} = 0$. Since $\text{supp } f_j \subset U_j \cap \overline{\Omega}$, we may assume that $\text{supp } \eta_i \subset U_j \cap \Omega$ for all i . Let

$$\tilde{\eta}_i(\mathbf{x}) = \begin{cases} \eta_i(\psi_j^{-1}(\mathbf{x})) & \text{if } \mathbf{x} \in B \\ 0 & \text{if } \mathbf{x} \in \overline{\mathbb{R}_+^n} \setminus B \end{cases}$$

We get a sequence $\{\tilde{\eta}_i\}_{i=1}^\infty \subset C_c^1(\overline{\mathbb{R}_+^n})$ such that $\text{supp } \tilde{\eta}_i \subset B$ and $\lim_{i \rightarrow \infty} \|\tilde{\eta}_i - \tilde{f}_j\|_{1,2,\overline{\mathbb{R}_+^n}} = 0$.

Consider $\mathbf{x} = (\check{\mathbf{x}}, x_n) \in \mathbb{R}_+^n$ where, as usual, $\check{\mathbf{x}} \in \mathbb{R}^{n-1}$. We have

$$|\tilde{\eta}_i(\check{\mathbf{x}}, x_n)| = \left| \int_0^{x_n} \frac{\partial \tilde{\eta}_i}{\partial x_n}(\check{\mathbf{x}}, s) ds \right| \leq \int_0^{x_n} \left| \frac{\partial \tilde{\eta}_i}{\partial x_n}(\check{\mathbf{x}}, s) \right| ds.$$

Hence, for $0 < \epsilon < 1$, we get

$$\begin{aligned} \frac{1}{\epsilon} \int_{\mathbb{R}^{n-1}} \int_0^\epsilon |\tilde{\eta}_i(\check{\mathbf{x}}, x_n)| dx_n d\check{\mathbf{x}} &\leq \frac{1}{\epsilon} \int_{\mathbb{R}^{n-1}} \int_0^\epsilon \left(\int_0^{x_n} \left| \frac{\partial \tilde{\eta}_i}{\partial x_n}(\check{\mathbf{x}}, s) \right| ds \right) dx_n d\check{\mathbf{x}} \\ &\leq \frac{1}{\epsilon} \int_{\mathbb{R}^{n-1}} \int_0^\epsilon \left(\int_0^\epsilon \left| \frac{\partial \tilde{\eta}_i}{\partial x_n}(\check{\mathbf{x}}, s) \right| ds \right) dx_n d\check{\mathbf{x}} = \int_{\mathbb{R}^{n-1}} \int_0^\epsilon \left| \frac{\partial \tilde{\eta}_i}{\partial x_n}(\check{\mathbf{x}}, s) \right| ds d\check{\mathbf{x}}. \end{aligned} \quad (11.8.4)$$

Since $\tilde{\eta}_i \rightarrow \tilde{f}_j$ as $i \rightarrow \infty$ with respect to the norm $\|\cdot\|_{1,2,\overline{\mathbb{R}_+^n}}$, and $\text{supp } \tilde{f}_j$ and $\text{supp } \tilde{\eta}_i$ are subsets of the bounded set B for all i , we get from Schwarz inequality that

$$\int_{\overline{\mathbb{R}_+^n}} |\tilde{\eta}_i - \tilde{f}_j| d\mathbf{x} = \int_B |\tilde{\eta}_i - \tilde{f}_j| d\mathbf{x} \leq \left(\int_B 1^2 d\mathbf{x} \right)^{1/2} \left(\int_B (\tilde{\eta}_i - \tilde{f}_j)^2 d\mathbf{x} \right)^{1/2} \rightarrow 0 \quad \text{as } i \rightarrow \infty.$$

and

$$\begin{aligned} \int_{\mathbb{R}_+^n} \left| \frac{\partial \tilde{\eta}_i}{\partial x_m} - D^{e_m} \tilde{f}_j \right| d\mathbf{x} &= \int_B \left| \frac{\partial \tilde{\eta}_i}{\partial x_m} - D^{e_m} \tilde{f}_j \right| d\mathbf{x} \\ &\leq \left(\int_B 1^2 d\mathbf{x} \right)^{1/2} \left(\int_B \left(\frac{\partial \tilde{\eta}_i}{\partial x_m} - D^{e_m} \tilde{f}_j \right)^2 d\mathbf{x} \right)^{1/2} \rightarrow 0 \quad \text{as } i \rightarrow \infty \end{aligned}$$

for $1 \leq m \leq n$, and D^{e_m} is as usual a derivative in the sense of distributions. If we let $i \rightarrow \infty$ in (11.8.4), it then follows that

$$\frac{1}{\epsilon} \int_{\mathbb{R}^{n-1}} \int_0^\epsilon |\tilde{f}_j(\mathbf{x}', x_n)| dx_n d\mathbf{x}' \leq \int_{\mathbb{R}^{n-1}} \int_0^\epsilon |D^{e_m} \tilde{f}_j(\mathbf{x}', s)| ds d\mathbf{x}' \quad (11.8.5)$$

If we let $\epsilon \rightarrow 0$ in (11.8.5), we obtain

$$\int_{\mathbb{R}^{n-1}} |\tilde{f}_j(\mathbf{x}', 0)| d\mathbf{x}' \leq 0 \quad (11.8.6)$$

because $\tilde{f}_j \in C(\overline{\mathbb{R}_+^n})$ and $D^{e_m} \tilde{f}_j \in L^1(\mathbb{R}_+^n)$ since $\text{supp } \tilde{f}_j$ is a subset of the bounded set B . Basic calculus is used to compute the limit on the left hand side of (11.8.5) and the Lebesgue Dominate Convergence Theorem is used to compute the limit on the right hand side.

Since $\tilde{f}_j \in C(\overline{\mathbb{R}_+^n})$, we get from (11.8.6) that $\tilde{f}_j|_{\mathbb{R}_0^n} = 0$. ■

11.9 Nonexistence of Solutions

We conclude this chapter with a generalization of Theorem 3.2.1 proved by L. Hörmander in [14].

Definition 11.9.1

A linear differential operator L with C^∞ coefficients in $\Omega \subset \mathbb{R}^n$ is **locally solvable** at $\mathbf{x}_0 \in \Omega$ if there exists an open neighbourhood $U \subset \Omega$ of $\mathbf{x}_0$ such that, given any $f \in \mathcal{D}(U)$, there exists a solution $u \in \mathcal{D}'(U)$ in the sense of distribution for $Lu = f$.

Theorem 11.9.2 (Hörmander)

Let L be a linear differential operator with C^∞ coefficients in $\Omega \subset \mathbb{R}^n$ and let $Q(\mathbf{x}, \boldsymbol{\xi})$ for $\mathbf{x} \in \Omega$ and $\boldsymbol{\xi} \in \mathbb{R}^n$ be the principal symbol associated to L . Let

$$P(\mathbf{x}, \boldsymbol{\xi}) = \sum_{j=1}^n \left(\frac{\partial Q}{\partial \xi_j}(\mathbf{x}, \boldsymbol{\xi}) \frac{\partial \overline{Q}}{\partial x_j}(\mathbf{x}, \boldsymbol{\xi}) - \frac{\partial Q}{\partial x_j}(\mathbf{x}, \boldsymbol{\xi}) \frac{\partial \overline{Q}}{\partial \xi_j}(\mathbf{x}, \boldsymbol{\xi}) \right).$$

- Given $\mathbf{x}_0 \in \Omega$, if there exists $\boldsymbol{\xi} \in \mathbb{R}^n$ such that $Q(\mathbf{x}_0, \boldsymbol{\xi}) = 0$ and $P(\mathbf{x}_0, \boldsymbol{\xi}) \neq 0$, then L is not locally solvable at $\mathbf{x}_0$.
- If for every $\mathbf{x} \in \Omega$ we can find $\boldsymbol{\xi} \in \mathbb{R}^n$ such that $Q(\mathbf{x}, \boldsymbol{\xi}) = 0$ and $P(\mathbf{x}, \boldsymbol{\xi}) \neq 0$,

then there exists $f \in C^\infty(\Omega)$ such that $Lu = f$ has no local solution $u \in \mathcal{D}'(U)$ whatever the open set $U \subset \Omega$.

11.10 Exercises

Suggested exercises:

- In [3]: number 8.1 to 8.22 in Chapter 8.
- In [10]: number 1 in Section 6A; numbers 1 and 2 in Section 6B; numbers 1 and 2 in Section 6E.
- In [22]: numbers 2 to 4, 6 and 7 in Section 6.4; numbers 2, 7(a) and 9 in Section 6.5.

Chapter 12

Linear Elliptic Partial Differential Equations

To simplify the presentation of the results about linear elliptic partial differential equations in this chapter, we introduce a new notation for the derivatives in the sense of distributions. Let Ω be an open subset of $\mathbb{R}^n$ and u be a distribution on Ω . If $\alpha = e_j$, then we denote the derivative $D^\alpha u$ in the sense of distribution by $D_{x_j} u$. If $\alpha = e_{j_1} + 3e_{j_2}$, then we denote the derivative $D^\alpha u$ in the sense of distribution by $D_{x_{j_1}x_{j_2}x_{j_2}x_{j_2}} u$. And so on. Obviously, if u is sufficiently differentiable with locally integrable derivatives, then $D_{x_j} u = \frac{\partial u}{\partial x_j}$ and

$$D_{x_{j_1}x_{j_2}x_{j_2}x_{j_2}} u = \frac{\partial^4 u}{\partial x_{j_1} \partial x_{j_2}^3}, \text{ and so on.}$$

The main source of information for this chapter was [3, 10, 28, 39]; in particular the first two citations.

12.1 Introduction to Variational Methods

There are four steps in a **variational method**. The steps can be roughly described as follows.

1. Choose a “variational problem” such that a classical solutions (if any) is also a solution of the “variational problem.” This includes choosing the right Sobolev space to use with the variational problem. Solutions of the “variational problem” are called “weak solution.”
2. Prove the existence and uniqueness of a weak solution u of the variational equation; for instance, with the help of the Lax-Milgram Theorem.
3. Determine the Sobolev space of the highest order that contains the weak solution u . “Regularity results” are generally used at this stage.

4. Determine the highest order k (if any) of the space of C^k functions that contains the weak solution u . “Imbedding results” for Sobolev spaces are generally used at this stage. If k is equal or larger than the order of the original partial differential equation, then the weak solution is a classical solution of the original partial differential equation.

We now illustrate the variational approach to solve a linear elliptic partial differential equations with a couple of simple examples. These examples will illustrate the approach that will be used to develop a general theory to solve linear elliptic partial differential equations. Obviously, more technical results will be needed to solve the general problems as we will see in the next sections. Nevertheless, the basic principle stays the same.

Example 12.1.1

Consider the boundary value problem

$$\begin{aligned} -u'' + u &= f, & a < x < b \\ u(a) &= u(b) = 0 \end{aligned} \quad (12.1.1)$$

where $f \in L^2(]a, b[)$. We assume that all functions are real valued functions.

If $f \in C([a, b])$, then the classical solution of (12.1.1) is a function $u \in C^2([a, b])$ that satisfies (12.1.1). Recall that $C^k([a, b])$, $k \in \mathbb{N}$, is the set of all $u \in C([a, b]) \cap C^k(]a, b[)$ such that $\frac{d^j u}{dx^j}$ can be continuously extended to $[a, b]$ for $0 \leq j \leq k$.

If $f \in C([a, b])$, then (12.1.1) is a simple second order ordinary differential equation with constant coefficients that can easily be solved using basic techniques to solve such ordinary differential equations. We ignore this fact.

Variational Problem: If $u \in C^2([a, b])$ is a classical solution of (12.1.1), then

$$\int_a^b (-u''(x) + u(x) - f(x)) v(x) dx = 0$$

for all $v \in C_c^1(]a, b[)$. A simple integration by parts gives

$$\int_a^b (u'(x) v'(x) + u(x) v(x) - f(x) v(x)) dx = 0$$

for all $v \in C_c^1(]a, b[)$. This expression makes sense if we assume that u and v are in the Sobolev space $H_0^1(]a, b[)$, and the derivatives are in the sense of distributions. We now return to the assumption that $f \in L^2([a, b])$. Recall that $H_0^1(]a, b[)$ is the completion of $\mathcal{D}(]a, b[)$ in $L^2(]a, b[)$ with respect to the norm

$$\|f\|_{1,2} = \left(\sum_{|\alpha| \leq 1} \|D^\alpha f\|_2^2 \right)^{1/2}, \quad (12.1.2)$$

where $\|\cdot\|_2$ is the usual norm on $L^2(]a, b[)$. $H_0^1(]a, b[)$ is a Banach space with respect to the norm (12.1.2). As we have seen in the previous chapter, $u \in H_0^1(]a, b[)$ generalizes the notion of $u \in C^1([a, b])$ with $u(a) = u(b) = 0$.

The **variational formulation** of (12.1.1) is to find $u \in H_0^1(]a, b[)$ such that

$$\int_a^b (D_x u D_x v + u v - f v) dx = 0 \quad (12.1.3)$$

for all $v \in H_0^1(]a, b[)$. This is also called a **variational problem**. (12.1.3) is a **variational equation**. The solution $u \in H_0^1(]a, b[)$ of (12.1.3) is called a **weak solution**.

We show that a classical solution is a weak solution. If $u \in C^2([a, b])$ is a solution of (12.1.1), then

$$\int_a^b (-u''(x) + u(x) - f(x)) v(x) dx = 0$$

for all $v \in C_c^1(]a, b[)$. A simple integration by parts yields

$$\int_a^b (u'(x) v'(x) + u(x) v(x) - f(x) v(x)) dx = 0$$

for all $v \in C_c^1(]a, b[)$. Since $C_c^1(]a, b[) \supset \mathcal{D}(]a, b[)$ is dense in $H_0^1(]a, b[)$, we get (12.1.3). To be more precise, if $\{v_i\}_{i=1}^\infty \subset C_c^1([a, b])$ converges to $v \in H_0^1(]a, b[)$ with respect to the norm $\|\cdot\|_{1,2,]a,b[}$, then $\|v_i - v\|_{2,]a,b[} \rightarrow 0$ and $\|D_x v_i - D_x v\|_{2,]a,b[} \rightarrow 0$ as $i \rightarrow \infty$. Hence, from Schwarz inequality, we get that

$$\begin{aligned} \left| \int_a^b D_x u D_x v_i dx - \int_a^b D_x u D_x v dx \right| &\leq \int_a^b |D_x u| |D_x v_i - D_x v| dx \\ &\leq \left(\int_a^b |D_x u|^2 dx \right)^{1/2} \left(\int_a^b |D_x v_i - D_x v|^2 dx \right)^{1/2} \rightarrow 0 \quad \text{as } i \rightarrow \infty, \end{aligned}$$

because $u' \in L^2(]a, b[)$ since $u \in C^2([a, b])$. Similarly,

$$\int_a^b u v_i dx \rightarrow \int_a^b u v dx \quad \text{and} \quad \int_a^b f v_i dx \rightarrow \int_a^b f v dx \quad \text{as } i \rightarrow \infty.$$

Thus,

$$\int_a^b (D_x u D_x v_i + u v_i - f v_i) dx \rightarrow \int_a^b (D_x u D_x v + u v - f v) dx \quad \text{as } i \rightarrow \infty.$$

It also follows from Proposition 11.8.14 that $u \in H_0^1(]a, b[)$ because $u \in C^2([a, b]) \subset H^1(]a, b[) \cap C([a, b])$ and $u(a) = u(b) = 0$.

Existence and uniqueness of the weak solution: We show that Lax-Milgram Theorem, Corollary 4.5.5, can be used to conclude that the variational problem (12.1.3) has a unique weak solution $u \in H_0^1(]a, b[)$ given by

$$\begin{aligned} &\frac{1}{2} \int_a^b ((D_x u)^2 + (u)^2) dx - \int_a^b f u dx \\ &= \min_{v \in H_0^1(]a, b[)} \left\{ \frac{1}{2} \int_a^b ((D_x v)^2 + (v)^2) dx - \int_a^b f v dx \right\}. \end{aligned} \quad (12.1.4)$$

This property is known as the **Dirichlet principle**.

Consider the bilinear form

$$B(v, u) = \int_a^b (D_x u D_x v + u v) dx = \int_a^b (D_x u, u) \cdot (D_x v, v) dx$$

for all $u, v \in H_0^1(]a, b[)$. It is clear that B is bilinear. To prove that B is bounded, we use Schwarz inequality in $\mathbb{R}^2$ and $L^2(]a, b[)$ to get

$$\begin{aligned} |B(v, u)| &\leq \int_a^b |(D_x u, u) \cdot (D_x v, v)| dx \leq \int_a^b ((D_x u)^2 + (u)^2)^{1/2} ((D_x v)^2 + (v)^2)^{1/2} dx \\ &\leq \left(\int_a^b ((D_x u)^2 + (u)^2) dx \right)^{1/2} \left(\int_a^b ((D_x v)^2 + (v)^2) dx \right)^{1/2} = \|u\|_{1,2} \|v\|_{1,2} \end{aligned}$$

for all $u, v \in H_0^1(]a, b[)$. Since $B(u, u) = \|u\|_{1,2}^2$ for $u \in H_0^1(]a, b[)$, it is obvious that B is coercive. It is also obvious that B is symmetric; namely, $B(u, v) = B(v, u)$ for all $u, v \in H_0^1(]a, b[)$. Moreover, the function $f \in L^2(]a, b[)$ yields a bounded functional on $H_0^1(]a, b[)$ defined by

$$\tilde{f}(u) = \langle u, f \rangle_2 = \int_a^b f(x)u(x) dx$$

for all $u \in H_0^1(]a, b[)$. We let the reader prove that f effectively defines a bounded functional. We therefore can apply Lax-Milgram Theorem to the bilinear form B to get (12.1.4).

Regularity of the solution: The regularity of the solution is deduced from Theorem 12.6.18 that will present later. We will see that $u \in H^2(]a, b[)$.

Classical Solution: We prove that if $f \in C([a, b])$, then the weak solution u belongs to $C^2([a, b])$. According to Corollary 11.7.9, we get that $u \in C^1([a, b])$ because $u \in H^2(]a, b[)$. We also have from $B(v, u) = \langle v, f \rangle_2$ with $v \in \mathcal{D}(]a, b[) \subset H_0^1(]a, b[)$ that $D_x(D_x u) = u - f \in C([a, b])$, where the derivatives are in the sense of distributions. Hence, from Proposition 6.4.5 and Theorem 11.7.2 (or Corollary 11.7.7), we get that $D_x u \in C^1([a, b])$. Thus $u \in C^2([a, b])$.

We now prove that a weak solution $u \in C^2([a, b]) \cap H_0^1(]a, b[)$ is a classical solution. If $u \in C^2([a, b]) \cap H_0^1(]a, b[)$ be a solution of (12.1.3), then

$$\int_a^b (u'(x)v'(x) + u(x)v(x) - f(x)v(x)) dx = 0$$

for all $v \in C_c^1(]a, b[)$. A simple integration by parts shows that

$$\int_a^b (-u''(x) + u(x) - f(x))v(x) dx = 0$$

for all $v \in C_c^1(]a, b[)$. Since $C_c^1(]a, b[) \supset \mathcal{D}(]a, b[)$ is dense in $L^2(]a, b[)$, we obtain that $-u'' + u - f = 0$ almost everywhere in $]a, b[$. Since $u \in C^2([a, b])$, we have that $-u'' + u \in C([a, b])$. So, we may assume that $f \in C([a, b])$ and $-u'' + u - f = 0$ everywhere on $]a, b[$. Finally, we get from Proposition 11.8.14 that $u(a) = u(b) = 0$ because $u \in H_0^{1,2}(]a, b[) \cap C([a, b])$. Thus u is the classical solution of (12.1.1). $\clubsuit$

Example 12.1.2

Consider the Sturm-Liouville problem

$$\begin{aligned} -(pu')' + qu &= f, \quad 0 < x < 1 \\ u(0) &= u(1) = 0 \end{aligned} \quad (12.1.5)$$

where $f \in L^2(]0, 1[)$, $q \in C([0, 1])$ and $p \in C^1([0, 1])$. Moreover, we assume that $q(x) \geq 0$ for all $x \in [0, 1]$ and that there exists $\alpha > 0$ such that $p(x) \geq \alpha$ for all $x \in [0, 1]$.

Variational Problem: The weak solution is the solution $u \in H_0^{1,2}(]0, 1[)$ of the equation

$$\int_0^1 (p D_x u D_x v + q u v - f v) dx = 0 \quad (12.1.6)$$

for all $v \in H_0^{1,2}(]0, 1[)$.

We prove that a classical solution is a weak solution. If $u \in C^2([0, 1])$ is a classical solution of (12.1.5), then it follows from integration by parts that

$$\int_0^1 (p u' v' + q u v - f v) dx = \int_0^1 (-(p u')' v + q u v - f v) dx = 0$$

for all $v \in C_c^1(]0, 1[)$. Since $C_c^1(]0, 1[) \supset \mathcal{D}(]0, 1[)$ is dense in $H_0^{1,2}(]0, 1[)$, we have that (12.1.6) is true. Moreover, it follows from Proposition 11.8.14 that $u \in H_0^{1,2}(]0, 1[)$ because $u \in C([0, 1])$ and $u(0) = u(1) = 0$.

Existence and uniqueness of the weak solution:

$$\begin{aligned} B : H_0^{1,2}(]0, 1[) \times H_0^{1,2}(]0, 1[) &\rightarrow \mathbb{R} \\ (u, v) &\mapsto \int_0^1 (p(x) D_x u(x) D_x v(x) + q(x) u(x) v(x)) dx \end{aligned}$$

is a coercive, symmetric bounded bilinear mapping. The prove that B is a symmetric bilinear form is easy and is left to the reader. To prove that B is bounded, let

$$M = \max \left\{ \max_{x \in [0, 1]} |p(x)|, \max_{x \in [0, 1]} |q(x)| \right\}.$$

Then, using Schwarz inequality in $\mathbb{R}^2$ and $L^2([0, 1])$, we get

$$\begin{aligned} |B(u, v)| &\leq \int_0^1 |p D_x u D_x v + q u v| dx \leq M \int_0^1 (|D_x u D_x v| + |u v|) dx \\ &= M \int_0^1 (|D_x u|, |u|) \cdot (|D_x v|, |v|) dx \leq M \int_0^1 \|(D_x u, u)\|_2 \|(D_x v, v)\|_2 dx \\ &\leq M \left(\int_0^1 \|(D_x u, u)\|_2^2 dx \right)^{1/2} \left(\int_0^1 \|(D_x v, v)\|_2^2 dx \right)^{1/2} \\ &= M \left(\int_0^1 (|D_x u|^2 + |u|^2) dx \right)^{1/2} \left(\int_0^1 (|D_x v|^2 + |v|^2) dx \right)^{1/2} = M \|u\|_{1,2} \|v\|_{1,2} \end{aligned}$$

for all $u, v \in H_0^{1,2}([0, 1[)$. To prove that B is coercive, we note that

$$B(u, u) = \int_0^1 (p |D_x u|^2 + q |u|^2) dx \geq \alpha \int_0^1 |D_x u|^2 dx \geq \frac{\alpha}{C} \|u\|_{1,2}^2$$

for all $u \in H_0^{1,2}([0, 1[)$, where C is the constant given by the Poincaré's inequality, Theorem 11.7.11. Moreover, the function $f \in L^2([0, 1[)$ yields a bounded functional on $H_0^{1,2}([0, 1[)$ defined by

$$\tilde{f}(u) = \langle u, f \rangle_2 = \int_{\Omega} f u dx$$

for all $u \in H_0^{1,2}([0, 1[)$. Thus, we may apply Lax-Milgram Theorem, Corollary 4.5.5, to (12.1.6) to conclude that there exists a unique weak solution $u \in H_0^{1,2}([0, 1[)$ given by

$$\begin{aligned} & \frac{1}{2} \int_0^1 (p D_x u D_x u + q u u) dx - \int_0^1 f u dx \\ &= \min_{v \in H_0^{1,2}([0, 1[)} \left\{ \frac{1}{2} \int_0^1 (p D_x v D_x v + q v v) dx - \int_0^1 f v dx \right\}. \end{aligned}$$

Regularity of the solution: As a consequence of Theorem 12.6.18 that we will present later, we have that $u \in H^{2,2}([0, 1[)$.

Classical solution: We prove that if $f \in C([0, 1])$, then the weak solution u belongs to $C^2([0, 1])$ and it is a classical solution of (12.1.5).

According to Corollary 11.7.9, $u \in H^{2,2}([0, 1[)$ implies that $u \in C^1([0, 1])$. Since $H_0^{1,2}([0, 1[) \supset \mathcal{D}([0, 1[)$, we also have from $B(v, u) = 0$ for all $v \in \mathcal{D}([0, 1[)$ that $D_x(p D_x u) = q u - f \in C([0, 1])$. Hence, from Proposition 6.4.5 and Theorem 11.7.2, $u' \in C^1([0, 1])$. It follows that $u \in C^2([0, 1])$.

To prove that our weak solution $u \in C^2([0, 1]) \cap H_0^1([0, 1[)$ satisfies (12.1.5) in the classical sense, we use integration by parts on (12.1.6) to get

$$\int_0^1 (-(p u')' + q u - f) v dx = \int_0^1 (p u' v' + q u v - f v) dx = 0$$

for all $v \in C_c^1([0, 1])$. Since $C_c^1([0, 1]) \supset \mathcal{D}([0, 1[)$ is dense in $L^2([0, 1])$, we get $-(p u')' + q u - f = 0$ almost everywhere. Since f, p' and q are all continuous functions, $-(p u')' + q u - f = 0$ everywhere on $]0, 1[$ because $u \in C^2([0, 1])$. It also follows from Proposition 11.8.14 that $u(0) = u(1) = 0$ because $u \in H_0^1([0, 1[) \cap C^2([0, 1])$. Thus u is the classical solution of (12.1.5). $\clubsuit$

12.2 Classical, Strong and Weak Solutions of Elliptic Problems

We consider strongly elliptic linear operators of the form

$$L(\mathbf{x}, D) = \sum_{|\gamma| \leq m} a_{\gamma}(\mathbf{x}) D^{\gamma}, \quad (12.2.1)$$

where the coefficients a_γ are all of class $C^\infty(\mathbb{R}^n)$. Recall that an elliptic linear operator paused on a bounded open set Ω is strongly elliptic on Ω if $a_\gamma : \overline{\Omega} \rightarrow \mathbb{R}$ is continuous for all γ . According to Proposition 2.4.11., it is also of even order; namely, m is even. If $n > 2$ and some of the coefficients are complex valued functions, then the order is also even. We assume that any one of these necessary conditions is satisfied.

It is useful to write $L(\mathbf{x}, D)$ in the **divergence form**

$$L(\mathbf{x}, D)u = \sum_{|\alpha|, |\beta| \leq m/2} (-1)^{|\alpha|} D^\alpha (b_{\alpha, \beta}(\mathbf{x}) D^\beta u) , \quad (12.2.2)$$

where $b_{\alpha, \beta} \in C^\infty(\Omega)$ for all multi-indices α and β .

If the linear differential operator $L(\mathbf{x}, D)$ is given in the divergence form (12.2.2), then the principal symbol takes the form

$$Q_{df}(\mathbf{x}, \xi) = \sum_{|\alpha|, |\beta| = m/2} \xi^\alpha b_{\alpha, \beta}(\mathbf{x}) \xi^\beta , \quad \mathbf{x} \in \mathbb{R}^n \text{ and } \xi \in \mathbb{R}^n .$$

If we compare (12.2.1) and (12.2.2), we find that $a_\gamma = (-1)^{m/2} \sum_{\alpha+\beta=\gamma} b_{\alpha, \beta}$ for $|\gamma| = m$. We may therefore modify Definition 2.4.12 as follows. The linear differential operator $L(\mathbf{x}, D)$ given in (12.2.2) is **strongly elliptic** on an open set $\Omega \subset \mathbb{R}^n$ if there exists a constant $C > 0$ such that $\operatorname{Re} Q_{df}(\mathbf{x}, \xi) \geq C \|\xi\|^m$ for all $\mathbf{x} \in \Omega$ and $\xi \in \mathbb{R}^n$. Note that there is no factor $(-1)^{m/2}$ on the left side because we have a factor of $(-1)^{m/2}$ in the definition of $Q(\mathbf{x}, \xi)$. So, the factor on the left hand side is $(-1)^m = 1$ since m is even. In the future, we will not use different notations for the principal symbol associated to a partial differential equation in standard form or in divergence form. The context will determine which form of the principal symbol is used.

For convenience, we also adopt in the following definition for an elliptic partial differential operator in divergence form. We will say that (12.2.2) is **elliptic** on an open set Ω if $Q_{df}(\mathbf{x}, \xi) > 0$ for all $\mathbf{x} \in \Omega$ and $\xi \in \mathbb{R}^n$ with $\xi \neq 0$.

Except otherwise stated, we always assume that the linear differential operators in this section are in divergence form (12.2.2).

12.2.1 Dirichlet Problem

In this section and the next one, we assume that Ω is an open subset of $\mathbb{R}^n$ and that $\partial\Omega$ is a C^∞ -manifold, or at least that Ω satisfies a strong q -extension property with $q \geq m/2$. Let $k = m/2$. We consider the Dirichlet problem

$$\begin{aligned} L(\mathbf{x}, D)u &= \sum_{|\alpha|, |\beta| \leq k} (-1)^{|\alpha|} D^\alpha (b_{\alpha, \beta}(\mathbf{x}) D^\beta u) = f , \quad \mathbf{x} \in \Omega , \\ \frac{\partial^j u}{\partial \nu^j}(\mathbf{x}) &= 0 , \quad \mathbf{x} \in \partial\Omega \text{ and } 0 \leq j < k , \end{aligned} \quad (12.2.3)$$

where $L(\mathbf{x}, D)$ is at least elliptic and $\nu(\mathbf{x})$ is the outward unit normal vector to $\partial\Omega$ at $\mathbf{x} \in \partial\Omega$.

There are many interpretations of solving this problem. We have already considered two interpretations of solving problems like (12.2.3) which we summarize in the following two definitions.

Definition 12.2.1

If $f \in C_b(\Omega)$, a **classical solution** for (12.2.3) is a function $u \in C_b^{2k}(\Omega) \cap C_b^{k-1}(\overline{\Omega})$ such that

$$L(\mathbf{x}, D)u(\mathbf{x}) = \sum_{|\alpha|, |\beta| \leq k} (-1)^{|\alpha|} D^\alpha (b_{\alpha, \beta}(\mathbf{x}) D^\beta u(\mathbf{x})) = f(\mathbf{x}) \quad , \quad \mathbf{x} \in \Omega \quad ,$$

and

$$\frac{\partial^j u}{\partial \nu^j}(\mathbf{x}) = 0 \quad , \quad \mathbf{x} \in \partial\Omega \text{ and } 0 \leq j < k \quad .$$

All partial and directional derivatives are in the classical sense.

Definition 12.2.2

If $f \in L_{loc}^2(\Omega)$, a **strong solution** for (12.2.3) is a distribution $u \in H^{2k,2}(\Omega) \cap H_0^{k,2}(\Omega) \subset \mathcal{D}'(\Omega)$ such that

$$\langle L(\mathbf{x}, D)u, \phi \rangle = \sum_{|\alpha|, |\beta| \leq k} \langle (-1)^{|\alpha|} D^\alpha (b_{\alpha, \beta}(\mathbf{x}) D^\beta u), \phi \rangle = \langle f, \phi \rangle$$

for all $\phi \in \mathcal{D}(\Omega)$, where $\langle v, \phi \rangle \equiv v(\phi)$ for $v \in \mathcal{D}'(\Omega)$ and $\phi \in \mathcal{D}(\Omega)$. Obviously, all derivatives are in the sense of distributions.

There is a third approach to study (12.2.3). It is the variational approach that we will consider in the rest of this section. We consider the **variational formulation** of (12.2.3) which consists in finding $u \in H_0^{k,2}(\Omega)$ such that

$$B(v, u) = \langle v, f \rangle_2 \quad (12.2.4)$$

for all $v \in H_0^{k,2}(\Omega)$, where

$$B : H^{k,2}(\Omega) \times H^{k,2}(\Omega) \rightarrow \mathbb{C} \\ (v, u) \mapsto \sum_{|\alpha|, |\beta| \leq k} \langle D^\alpha v, b_{\alpha, \beta}(\mathbf{x}) D^\beta u \rangle_2 \quad (12.2.5)$$

and $\langle v, u \rangle_2 = \int_\Omega v(\mathbf{x}) \overline{u(\mathbf{x})} d\mathbf{x}$ for $u, v \in L^2(\Omega)$ is the scalar product in $L^2(\Omega)$. Moreover, we assume that the coefficients $b_{\alpha, \beta}$ are bounded on Ω .

The mapping B is well defined. Suppose that $|b_{\alpha, \beta}(\mathbf{x})| \leq M$ for all $\mathbf{x} \in \Omega$, where M is a constant. Then,

$$|B(v, u)| = \left| \sum_{|\alpha|, |\beta| \leq k} \int_\Omega \overline{b_{\alpha, \beta}(\mathbf{x}) D^\beta u(\mathbf{x})} D^\alpha v(\mathbf{x}) d\mathbf{x} \right| \leq \sum_{|\alpha|, |\beta| \leq k} M \int_\Omega |\overline{D^\beta u(\mathbf{x})} D^\alpha v(\mathbf{x})| d\mathbf{x}$$

$$\begin{aligned}
&\leq \sum_{|\alpha|, |\beta| \leq k} M \left(\int_{\Omega} |D^{\beta} u(\mathbf{x})|^2 d\mathbf{x} \right)^{1/2} \left(\int_{\Omega} |D^{\alpha} v(\mathbf{x})|^2 d\mathbf{x} \right)^{1/2} \\
&\leq \left(\sum_{|\alpha|, |\beta| \leq k} M \right) \|u\|_{k,2} \|v\|_{k,2} = C \|u\|_{k,2} \|v\|_{k,2} < \infty
\end{aligned}$$

for all $(v, u) \in H^{k,2}(\Omega) \times H^{k,2}(\Omega)$, where we have used Schwarz inequality for the second inequality and set $C = \sum_{|\alpha|, |\beta| \leq k} M$.

Since

1. the mapping $u \mapsto B(v, u)$ is conjugate-linear for each $u \in H^{k,2}(\Omega)$,
2. the mapping $v \mapsto B(v, u)$ is linear for each $v \in H^{k,2}(\Omega)$, and
3. there exists a constant C such that $|B(v, u)| \leq C \|u\|_{k,2} \|v\|_{k,2}$ for $(v, u) \in H^{k,2}(\Omega) \times H^{k,2}(\Omega)$,

we get the following result.

Proposition 12.2.3

The mapping B defined in (12.2.5) is a bounded sesquilinear form.

The sesquilinear form B defined in (12.2.5) is called a **Dirichlet form** of order k .

Remark 12.2.4

A partial differential operator $L(\mathbf{x}, D)$ may have more than one Dirichlet form associated to it. Consider the Laplacian Δ in $\mathbb{R}^2$. Let $\eta = x - y$ and $\xi = x + y$. Then $B(v, u) = \langle D_x v, D_x u \rangle + \langle D_y v, D_y u \rangle$ and $B(v, u) = 4 \langle D_{\xi} v, D_{\eta} u \rangle$ are two Dirichlet forms associated to the Laplacian. Recall that $D_x, D_y, \dots$ denote derivatives in the sense of distributions. ♠

Suppose that $b_{\alpha, \beta} \in C_b^{|\alpha|}(\Omega)$ for all multi-indices α and β involved in the sum in (12.2.3). Moreover, suppose that $f \in C_b(\Omega)$, and $u \in C_b^{2k}(\Omega) \cap C_b^{k-1}(\bar{\Omega})$ satisfy (12.2.4) for all $v \in C_c^k(\Omega)$. Using integration by parts, (12.2.4) yields

$$\int_{\Omega} \left(\overline{\sum_{|\alpha|, |\beta| \leq k} (-1)^{\alpha} D^{\alpha} (b_{\alpha, \beta}(\mathbf{x}) D^{\beta} u(\mathbf{x}))} \right) v(\mathbf{x}) d\mathbf{x} = \int_{\Omega} \overline{f(\mathbf{x})} v(\mathbf{x}) d\mathbf{x}$$

for all $v \in C_c^k(\Omega)$. Since $C_c^k(\Omega) \cap \mathcal{D}(\Omega)$ is dense in $L^2(\Omega)$, and $L(\mathbf{x}, D)u$ and f are continuous on Ω , we must have that

$$L(\mathbf{x}, D)u(\mathbf{x}) = \sum_{|\alpha|, |\beta| \leq k} (-1)^{|\alpha|} D^{\alpha} (b_{\alpha, \beta}(\mathbf{x}) D^{\beta} u(\mathbf{x})) = f(\mathbf{x}) \quad , \quad \mathbf{x} \in \Omega .$$

Moreover, since $u \in H_0^{k,2}(\Omega)$ because we assume that u is a solution of (12.2.4), it follows from Proposition 11.8.14 that the boundary conditions in (12.2.3) are satisfied in the classical sense. So u is the classical solution of (12.2.3).

Definition 12.2.5

If $b_{\alpha,\beta} \in C_b^\infty(\Omega)$ and $f \in L^2(\Omega)$, a **weak solution** for (12.2.4) is a function $u \in H_0^{k,2}(\Omega)$ such that $B(v, u) = \langle v, f \rangle_2$ for all $v \in H_0^{k,2}(\Omega)$.

The variational approach is more general than the classical approach because we look for solutions in a larger space of functions; namely, $H_0^{k,2}(\Omega)$ instead of $C_b^{2k}(\Omega) \cap C_b^{k-1}(\overline{\Omega})$.

12.2.2 Neumann Problem

In this section, we consider the Neumann problem

$$\begin{aligned} L(\mathbf{x}, D)u(\mathbf{x}) &= - \sum_{i,j=1}^n D_{x_i} (b_{i,j} D_{x_j} u) (\mathbf{x}) + b_0(\mathbf{x})u(\mathbf{x}) = f(\mathbf{x}) \quad , \quad \mathbf{x} \in \Omega \quad , \\ \sum_{i,j=1}^n b_{i,j}(\mathbf{x}) (D_{x_j} u(\mathbf{x})) \nu_i(\mathbf{x}) &= 0 \quad , \quad \mathbf{x} \in \partial\Omega \quad , \end{aligned} \quad (12.2.6)$$

where $\nu(\mathbf{x})$ is the outward unit normal to $\partial\Omega$ at $\mathbf{x} \in \partial\Omega$. We also assume that $b_{i,j} \in C^\infty(\overline{\Omega})$ for all i and j , and $b_0 \in C^\infty(\overline{\Omega})$.

If we let B be the $n \times n$ matrix with components $b_{i,j}$ for $1 \leq i, j \leq n$, then the boundary condition in (12.2.6) can be rewritten

$$(\nabla u(\mathbf{x}))^\top B(\mathbf{x}) \nu(\mathbf{x}) = 0 \quad , \quad \mathbf{x} \in \partial\Omega \quad .$$

Example 12.2.6

If $b_{i,j} = 0$ for $i \neq j$ and $b_{i,i} = h$ for $1 \leq i \leq n$, where $h \in C^\infty(\Omega)$ and $h(\mathbf{x}) \neq 0$ for all $\mathbf{x} \in \Omega$, then

$$L(\mathbf{x}, D)u(\mathbf{x}) = -\Delta u(\mathbf{x}) + \sum_{i=1}^n b_i(\mathbf{x}) D_{x_i} u(\mathbf{x}) + \frac{b_0(\mathbf{x})}{h(\mathbf{x})} u(\mathbf{x}) = \frac{f(\mathbf{x})}{h(\mathbf{x})} \quad , \quad \mathbf{x} \in \Omega \quad ,$$

where $b_i = \frac{1}{h} \frac{\partial h}{\partial x_i}$ for $1 \leq i \leq n$, and the boundary condition becomes $\frac{\partial u}{\partial \nu} = 0$ on $\partial\Omega$. ♣

As for the Dirichlet Problem, We have already considered two interpretations of solving problems like (12.2.6) which we summarize in the following two definitions.

Definition 12.2.7

If $f \in C_b(\Omega)$, a **classical solution** for (12.2.6) is a function $u \in C_b^2(\Omega) \cap C_b^1(\overline{\Omega})$ such that

$$L(\mathbf{x}, D)u(\mathbf{x}) = - \sum_{i,j=1}^n D_{x_i} (b_{i,j} D_{x_j} u) (\mathbf{x}) + b_0(\mathbf{x})u(\mathbf{x}) = f(\mathbf{x}) \quad , \quad \mathbf{x} \in \Omega \quad ,$$

and

$$\sum_{i,j=1}^n b_{i,j}(\mathbf{x}) (D_{x_j} u(\mathbf{x})) \nu_i(\mathbf{x}) = 0 \quad , \quad \mathbf{x} \in \partial\Omega \quad .$$

All partial derivatives are in the classical sense.

Definition 12.2.8

If $f \in L^2_{loc}(\Omega)$, a **strong solution** for (12.2.6) is a distribution $u \in H^{2,2}(\Omega) \subset \mathcal{D}'(\Omega)$ such that

$$\langle L(\mathbf{x}, D)u, \phi \rangle = - \sum_{i,j=1}^n \langle D_{x_i} (b_{i,j} D_{x_j} u), \phi \rangle + \langle b_0 u, \phi \rangle = \langle f, \phi \rangle$$

for all $\phi \in \mathcal{D}(\Omega)$, and

$$\sum_{i,j=1}^n b_{i,j} (D_{x_j} u) \nu_i = 0$$

on $\partial\Omega$. Obviously, all derivatives are in the sense of distributions.

Since we consider elliptic linear partial differential equations, we can use a change of coordinates as we did in Chapter 2 to reduce the problem in (12.2.6) to one of the form given in Example 12.2.6.

As for the Dirichlet problem, there is a variational formulation of (12.2.6). The **variational formulation** of (12.2.6) consists in finding $u \in H^{1,2}(\Omega)$ such that

$$B(v, u) = \langle v, f \rangle_2 \quad (12.2.7)$$

for all $v \in H^{1,2}(\Omega)$, where

$$B : H^{1,2}(\Omega) \times H^{1,2}(\Omega) \rightarrow \mathbb{C} \quad (12.2.8)$$

$$(v, u) \mapsto \sum_{i,j=1}^n \langle D_{x_i} v, b_{i,j} D_{x_j} u \rangle_2 + \langle v, b_0 u \rangle_2$$

and, as usual, the scalar product $\langle \cdot, \cdot \rangle_2$ is defined by $\langle v, u \rangle_2 = \int_{\Omega} v(\mathbf{x}) \overline{u(\mathbf{x})} d\mathbf{x}$ for $u, v \in L^2(\Omega)$. Moreover, we assume that the coefficients $b_{i,j}$ are bounded on Ω . As we did for the Dirichlet problem, we can show that B is a bounded sesquilinear form on $H^{1,2}(\Omega)$.

Definition 12.2.9

If $b_{i,j} \in C_b^\infty(\overline{\Omega})$, $b_i \in C_b^\infty(\Omega)$ and $f \in L^2(\Omega)$, a **weak solution** for (12.2.7) is a function $u \in H^{1,2}(\Omega)$ such that $B(v, u) = \langle v, f \rangle_2$ for all $v \in H^{1,2}(\Omega)$.

Consider the variational problem (12.2.7), where $b_{i,j} \in C_b^1(\overline{\Omega})$, $b_0 \in C_b(\Omega)$ and $f \in C_b(\Omega)$. Suppose that $u \in C_b^2(\Omega) \cap C_b^1(\overline{\Omega})$ is a solution of 12.2.7.

For all $v \in C_b^1(\overline{\Omega})$, the divergence theorem yields

$$\begin{aligned} \int_{\partial\Omega} \sum_{i=1}^n \left(\overline{b_{i,j}} \frac{\partial \overline{u}}{\partial x_j} v \nu_i \right) d\mathbf{x} &= \int_{\Omega} \sum_{i=1}^n \frac{\partial}{\partial x_i} \left(\overline{b_{i,j}} \frac{\partial \overline{u}}{\partial x_j} v \right) d\mathbf{x} \\ &= \int_{\Omega} \sum_{i=1}^n \left(\frac{\partial}{\partial x_i} \left(\overline{b_{i,j}} \frac{\partial \overline{u}}{\partial x_j} \right) v + \overline{b_{i,j}} \frac{\partial \overline{u}}{\partial x_j} \frac{\partial v}{\partial x_i} \right) d\mathbf{x} \end{aligned}$$

$$= \sum_{i=1}^n \int_{\Omega} \frac{\partial}{\partial x_i} \left(\overline{b_{i,j}} \frac{\partial \bar{u}}{\partial x_j} \right) v \, d\mathbf{x} + \sum_{i=1}^n \int_{\Omega} \overline{b_{i,j}} \frac{\partial \bar{u}}{\partial x_j} \frac{\partial v}{\partial x_i} \, d\mathbf{x}.$$

Hence,

$$\sum_{i,j=1}^n \int_{\Omega} \overline{b_{i,j}} \frac{\partial \bar{u}}{\partial x_j} \frac{\partial v}{\partial x_i} \, d\mathbf{x} = \int_{\partial\Omega} \sum_{i,j=1}^n \left(\overline{b_{i,j}} \frac{\partial \bar{u}}{\partial x_j} v \nu_i \right) \, d\mathbf{x} - \sum_{i,j=1}^n \int_{\Omega} \frac{\partial}{\partial x_i} \left(\overline{b_{i,j}} \frac{\partial \bar{u}}{\partial x_j} \right) v \, d\mathbf{x}$$

for all $v \in C_b^1(\bar{\Omega})$. It then follows from $B(v, u) = \langle v, f \rangle$ that

$$\langle v, f \rangle_2 = \int_{\Omega} \overline{\left(- \sum_{i,j=1}^n \frac{\partial}{\partial x_i} \left(b_{i,j} \frac{\partial u}{\partial x_j} \right) + b_0 u \right)} v \, d\mathbf{x} + \int_{\partial\Omega} \sum_{i,j=1}^n \left(\overline{b_{i,j}} \frac{\partial \bar{u}}{\partial x_j} v \nu_i \right) \, d\mathbf{x} \quad (12.2.9)$$

for $v \in C_b^1(\bar{\Omega})$. From (12.2.9) with $v \in C_b^1(\Omega)$, we get

$$\langle v, f \rangle_2 = \int_{\Omega} \overline{\left(- \sum_{i,j=1}^n \frac{\partial}{\partial x_i} \left(b_{i,j} \frac{\partial u}{\partial x_j} \right) + b_0 u \right)} v \, d\mathbf{x}$$

for $v \in C_b^1(\Omega)$. Since $C_b^1(\Omega) \supset \mathcal{D}(\Omega)$ is dense in $L^2(\Omega)$, and $L(\mathbf{x}, D)u$ and f are continuous on Ω , we find that

$$L(\mathbf{x}, D)u = - \sum_{i,j=1}^n \frac{\partial}{\partial x_i} \left(b_{i,j} \frac{\partial u}{\partial x_j} \right)(\mathbf{x}) + b_0(\mathbf{x})u(\mathbf{x}) = f(\mathbf{x}) \quad (12.2.10)$$

for $\mathbf{x} \in \Omega$. This means that the integrals over Ω in (12.2.9), including $\langle v, f \rangle = \int_{\Omega} v f \, d\mathbf{x}$, do not depend on the value of $v \in C_b^1(\bar{\Omega})$ on the boundary $\partial\Omega$ (a set of measure zero). For (12.2.9) to be true for all $v \in C_b^1(\bar{\Omega})$, we must have that

$$\int_{\partial\Omega} \sum_{i,j=1}^n \left(\overline{b_{i,j}} \frac{\partial \bar{u}}{\partial x_j} v \nu_i \right) \, d\mathbf{x} = 0$$

for all $v \in C_b^1(\bar{\Omega})$. Finally, since $C_b^1(\bar{\Omega})|_{\partial\Omega}$ is dense in $L^2(\partial\Omega)$, and the $b_{i,j}$ and $\frac{\partial u}{\partial x_j}$ are continuous on $\partial\Omega$, we get the boundary condition

$$\sum_{i,j=1}^n b_{i,j}(\mathbf{x}) \frac{\partial u}{\partial x_j}(\mathbf{x}) \nu_i(\mathbf{x}) = 0 \quad (12.2.11)$$

for all $\mathbf{x} \in \partial\Omega$. So u is the classical solution of (12.2.3).

In Example 12.5.5 below, we will illustrate for the Laplace operator why no boundary conditions are needed for the variational formulation.

12.3 Adjoint Boundary Conditions

For the Dirichlet and Neumann problems presented in the previous section, the boundary conditions were such that the variational formulation of the partial differential equation did not involve any boundary conditions. This is not always possible or desirable.

In this section, we present a general theory to deduce the variational problem with its adjoint boundary conditions associated to an elliptic partial differential equation. However, as we have seen in previous examples and as we will see in future examples, it is often easier to directly deduce the variational problem with its adjoint boundary conditions without referring to the general results of the present section.

We assume that Ω is a bounded open subset of $\mathbb{R}^n$ satisfying at least a strong m -extension property with $m \geq k$, and X is a subspace of $H^{k,2}(\Omega)$ containing $H_0^{k,2}(\Omega)$. We consider the elliptic partial differential equation

$$L(\mathbf{x}, D)u = \sum_{|\alpha|, |\beta| \leq k} (-1)^{|\alpha|} D^\alpha (b_{\alpha, \beta}(\mathbf{x}) D^\beta u) = f$$

on Ω , where the coefficients $b_{\alpha, \beta}$ are all of class $C_b^\infty(\Omega)$ and $f \in L^2(\Omega)$.

Definition 12.3.1

A set $\{P_j(\mathbf{x}, D)\}_{j \in J}$ of differential operators defined on an open set containing $\partial\Omega$, where J is some index set, is **normal** on $\partial\Omega$ if

1. $P_j(\mathbf{x}, D)$ is a differential operator of exactly order j for all j .
2. For each j , $\partial\Omega$ is not a characteristic surface of the operator $P_j(\mathbf{x}, D)$.

The best known normal set of operators for a partial differential equation of order $m = 2k$ in divergence form is $\left\{ \frac{\tilde{\partial}^j}{\tilde{\partial} \nu^j} \right\}_{j=0}^{k-1}$ in a tubular neighbourhood O_μ of $\partial\Omega$, where $\nu(\mathbf{x})$ is the outward unit normal to $\partial\Omega$ at $\mathbf{x} \in \partial\Omega$. Recall that we have defined $\frac{\tilde{\partial} u}{\tilde{\partial} \nu}$ for $u : \Omega_\mu \rightarrow \mathbb{R}$ in Section 9.9.6 as

$$\frac{\tilde{\partial} u}{\tilde{\partial} \nu}(\mathbf{z}_t) \equiv \frac{d}{dt} u(\mathbf{x} + t\nu(\mathbf{x})) = \nabla u(\mathbf{z}_t) \cdot \nu(\mathbf{x}) = \nabla u(\mathbf{x} + t\nu(\mathbf{x})) \cdot \nu(\mathbf{x})$$

for $\mathbf{z}_t = \mathbf{x} + t\nu(\mathbf{x})$ with $\mathbf{x} \in \partial\Omega$ and $|t| < \mu$.

Theorem 12.3.2

Let

$$\begin{aligned} B : H^{k,2}(\Omega) \times H^{k,2}(\Omega) &\rightarrow \mathbb{C} \\ (v, u) &\mapsto \sum_{|\alpha|, |\beta| \leq k} \langle D^\alpha v, b_{\alpha, \beta}(\mathbf{x}) D^\beta u \rangle_2 \end{aligned} \quad (12.3.1)$$

be the Dirichlet form associated to $L(\mathbf{x}, D)$. Given a normal set of differential operators $\{P_j(\mathbf{x}, D)\}_{j=0}^{k-1}$, there exists a set of differential operators $\{\tilde{P}_j(\mathbf{x}, D)\}_{j=k}^{2k-1}$ such that

$$B(v, u) - \langle v, L(\mathbf{x}, D)u \rangle_2 = \sum_{j=0}^{k-1} \int_{\partial\Omega} P_j(\mathbf{x}, D)v(\mathbf{x}) \overline{\tilde{P}_{2k-1-j}(\mathbf{x}, D)u(\mathbf{x})} dS$$

for all $u, v \in C^{2k}(\overline{\Omega})$. Moreover, if B is coercive, the set of differential operators $\{\tilde{P}_j(\mathbf{x}, D)\}_{j=k}^{2k-1}$ is normal.

A long proof of this theorem is given in [10].

Theorem 12.3.3

Suppose that Ω is a bounded open subset of $\mathbb{R}^n$ satisfying at a strong k -extension property. Let $\{P_j(\mathbf{x}, D)\}_{j=0}^{k-1}$ be a normal set of differential operators associated to $L(\mathbf{x}, D)u = f$. Given $g_j \in C^\infty(\partial\Omega)$ with $0 \leq j < k$, there exists $g \in C_c^\infty(\mathbb{R}^n)$ such that $P_j(\mathbf{x}, D)g = g_j$ on $\partial\Omega$ for $0 \leq j < k$.

Proof.

i) Suppose that $P_j(\mathbf{x}, D) = \frac{\tilde{\partial}^j}{\tilde{\partial}\nu^j}$ on a tubular neighbourhood O_ϵ of $\partial\Omega$ as given by Theorem 9.9.8. Choose $\phi \in C_c^\infty(\mathbb{R})$ such that $\text{supp } \phi \subset]-\epsilon/2, \epsilon/2[$ and $\phi(x) = 1$ for $x \in]-\epsilon/4, \epsilon/4[$.

Consider the function $g : \mathbb{R}^n \rightarrow \mathbb{R}$ defined by

$$g(\mathbf{x}) = \begin{cases} \phi(t) \sum_{j=0}^k \frac{t^j}{j!} g_j(\mathbf{y}) & \text{if } \mathbf{x} = \mathbf{y} + t\nu(\mathbf{y}) \text{ with } \mathbf{y} \in \partial\Omega \text{ and } -\epsilon < t < \epsilon \\ 0 & \text{if } \mathbf{x} \notin O_\epsilon \end{cases}$$

We have that $g \in C_c^\infty(\mathbb{R}^n)$, $\text{supp } g \subset O_\epsilon$ and

$$(P_j(\mathbf{x}, D)g)(\mathbf{y}) = \frac{\tilde{\partial}^j g}{\tilde{\partial}\nu^j}(\mathbf{y}) = \frac{\partial^j g}{\partial t^j}(\mathbf{y} + t\nu(\mathbf{y})) \Big|_{t=0} = g_j(\mathbf{y})$$

for $\mathbf{y} \in \partial\Omega$.

ii) The general form of $P_j(\mathbf{x}, D)$ is

$$P_j(\mathbf{x}, D) = \sum_{i=0}^j Q_{j,i}(\mathbf{x}, D) \frac{\partial^i}{\partial \nu^i},$$

where $Q_{j,i}(\mathbf{x}, D)$ is a differential operator involving only directional derivatives in directions tangent to $\partial\Omega$ and of order smaller than or equal to $j-i$. In particular, $Q_{j,j}(\mathbf{x}, D)$ is simply a function $q_j : \mathbb{R}^n \rightarrow \mathbb{R}$ such that $q_j(\mathbf{x}) \neq 0$ for all $\mathbf{x} \in \partial\Omega$ because $\partial\Omega$ is not a characteristic surface for the operator $P_j(\mathbf{x}, D)$ by definition of a normal set of differential operators.

Hence $P_j(\mathbf{x}, D)g = g_j$ for $0 \leq j < k$ is equivalent to $\frac{\partial^j g}{\partial \nu^j} = f_j$ for $0 \leq j < k$, where the f_j are given recursively by $f_0 = \frac{g_0}{q_0}$ and $f_j = \frac{1}{q_j} \left(g_j - \sum_{i=0}^{j-1} Q_{j,i}(\mathbf{x}, D)f_i \right)$ for $j = 1, 2, \dots, k-1$.

The problem is reduced to (i); namely, to find g such that $\frac{\partial^j g}{\partial \nu^j} = f_j$ for $0 \leq j < k$. ■

The following examples were suggested in [10].

Example 12.3.4

Consider the partial differential equation

$$L(\mathbf{x}, D)u(\mathbf{x}) = \sum_{|\alpha|, |\beta| \leq k} (-1)^{|\alpha|} D^\alpha (b_{\alpha, \beta}(\mathbf{x}) D^\beta u(\mathbf{x})) = f(\mathbf{x}) \quad , \quad \mathbf{x} \in \Omega \quad , \quad (12.3.2)$$

where no boundary conditions are given. To find the **free or natural boundary conditions** associated to this problem, consider the normal family of differential operators $\left\{ \frac{\tilde{\partial}^j}{\tilde{\partial} \nu^j} \right\}_{j=0}^{k-1}$. According to Theorem 12.3.1, there exists a set of differential operators $\{\tilde{P}_j\}_{j=k}^{2k-1}$ such that

$$B(v, u) - \langle v, L(\mathbf{x}, D)u \rangle = \sum_{j=0}^{k-1} \int_{\partial\Omega} \frac{\tilde{\partial}^j v}{\tilde{\partial} \nu^j}(\mathbf{x}) \overline{\tilde{P}_{2k-1-j}(u)(\mathbf{x})} dS \quad (12.3.3)$$

for all $v, u \in C_c^{2k}(\overline{\Omega})$, where B is defined in (12.3.1). Hence the variational formulation of $L(\mathbf{x}, D)u = f$ is to find $u \in H^{k,2}(\Omega)$ such that

$$B(v, u) = \langle v, f \rangle \quad (12.3.4)$$

for all $v \in H^{k,2}(\Omega)$ with

$$\tilde{P}_j(u)(\mathbf{x}) = 0 \quad , \quad \mathbf{x} \in \partial\Omega \text{ and } k \leq j < 2k \quad . \quad (12.3.5)$$

It follows from (12.3.3) that a solution u of (12.3.4) and (12.3.5) is a solution of

$$\langle v, L(\mathbf{x}, D)u \rangle = \langle v, f \rangle$$

for all $v \in H^{k,2}(\Omega)$. One can show that the solution u is in fact in $H^{2k,2}(\Omega)$ if f is nice enough¹, and so the previous free boundary conditions $\tilde{P}_j(u) = 0$ on $\partial\Omega$ for $k \leq j < 2k$ make sense. ♣

Example 12.3.5

We may create more complex Dirichlet problems than those mentioned before. Consider a normal family of differential operators $\{P_j\}_{j=0}^{k-1}$ for $L(\mathbf{x}, D)u = f$ in Ω , where $f \in L^2(\Omega)$.

Let J_1 and J_2 be two non-empty, disjoint subsets of $J = \{0, 1, 2, \dots, k-1\}$ such that $J = J_1 \cup J_2$. Our goal is to find the variational formulation of

$$\begin{aligned} L(\mathbf{x}, D)u(\mathbf{x}) &= f(\mathbf{x}) \quad , \quad \mathbf{x} \in \Omega \quad , \\ P_j(u)(\mathbf{x}) &= 0 \quad , \quad \mathbf{x} \in \partial\Omega \text{ and } j \in J_1 \quad , \end{aligned} \quad (12.3.6)$$

¹The main result of Section 12.6.2 on regularity at the boundary may be used when $k = 1$.

where the linear partial differential operator $L(\mathbf{x}, D)$ is defined in (12.3.2). Consider the space X defined as the closure in $H^{k,2}(\Omega)$ (always with respect to the norm $\|\cdot\|_{k,2}$) of the set

$$S = \{u \in C^\infty(\overline{\Omega}) : P_j(u) = 0 \text{ on } \partial\Omega \text{ for } j \in J_1\}.$$

According to Theorem 12.3.1, there exists a set of differential operators $\{\tilde{P}_j\}_{j=k}^{2k-1}$ such that

$$B(v, u) - \langle v, L(\mathbf{x}, D)u \rangle = \sum_{j \in J} \int_{\partial\Omega} P_j(v)(\mathbf{x}) \overline{\tilde{P}_{2k-1-j}(u)(\mathbf{x})} dS$$

for all $v, u \in C^\infty(\overline{\Omega})$. Hence

$$B(v, u) - \langle v, L(\mathbf{x}, D)u \rangle = \sum_{j \in J_2} \int_{\partial\Omega} P_j(v)(\mathbf{x}) \overline{\tilde{P}_{2k-1-j}(u)(\mathbf{x})} dS \quad (12.3.7)$$

for all $v, u \in X \cap C^\infty(\overline{\Omega})$. Given $u \in X \cap C^\infty(\overline{\Omega})$, we may use Theorem 12.3.3 to find $v \in C^\infty(\overline{\Omega})$ such that $P_j(v) = 0$ for $j \in J_1$ and $P_j(v) = \tilde{P}_{2k-1-j}(u)$ for $j \in J_2$. However, to get the variational formulation, the right hand side of (12.3.7) must be null for all $v \in X \cap C^\infty(\overline{\Omega})$. Therefore, we must require that $\tilde{P}_{2k-1-j}(u) = 0$ on $\partial\Omega$ for all $j \in J_2$ to get the variational formulation. Hence, the variational formulation of (12.3.6) is to find $u \in X$ such that

$$B(v, u) = \langle v, f \rangle \quad (12.3.8)$$

for all $v \in X$ with

$$\tilde{P}_{2k-1-j}(u)(\mathbf{x}) = 0 \quad , \quad \mathbf{x} \in \partial\Omega \text{ and } j \in J_2. \quad (12.3.9)$$

It follows from (12.3.7) that a solution u of (12.3.8) and (12.3.9) is a solution of

$$\langle v, L(\mathbf{x}, D)u \rangle = \langle v, f \rangle$$

for all $v \in X$. As for the previous example, the missing step is to show that the solution u is in fact in $H^{2k,2}(\Omega)$ and so the boundary conditions on $\partial\Omega$ for $k \leq j < 2k$ make sense. ♣

12.4 Weakly Coercive Bilinear Mappings

As usual, let Ω be a bounded open subset of $\mathbb{R}^n$ with a sufficiently smooth boundary. Recall that a bounded sesquilinear mapping $B : H^{k,2}(\Omega) \times H^{k,2}(\Omega) \rightarrow \mathbb{C}$ is coercive if there exists a constant $C > 0$ such that $\operatorname{Re} B(f, f) \geq C\|f\|_{k,2}^2$ for all $f \in H^{k,2}(\Omega)$.

Definition 12.4.1

Let $H = H^{k,2}(\Omega)$ or $H = H_0^{k,2}(\Omega)$. We say that a bounded sesquilinear mapping $B : H \times H \rightarrow \mathbb{C}$ is **weakly coercive** if there exist two constants $C > 0$ and $\lambda > 0$ such that $\operatorname{Re} B(f, f) \geq C\|f\|_{k,2,\Omega}^2 - \lambda\|f\|_{2,\Omega}^2$ for $f \in H$.

The main result of this section is Gårding's inequality. Before stating and proving this result, we need several lemmas.

Lemma 12.4.2

For $r < s < t$ and $\epsilon > 0$, there exists $C > 0$ such that

$$\|f\|_{s,\rho}^2 \leq \epsilon \|f\|_{t,\rho}^2 + C \|f\|_{r,\rho}^2$$

for $f \in H^{t,2}(\mathbb{R}^n)$. We also have for $r < s < t$, non-negative integers, that there exists $C > 0$ such that

$$\|f\|_{s,2,\mathbb{R}^n}^2 \leq \epsilon \|f\|_{t,2,\mathbb{R}^n}^2 + C \|f\|_{r,2,\mathbb{R}^n}^2$$

for $f \in H^{t,2}(\mathbb{R}^n)$.

Proof.

i) Choose $R > 0$ such that $(1 + r^2)^s < \epsilon(1 + r^2)^t$ for $r > R$. This is possible because

$$\lim_{r \rightarrow \infty} \frac{(1 - r^2)^s}{(1 - r^2)^t} = 0$$

for $t > s$. Let $C = (1 + R^2)^{s-r}$. Then

$$(1 + \|\mathbf{y}\|^2)^s \leq C(1 + \|\mathbf{y}\|^2)^r \leq \epsilon(1 + \|\mathbf{y}\|^2)^t + C(1 + \|\mathbf{y}\|^2)^r$$

for $\|\mathbf{y}\| \leq R$ and

$$(1 + \|\mathbf{y}\|^2)^s \leq \epsilon(1 + \|\mathbf{y}\|^2)^t \leq \epsilon(1 + \|\mathbf{y}\|^2)^t + C(1 + \|\mathbf{y}\|^2)^r$$

for $\|\mathbf{y}\| > R$. It follows that

$$\int_{\mathbb{R}^n} |\hat{f}(\mathbf{y})|^2 (1 + \|\mathbf{y}\|^2)^s d\mathbf{y} \leq \epsilon \int_{\mathbb{R}^n} |\hat{f}(\mathbf{y})|^2 (1 + \|\mathbf{y}\|^2)^t d\mathbf{y} + C \int_{\mathbb{R}^n} |\hat{f}(\mathbf{y})|^2 (1 + \|\mathbf{y}\|^2)^r d\mathbf{y}$$

for $f \in H^{t,2}(\mathbb{R}^n)$.

ii) Suppose that $r < s < t$ are non-negative integers. According to Theorem 11.1.3 about the equivalence of the Sobolev norms, there exist constants Q_1 , Q_2 and Q_3 such that $\|\cdot\|_{s,2} \leq Q_1 \|\cdot\|_{s,\rho}$, $\|\cdot\|_{t,\rho} \leq Q_2 \|\cdot\|_{t,2}$ and $\|\cdot\|_{t,\rho} \leq Q_3 \|\cdot\|_{r,2}$.

If we apply (i) with ϵ replaced by $\epsilon/(Q_1 Q_2)$, we find that there exists a constant C_1 such that

$$\|f\|_{s,\rho}^2 \leq \frac{\epsilon}{Q_1 Q_2} \|f\|_{t,\rho}^2 + C_1 \|f\|_{r,\rho}^2$$

for $f \in H^{t,2}(\mathbb{R}^n)$. Hence,

$$\frac{1}{Q_1} \|f\|_{s,2,\mathbb{R}^n}^2 \leq \frac{\epsilon}{Q_1} \|f\|_{t,2,\mathbb{R}^n}^2 + C_1 Q_3 \|f\|_{r,2,\mathbb{R}^n}^2$$

for $f \in H^{t,2}(\mathbb{R}^n)$. We then get that

$$\|f\|_{s,2,\mathbb{R}^n}^2 \leq \epsilon \|f\|_{t,2,\mathbb{R}^n}^2 + C \|f\|_{r,2,\mathbb{R}^n}^2$$

for $f \in H^{t,2}(\mathbb{R}^n)$, where $C = C_1 Q_1 Q_3$. ■

Lemma 12.4.3

Let Ω be a bounded open subset of $\mathbb{R}^n$. Suppose that

$$L(\mathbf{x}, D)u = \sum_{|\alpha|, |\beta|=k} (-1)^\alpha D^\alpha (b_{\alpha, \beta} D^\beta u)$$

is strongly elliptic on $\overline{\Omega}$, where $b_{\alpha, \beta} \in \mathbb{C}$ for all α and β . Then

$$\begin{aligned} B : H_0^{k,2}(\Omega) \times H_0^{k,2}(\Omega) &\rightarrow \mathbb{C} \\ (v, u) &\mapsto \sum_{|\alpha|, |\beta|=k} \langle D^\alpha v, b_{\alpha, \beta} D^\beta u \rangle_2 \end{aligned}$$

is coercive. Namely, there exists $K > 0$ such that

$$\operatorname{Re} B(u, u) = \operatorname{Re} \sum_{|\alpha|, |\beta|=k} \langle D^\alpha u, b_{\alpha, \beta} D^\beta u \rangle_2 \geq K \|u\|_{k,2,\Omega}^2$$

for $u \in H_0^{k,2}(\Omega)$.

Proof.

For $u \in H_0^{k,2}(\Omega)$, let $\underline{u} : \mathbb{R}^n \rightarrow \mathbb{C}$ be the function defined by

$$\underline{u}(\mathbf{x}) = \begin{cases} u(\mathbf{x}) & \text{if } \mathbf{x} \in \Omega \\ 0 & \text{if } \mathbf{x} \in \mathbb{R}^n \setminus \Omega \end{cases}$$

From Proposition 11.2.2, we have that $\underline{u} \in H^{k,2}(\mathbb{R}^n)$. Using Plancherel theorem, Theorem 6.6.6, we may write

$$\begin{aligned} B(u, u) &= \sum_{|\alpha|, |\beta|=k} \overline{b_{\alpha, \beta}} \int_{\Omega} \overline{D^\beta u(\mathbf{x})} D^\alpha u(\mathbf{x}) d\mathbf{x} = \sum_{|\alpha|, |\beta|=k} \overline{b_{\alpha, \beta}} \int_{\mathbb{R}^n} \overline{D^\beta \underline{u}(\mathbf{x})} D^\alpha \underline{u}(\mathbf{x}) d\mathbf{x} \\ &= \sum_{|\alpha|, |\beta|=k} \overline{b_{\alpha, \beta}} \int_{\mathbb{R}^n} \overline{(D^\beta \underline{u})^\wedge(\mathbf{x})} (D^\alpha \underline{u})^\wedge(\mathbf{x}) d\mathbf{x} = \sum_{|\alpha|, |\beta|=k} \overline{b_{\alpha, \beta}} \int_{\mathbb{R}^n} (i\mathbf{x})^\beta (-i\mathbf{x})^\alpha |\hat{\underline{u}}(\mathbf{x})|^2 d\mathbf{x} \\ &= \int_{\mathbb{R}^n} \left(\sum_{|\alpha|, |\beta|=k} \overline{b_{\alpha, \beta}} \mathbf{x}^\beta \mathbf{x}^\alpha \right) |\hat{\underline{u}}(\mathbf{x})|^2 d\mathbf{x} \end{aligned}$$

for $u \in H_0^{k,2}(\Omega)$. Since $L(\mathbf{x}, D)$ is strongly elliptic on $\overline{\Omega}$, there exists $\theta_1 > 0$ such that

$$\operatorname{Re} Q(\mathbf{x}, \boldsymbol{\xi}) = \operatorname{Re} \sum_{|\alpha|, |\beta|=k} b_{\alpha, \beta} \boldsymbol{\xi}^\beta \boldsymbol{\xi}^\alpha \geq \theta_1 \|\boldsymbol{\xi}\|^{2k} \quad (12.4.1)$$

for $\boldsymbol{\xi} \in \mathbb{R}^n$. Thus,

$$\begin{aligned} \operatorname{Re} B(u, u) &\geq \theta_1 \int_{\mathbb{R}^n} \|\mathbf{x}\|^{2k} |\hat{\underline{u}}(\mathbf{x})|^2 d\mathbf{x} \geq \theta_2 \sum_{|\alpha|=k} \int_{\mathbb{R}^n} |\mathbf{x}^\alpha|^2 |\hat{\underline{u}}(\mathbf{x})|^2 d\mathbf{x} \\ &= \theta_2 \sum_{|\alpha|=k} \int_{\mathbb{R}^n} |(D^\alpha \underline{u})^\wedge(\mathbf{x})|^2 d\mathbf{x} = \theta_2 \sum_{|\alpha|=k} \int_{\mathbb{R}^n} |D^\alpha \underline{u}(\mathbf{x})|^2 d\mathbf{x} \geq \theta_3 \|\underline{u}\|_{k,2,\mathbb{R}^n}^2 \end{aligned}$$

for $u \in H_0^{k,2}(\Omega)$, where we have used (12.4.1) for the first inequality and Plancherel theorem for the second equality. The constant θ_3 for the last inequality is $\theta_3 = \theta_2 C$ where C comes from the Poincaré's inequality, Theorem 11.7.11. Moreover, for the second inequality, we have used the relation

$$\|\mathbf{x}\|^{2k} = \left(\sum_{j=1}^n x_j^2 \right)^k \geq C \left(\sum_{|\alpha|=k} \mathbf{x}^{2\alpha} \right) = C \left(\sum_{|\alpha|=k} |\mathbf{x}^\alpha|^2 \right)$$

for $\mathbf{x} \in \mathbb{R}^n$, where $1/C = \max_{\|\mathbf{x}\|=1} \sum_{|\alpha|=k} \mathbf{x}^{2\alpha}$, and taken $\theta_2 = C\theta_1$.

Since $\underline{u}(\mathbf{x}) = 0$ for $\mathbf{x} \in \mathbb{R}^n \setminus \Omega$, we get $\operatorname{Re} B(u, u) \geq \theta_3 \|u\|_{k,2,\Omega}^2$ for all $u \in H_0^{k,2}(\Omega)$ with $\theta_3 > 0$. We take $K = \theta_3$ to get the conclusion of the lemma. ■

Lemma 12.4.4

Let Ω be a bounded open subset of $\mathbb{R}^n$. Suppose that

$$L(\mathbf{x}, D)u = \sum_{|\alpha|, |\beta|=k} (-1)^\alpha D^\alpha (b_{\alpha,\beta}(\mathbf{x}) D^\beta u)$$

is strongly elliptic on $\overline{\Omega}$, where $b_{\alpha,\beta} \in C(\overline{\Omega})$ for all α and β . Then

$$\begin{aligned} B : H_0^{k,2}(\Omega) \times H_0^{k,2}(\Omega) &\rightarrow \mathbb{R} \\ (v, u) &\mapsto \sum_{|\alpha|, |\beta|=k} \langle D^\alpha v, b_{\alpha,\beta}(\mathbf{x}) D^\beta u \rangle_2 \end{aligned}$$

is weakly coercive.

Proof.

i) We prove that there exists a constant $C_1 > 0$ such that $\operatorname{Re} B(u, u) \geq C_1 \|u\|_{k,2,\Omega}^2$ for all $u \in H_0^{k,2}(\Omega)$ such that $\operatorname{supp} u$ is small enough.

Given $\mathbf{z} \in \Omega$, we write $B(u, u)$ as

$$B(u, u) = \sum_{|\alpha|, |\beta|=k} \langle D^\alpha u, (b_{\alpha,\beta}(\mathbf{x}) - b_{\alpha,\beta}(\mathbf{z})) D^\beta u \rangle_2 + \sum_{|\alpha|, |\beta|=k} \langle D^\alpha u, b_{\alpha,\beta}(\mathbf{z}) D^\beta u \rangle_2 \quad (12.4.2)$$

for $u \in H_0^{k,2}(\Omega)$. Since $L(\mathbf{x}, D)$ is strongly elliptic on $\overline{\Omega}$, there exists $\theta_1 > 0$ such that

$$\operatorname{Re} Q(\mathbf{z}, \boldsymbol{\xi}) = \operatorname{Re} \sum_{|\alpha|, |\beta|=k} b_{\alpha,\beta}(\mathbf{z}) \boldsymbol{\xi}^\beta \boldsymbol{\xi}^\alpha \geq \theta_1 \|\boldsymbol{\xi}\|^{2k} \quad (12.4.3)$$

for all $\boldsymbol{\xi} \in \mathbb{R}^n$ and $\mathbf{z} \in \overline{\Omega}$. It follows that θ_1 in (12.4.1) in the proof of Lemma 12.4.3 may be assumed to be independent of $\mathbf{z} \in \overline{\Omega}$. Therefore, we may assume that the constant $K > 0$ for the coercive property in Lemma 12.4.3 is independent of $\mathbf{z} \in \overline{\Omega}$; namely,

$$\operatorname{Re} \sum_{|\alpha|, |\beta|=k} \langle D^\alpha u, b_{\alpha,\beta}(\mathbf{z}) D^\beta u \rangle_2 \geq K \|u\|_{k,2,\Omega}^2 \quad (12.4.4)$$

for $u \in H_0^{k,2}(\Omega)$ and $\mathbf{z} \in \Omega$. Since the $b_{\alpha,\beta}$ are continuous on $\overline{\Omega}$ by hypothesis and $\overline{\Omega}$ is compact because Ω is bounded, we have that the $b_{\alpha,\beta}$ are uniformly continuous on $\overline{\Omega}$. Thus, there exists $\delta > 0$ such that $\sum_{|\alpha|,|\beta|=k} |b_{\alpha,\beta}(\mathbf{x}_1) - b_{\alpha,\beta}(\mathbf{x}_2)| < \frac{K}{2}$ for $|\mathbf{x}_1 - \mathbf{x}_2| < \delta$ and $\mathbf{x}_1, \mathbf{x}_2 \in \overline{\Omega}$.

Hence, for all $\mathbf{z} \in \Omega$ and all $u \in H_0^{k,2}(\Omega)$ with $\text{supp } u \subset B_\delta(\mathbf{z}) \cap \Omega$, we have

$$\begin{aligned}
& \left| \sum_{|\alpha|,|\beta|=k} \int_{\Omega} \overline{(b_{\alpha,\beta}(\mathbf{x}) - b_{\alpha,\beta}(\mathbf{z}))} D^\beta u(\mathbf{x}) D^\alpha u(\mathbf{x}) d\mathbf{x} \right| \\
& \leq \sum_{|\alpha|,|\beta|=k} \int_{\Omega} |b_{\alpha,\beta}(\mathbf{x}) - b_{\alpha,\beta}(\mathbf{z})| |D^\beta u(\mathbf{x})| |D^\alpha u(\mathbf{x})| d\mathbf{x} \\
& \leq \sup_{\mathbf{x} \in B_\delta(\mathbf{z}) \cap \Omega} \left(\sum_{|\alpha|,|\beta|=k} |b_{\alpha,\beta}(\mathbf{x}) - b_{\alpha,\beta}(\mathbf{z})| \right) \int_{\Omega} |D^\beta u(\mathbf{x})| |D^\alpha u(\mathbf{x})| d\mathbf{x} \\
& \leq \sup_{\mathbf{x} \in B_\delta(\mathbf{z}) \cap \Omega} \left(\sum_{|\alpha|,|\beta|=k} |b_{\alpha,\beta}(\mathbf{x}) - b_{\alpha,\beta}(\mathbf{z})| \right) \left(\int_{\Omega} |D^\beta u(\mathbf{x})|^2 d\mathbf{x} \right)^{1/2} \left(\int_{\Omega} |D^\alpha u(\mathbf{x})|^2 d\mathbf{x} \right)^{1/2} \\
& < \frac{K}{2} \|u\|_{k,2,\Omega}^2. \tag{12.4.5}
\end{aligned}$$

It follows from (12.4.2), (12.4.4) and (12.4.5) that

$$\text{Re } B(u, u) \geq \frac{K}{2} \|u\|_{k,2,\Omega}^2 \tag{12.4.6}$$

for all $u \in H_0^{k,2}(\Omega)$ such that there exists $\mathbf{z} \in \Omega$ with $\text{supp } u \subset B_\delta(\mathbf{z}) \cap \Omega$. Let $C_1 = K/2$.

ii) We now prove the lemma. Since $\overline{\Omega}$ is compact, there exists a finite subset $\{\mathbf{z}_j\}_{j=1}^J$ of Ω such that $\{B_\delta(\mathbf{z}_j)\}_{j=1}^J$ is an open cover of $\overline{\Omega}$. Let $\{\phi_j\}_{j=1}^J$ be a partition of unity subordinate to this open cover. Since $\sum_{j=1}^J \phi(\mathbf{x}) = 1$ for all $\mathbf{x} \in \overline{\Omega}$, we have that $\sum_{j=1}^J |\phi(\mathbf{x})|^2 > 0$ for all $\mathbf{x} \in \overline{\Omega}$. Let $\psi_j(\mathbf{x}) = \phi_j(\mathbf{x}) \left(\sum_{j=1}^J |\phi(\mathbf{x})|^2 \right)^{-1}$ for $\mathbf{x} \in \overline{\Omega}$. We have that $\psi_j \in C^\infty(\overline{\Omega})$ for all j and $\sum_{j=1}^J |\psi_j(\mathbf{x})|^2 = 1$ for all $\mathbf{x} \in \overline{\Omega}$. Hence,

$$\begin{aligned}
B(u, u) &= \sum_{|\alpha|,|\beta|=k} \left\langle D^\alpha u, \sum_{j=1}^J |\psi_j(\mathbf{x})|^2 b_{\alpha,\beta}(\mathbf{x}) D^\beta u \right\rangle_2 = \sum_{|\alpha|,|\beta|=k} \sum_{j=1}^J \langle D^\alpha u, |\psi_j(\mathbf{x})|^2 b_{\alpha,\beta}(\mathbf{x}) D^\beta u \rangle_2 \\
&= \sum_{|\alpha|,|\beta|=k} \sum_{j=1}^J \langle \psi_j D^\alpha u, b_{\alpha,\beta}(\mathbf{x}) \psi_j D^\beta u \rangle_2 = Q_1(u) + Q_2(u)
\end{aligned}$$

for $u \in H_0^{k,2}(\Omega)$, where

$$Q_1(u) = \sum_{|\alpha|,|\beta|=k} \sum_{j=1}^J \langle D^\alpha (\psi_j u), b_{\alpha,\beta}(\mathbf{x}) D^\beta (\psi_j u) \rangle_2$$

and

$$\begin{aligned} Q_2(u) = & \sum_{|\alpha|, |\beta|=k} \sum_{j=1}^J \langle D^\alpha(\psi_j u), b_{\alpha, \beta}(\mathbf{x}) (\psi_j D^\beta u - D^\beta(\psi_j u)) \rangle_2 \\ & + \sum_{|\alpha|, |\beta|=k} \sum_{j=1}^J \langle \psi_j D^\alpha u - D^\alpha(\psi_j u), b_{\alpha, \beta}(\mathbf{x}) \psi_j D^\beta u \rangle_2 . \end{aligned}$$

Since $\psi_j u \in H_0^{k,2}(\Omega)$ has compact support in $B_\delta(\mathbf{z}_j)$ for all j , we have from part (i) that

$$\operatorname{Re} Q_1(u) \geq C_1 \sum_{j=1}^J \|\psi_j u\|_{k,2,\Omega}^2 . \quad (12.4.7)$$

However,

$$\begin{aligned} \sum_{j=1}^J \|\psi_j u\|_{k,2,\Omega}^2 & \geq \sum_{j=1}^J \sum_{|\alpha|=k} \langle D^\alpha(\psi_j u), D^\alpha(\psi_j u) \rangle_2 \\ & = \sum_{j=1}^J \sum_{|\alpha|=k} \langle \psi_j D^\alpha u, \psi_j D^\alpha u \rangle_2 + \sum_{j=1}^J \sum_{|\alpha|=k} \langle D^\alpha(\psi_j u), D^\alpha(\psi_j u) - \psi_j D^\alpha u \rangle_2 \\ & \quad + \sum_{j=1}^J \sum_{|\alpha|=k} \langle D^\alpha(\psi_j u) - \psi_j D^\alpha u, \psi_j D^\alpha u \rangle_2 , \end{aligned} \quad (12.4.8)$$

where

$$\sum_{j=1}^J \sum_{|\alpha|=k} \langle \psi_j D^\alpha u, \psi_j D^\alpha u \rangle_2 = \sum_{|\alpha|=k} \left\langle D^\alpha u, \sum_{j=1}^J |\psi_j|^2 D^\alpha u \right\rangle_2 = \sum_{|\alpha|=k} \langle D^\alpha u, D^\alpha u \rangle_2 \geq C_2 \|u\|_{k,2,\Omega}^2$$

for $u \in H_0^{k,2}(\Omega)$ and some constant $C_2 > 0$ provided by the Poincaré's inequality, Theorem 11.7.11, and

$$\begin{aligned} & \left| \sum_{j=1}^J \sum_{|\alpha|=k} \langle D^\alpha(\psi_j u), D^\alpha(\psi_j u) - \psi_j D^\alpha u \rangle_2 + \sum_{j=1}^J \sum_{|\alpha|=k} \langle D^\alpha(\psi_j u) - \psi_j D^\alpha u, \psi_j D^\alpha u \rangle_2 \right| \\ & \leq \sum_{\substack{|\alpha|, |\beta| \leq k \\ |\alpha|+|\beta| < 2k}} \int_{\Omega} |A_{\alpha, \beta}(\mathbf{x})| |D^\alpha u(\mathbf{x})| |D^\beta u(\mathbf{x})| \, d\mathbf{x} \\ & \leq \sum_{\substack{|\alpha|, |\beta| \leq k \\ |\alpha|+|\beta| < 2k}} \sup_{\mathbf{x} \in \bar{\Omega}} |A_{\alpha, \beta}(\mathbf{x})| \int_{\Omega} |D^\alpha u(\mathbf{x})| |D^\beta u(\mathbf{x})| \, d\mathbf{x} \\ & \leq \sum_{\substack{|\alpha|, |\beta| \leq k \\ |\alpha|+|\beta| < 2k}} \sup_{\mathbf{x} \in \bar{\Omega}} |A_{\alpha, \beta}(\mathbf{x})| \left(\int_{\Omega} |D^\alpha u(\mathbf{x})|^2 \, d\mathbf{x} \right)^{1/2} \left(\int_{\Omega} |D^\beta u(\mathbf{x})|^2 \, d\mathbf{x} \right)^{1/2} \\ & \leq \sum_{\substack{|\alpha|, |\beta| \leq k \\ |\alpha|+|\beta| < 2k}} \sup_{\mathbf{x} \in \bar{\Omega}} |A_{\alpha, \beta}(\mathbf{x})| \|u\|_{k,2,\Omega} \|u\|_{k-1,2,\Omega} \leq C_3 \|u\|_{k,2,\Omega} \|u\|_{k-1,2,\Omega} \end{aligned}$$

for $u \in H_0^{k,2}(\Omega)$, where

$$C_3 = \sum_{\substack{|\alpha|, |\beta| \leq k \\ |\alpha| + |\beta| < 2k}} \sup_{\mathbf{x} \in \bar{\Omega}} |A_{\alpha, \beta}(\mathbf{x})| \geq 0.$$

Note that $C_3 < \infty$ because the $A_{\alpha, \beta}$ are polynomial expressions in $D^\gamma \psi_j$ for $1 \leq j \leq J$ and multi-indices γ such that $|\gamma| \leq k$. Thus the $A_{\alpha, \beta}$ are continuous functions on $\bar{\Omega}$. Hence, for all multi-indices α and β , the maximum of $|A_{\alpha, \beta}|$ on $\bar{\Omega}$ exists. We have shown that

$$\operatorname{Re} Q_1(u) \geq C_1 C_2 \|u\|_{k,2,\Omega}^2 - C_1 C_3 \|u\|_{k,2,\Omega} \|u\|_{k-1,2,\Omega}$$

for $u \in H_0^{k,2}(\Omega)$.

A reasoning similar to the one used to study the last two series in (12.4.8) shows that there exists a finite constant $C_4 \geq 0$ such that

$$\operatorname{Re} Q_2(u) \leq C_4 \|u\|_{k,2,\Omega} \|u\|_{k-1,2,\Omega}$$

for $u \in H_0^{k,2}(\Omega)$.

Hence

$$\operatorname{Re} B(u, u) \geq C_1 C_2 \|u\|_{k,2,\Omega}^2 - (C_1 C_3 + C_4) \|u\|_{k,2,\Omega} \|u\|_{k-1,2,\Omega} \quad (12.4.9)$$

for $u \in H_0^{k,2}(\Omega)$.

If we substitute $a = \sqrt{C_1 C_2} \|u\|_{k,2,\Omega}$ and $b = \frac{C_1 C_3 + C_4}{\sqrt{C_1 C_2}} \|u\|_{k-1,2,\Omega}$ in $ab \leq \frac{1}{2}(a^2 + b^2)$, we get

$$(C_1 C_3 + C_4) \|u\|_{k,2,\Omega} \|u\|_{k-1,2,\Omega} \leq \frac{C_1 C_2}{2} \|u\|_{k,2,\Omega}^2 + \frac{(C_1 C_3 + C_4)^2}{2C_1 C_2} \|u\|_{k-1,2,\Omega}^2$$

for $u \in H_0^{k,2}(\Omega)$.

The relation (12.4.9) becomes

$$\operatorname{Re} B(u, u) \geq \frac{C_1 C_2}{2} \|u\|_{k,2,\Omega}^2 - \frac{(C_1 C_3 + C_4)^2}{2C_1 C_2} \|u\|_{k-1,2,\Omega}^2 \quad (12.4.10)$$

for $u \in H_0^{k,2}(\Omega)$.

If $\underline{u} : \mathbb{R}^n \rightarrow \mathbb{R}$ is defined by

$$\underline{u}(\mathbf{x}) = \begin{cases} u(\mathbf{x}) & \text{if } \mathbf{x} \in \Omega \\ 0 & \text{if } \mathbf{x} \in \mathbb{R}^n \setminus \Omega \end{cases}$$

as in Proposition 11.2.2, then $\|u\|_{s,2,\Omega} = \|\underline{u}\|_{s,2,\mathbb{R}^n}$ for all $s \leq k$. Hence, it follows from Lemma 12.4.2 with $r = 0$, $s = k - 1$, $t = k$ and $\epsilon = \frac{(C_1 C_2)^2}{2(C_1 C_3 + C_4)^2}$ that there exists C_5 such that

$$\frac{(C_1 C_3 + C_4)^2}{2C_1 C_2} \|u\|_{k-1,2,\Omega}^2 \leq \frac{C_1 C_2}{4} \|u\|_{k,2,\Omega}^2 + \frac{(C_1 C_3 + C_4)^2 C_5}{2C_1 C_2} \|u\|_{2,\Omega}^2$$

for $u \in H_0^{k,2}(\Omega)$. Using this relation, we finally get from (12.4.10) that

$$\operatorname{Re} B(u, u) \geq \frac{C_1 C_2}{4} \|u\|_{k,2,\Omega}^2 - \frac{(C_1 C_3 + C_4)^2 C_5}{2 C_1 C_2} \|u\|_{2,\Omega}^2$$

for $u \in H_0^{k,2}(\Omega)$. This is Definition 12.4.1 with $C = C_1 C_2 / 4$ and $\lambda = (C_1 C_3 + C_4)^2 C_5 / (2 C_1 C_2)$. ■

Theorem 12.4.5 (Gårding's inequality)

Let Ω be a bounded open subset of $\mathbb{R}^n$ with a boundary of class C^2 . Suppose that

$$L(\mathbf{x}, D)u = \sum_{|\alpha|, |\beta| \leq k} (-1)^\alpha D^\alpha (b_{\alpha,\beta}(\mathbf{x}) D^\beta u) ,$$

is strongly elliptic on $\overline{\Omega}$. Moreover, suppose that $b_{\alpha,\beta} \in C(\overline{\Omega})$ for $|\alpha| = |\beta| = k$ and $b_{\alpha,\beta} \in L^\infty(\Omega)$ for $|\alpha| + |\beta| < 2k$ with $|\alpha|, |\beta| \leq k$. Then $B : H_0^{k,2}(\Omega) \times H_0^{k,2}(\Omega) \rightarrow \mathbb{C}$ defined in (12.2.5) is weakly coercive.

Proof.

Let

$$B(v, u) = \sum_{|\alpha|, |\beta| \leq k} \langle D^\alpha v, b_{\alpha,\beta}(\mathbf{x}) D^\beta u \rangle_2 = B_1(v, u) + B_2(v, u)$$

for $u, v \in H_0^{k,2}(\Omega)$, where

$$B_1(v, u) = \sum_{|\alpha|, |\beta| = k} \langle D^\alpha v, b_{\alpha,\beta}(\mathbf{x}) D^\beta u \rangle_2 \quad \text{and} \quad B_2(v, u) = \sum_{\substack{|\alpha|, |\beta| \leq k \\ |\alpha| + |\beta| < 2k}} \langle D^\alpha v, b_{\alpha,\beta}(\mathbf{x}) D^\beta u \rangle_2 .$$

From Lemma 12.4.4, there exist $C_1 > 0$ and $\lambda_1 > 0$ such that

$$\operatorname{Re} B_1(u, u) \geq C_1 \|u\|_{k,2,\Omega}^2 - \lambda_1 \|u\|_{2,\Omega}^2 \tag{12.4.11}$$

for $u \in H_0^{k,2}(\Omega)$.

Moreover, using Schwarz inequality, we get

$$\begin{aligned} |B_2(u, u)| &= \left| \sum_{\substack{|\alpha|, |\beta| \leq k \\ |\alpha| + |\beta| < 2k}} \langle D^\alpha u, b_{\alpha,\beta}(\mathbf{x}) D^\beta u \rangle_2 \right| \leq \sum_{\substack{|\alpha|, |\beta| \leq k \\ |\alpha| + |\beta| < 2k}} \int_{\Omega} |b_{\alpha,\beta}(\mathbf{x})| |D^\beta u(\mathbf{x})| |D^\alpha u(\mathbf{x})| \, d\mathbf{x} \\ &\leq \sum_{\substack{|\alpha|, |\beta| \leq k \\ |\alpha| + |\beta| < 2k}} \|b_{\alpha,\beta}\|_\infty \int_{\Omega} |D^\beta u(\mathbf{x})| |D^\alpha u(\mathbf{x})| \, d\mathbf{x} \\ &\leq \sum_{\substack{|\alpha|, |\beta| \leq k \\ |\alpha| + |\beta| < 2k}} \|b_{\alpha,\beta}\|_\infty \left(\int_{\Omega} |D^\alpha u(\mathbf{x})|^2 \, d\mathbf{x} \right)^{1/2} \left(\int_{\Omega} |D^\beta u(\mathbf{x})|^2 \, d\mathbf{x} \right)^{1/2} \end{aligned}$$

$$\leq \sum_{\substack{|\alpha|, |\beta| \leq k \\ |\alpha| + |\beta| < 2k}} \|b_{\alpha, \beta}\|_{\infty} \|u\|_{k, 2, \Omega} \|u\|_{k-1, 2, \Omega} \leq C_2 \|u\|_{k, 2, \Omega} \|u\|_{k-1, 2, \Omega} \quad (12.4.12)$$

for $u \in H_0^{k,2}(\Omega)$, where

$$C_2 = \sum_{\substack{|\alpha|, |\beta| \leq k \\ |\alpha| + |\beta| < 2k}} \|b_{\alpha, \beta}\|_{\infty} \geq 0.$$

Note that the essential supremum $\|b_{\alpha, \beta}\|_{\infty}$ is finite for all multi-indices α and β because $b_{\alpha, \beta} \in L^{\infty}(\Omega)$ for $|\alpha|, |\beta| \leq k$ and $|\alpha| + |\beta| < 2k$.

If we substitute $a = \sqrt{C_1} \|u\|_{k, 2, \Omega}$ and $b = \frac{C_2}{\sqrt{C_1}} \|u\|_{k-1, 2, \Omega}$ in $ab \leq \frac{1}{2}(a^2 + b^2)$, we get

$$C_2 \|u\|_{k, 2, \Omega} \|u\|_{k-1, 2, \Omega} \leq \frac{C_1}{2} \|u\|_{k, 2, \Omega}^2 + \frac{C_2^2}{2C_1} \|u\|_{k-1, 2, \Omega}^2$$

for $u \in H_0^{k,2}(\Omega)$. It follows from (12.4.12) that

$$|B_2(u, u)| \leq \frac{C_1}{2} \|u\|_{k, 2, \Omega}^2 + \frac{C_2^2}{2C_1} \|u\|_{k-1, 2, \Omega}^2 \quad (12.4.13)$$

for $u \in H_0^{k,2}(\Omega)$.

If $\underline{u} : \mathbb{R}^n \rightarrow \mathbb{R}$ is defined by

$$\underline{u}(\mathbf{x}) = \begin{cases} u(\mathbf{x}) & \text{if } \mathbf{x} \in \Omega \\ 0 & \text{if } \mathbf{x} \in \mathbb{R}^n \setminus \Omega \end{cases}$$

as in Proposition 11.2.2, then $\|u\|_{s, 2, \Omega} = \|\underline{u}\|_{s, 2, \mathbb{R}^n}$ for $s \leq k$. Hence, it follows from Lemma 12.4.2 with $r = 0$, $s = k - 1$, $t = k$ and $\epsilon = \frac{C_1}{2C_2^2}$ that there exists C_3 such that

$$\frac{C_2^2}{2C_1} \|u\|_{k-1, 2, \Omega}^2 \leq \frac{C_1}{4} \|u\|_{k, 2, \Omega}^2 + \frac{C_2^2 C_3}{2C_1} \|u\|_{2, \Omega}^2$$

for $u \in H_0^{k,2}(\Omega)$. We then get from (12.4.13) that

$$|B_2(u, u)| \leq \frac{3C_1}{4} \|u\|_{k, 2, \Omega}^2 + \frac{C_2^2 C_3}{2C_1} \|u\|_{2, \Omega}^2 \quad (12.4.14)$$

for $u \in H_0^{k,2}(\Omega)$.

We conclude from (12.4.11) and (12.4.14) that

$$\begin{aligned} \operatorname{Re} B(u, u) &= \operatorname{Re} B_1(u, u) + \operatorname{Re} B_2(u, u) \geq C_1 \|u\|_{k, 2, \Omega}^2 - \lambda_1 \|u\|_{2, \Omega}^2 - \frac{3C_1}{4} \|u\|_{k, 2, \Omega}^2 - \frac{C_2^2 C_3}{2C_1} \|u\|_{2, \Omega}^2 \\ &\geq \frac{C_1}{4} \|u\|_{k, 2, \Omega}^2 - \left(\lambda_1 + \frac{C_2^2 C_3}{2C_1} \right) \|u\|_{2, \Omega}^2 \end{aligned}$$

for $u \in H_0^{k,2}(\Omega)$. This is Definition 12.4.1 with $C = C_1/4$ and $\lambda = \lambda_1 + (C_2^2 C_3)/(2C_1 C_2)$. ■

It is interesting to note that there is a converse to Gårding's inequality.

Theorem 12.4.6

Let Ω be a bounded open subset of $\mathbb{R}^n$. consider

$$L(\mathbf{x}, D)u = \sum_{|\alpha|, |\beta| \leq k} (-1)^\alpha D^\alpha (b_{\alpha, \beta}(\mathbf{x}) D^\beta u) ,$$

where $b_{\alpha, \beta} \in C(\overline{\Omega})$ for $|\alpha| = |\beta| = k$ and $b_{\alpha, \beta} \in L^\infty(\Omega)$ for $|\alpha| + |\beta| < 2k$ with $|\alpha|, |\beta| \leq k$. If $B : H_0^{k,2}(\Omega) \times H_0^{k,2}(\Omega) \rightarrow \mathbb{C}$ defined in (12.2.5) is weakly coercive, then $L(\mathbf{x}, D)$ is strongly elliptic on Ω .

A proof of this result can be found in [10] for instance.

12.5 Existence and Uniqueness of Solutions

We begin this section by presenting some results concerning the existence of classical solutions.

Consider the following partial differential equation on a bounded open set Ω .

$$\begin{aligned} L(\mathbf{x}, D)u(\mathbf{x}) &= \sum_{i,j=1}^n a_{i,j}(\mathbf{x}) D_{x_i, x_j} u(\mathbf{x}) + \sum_{i=1}^n a_i(\mathbf{x}) D_{x_i} u(\mathbf{x}) \\ &\quad + a_0(\mathbf{x})u(\mathbf{x}) = f(\mathbf{x}) \quad , \quad \mathbf{x} \in \Omega \\ u(\mathbf{x}) &= g(\mathbf{x}) \quad , \quad \mathbf{x} \in \partial\Omega \end{aligned} \tag{12.5.1}$$

where $b_{i,j} \in C^{0,\alpha}(\overline{\Omega})$ for all $1 \leq i, j \leq n$ and $b_i \in C^{0,\alpha}(\overline{\Omega})$ for all $0 \leq i \leq n$ and some $\alpha \geq 0$. If $L(\mathbf{x}, D)$ is strongly elliptic in Ω , $a_0(\mathbf{x}) \leq 0$ for all $\mathbf{x} \in \Omega$, $g \in C(\partial\Omega)$ and $f \in C^{0,\alpha}(\overline{\Omega})$, then there exists a unique solution $u \in C^{2,\alpha}(\overline{\Omega})$ to (12.5.1). The uniqueness of the solution comes from the relation

$$\|u\|_{\infty, \overline{\Omega}} \leq \|g\|_{\infty, \partial\Omega} + Q\|f\|_{\infty, \overline{\Omega}} ,$$

where the constant Q depends on Ω , the constant C in the definition of strongly elliptic in Definition 2.4.12, and the maximum of $a_{i,j}$, a_i on $\overline{\Omega}$. An accessible proof of these results can be found in [35].

The next results in this section are more general and are not limited to classical solutions.

Theorem 12.5.1

Let Ω be an open subset of $\mathbb{R}^n$ and let X be a closed subspace of $H^{k,2}(\Omega)$ such that $H_0^{k,2}(\Omega) \subset X$. Given $f \in L^2(\Omega)$, let

$$B(v, u) = \sum_{|\alpha|, |\beta| \leq k} \langle D^\alpha v, b_{\alpha, \beta}(\mathbf{x}) D^\beta u \rangle_2 = \langle v, f \rangle_2 \quad , \quad u, v \in X , \tag{12.5.2}$$

be a variational formulation of the problem

$$L(\mathbf{x}, D)u = \sum_{|\alpha|, |\beta| \leq k} (-1)^\alpha D^\alpha (b_{\alpha, \beta}(\mathbf{x}) D^\beta u) = f , \tag{12.5.3}$$

where $b_{\alpha,\beta} \in L^\infty(\Omega)$ for all α and β . If B is coercive on X , then there exists a one-to-one, bounded linear mapping $F : L^2(\Omega) \rightarrow X$ such that $F(f)$ is a weak solution of (12.5.2); namely, $B(v, F(f)) = \langle v, f \rangle_2$ for all $v \in X$.

Proof.

Given $g \in L^2(\Omega)$, we have that $\tilde{g} : X \rightarrow \mathbb{R}$ defined by $\tilde{g}(v) = \langle v, g \rangle_2$ for all $v \in X$ is a bounded linear functional. To prove this statement, we use Schwarz inequality to get $|\tilde{g}(v)| = |\langle v, g \rangle_2| \leq \|g\|_2 \|v\|_2 \leq \|g\|_2 \|v\|_{k,2}$ for all $v \in H^{k,2}(\Omega)$. Thus $\|\tilde{g}\| \leq \|g\|_2$. In particular, this shows that $T_1 : L^2(\Omega) \rightarrow X^*$ defined by $T_1(g) = \tilde{g}$ is a continuous linear mapping of norm at most 1.

The application T_1 is one-to-one onto its image. Suppose that $f_1, f_2 \in L^2(\Omega)$ satisfy $\langle v, f_1 \rangle_2 = \langle v, f_2 \rangle_2$ for all $v \in X$, then $\langle v, f_1 \rangle_2 = \langle v, f_2 \rangle_2$ for all $v \in \mathcal{D}(\Omega)$ because $X \supset H_0^{k,2}(\Omega)$. Thus $f_1 = f_2$ almost everywhere on Ω because $\mathcal{D}(\Omega)$ is dense in $L^2(\Omega)$.²

Since X is an Hilbert space with respect to the norm $\|\cdot\|_{k,2}$ and $B : X \times X \rightarrow \mathbb{C}$ is a bounded coercive sesquilinear form according to Proposition 12.2.3, then we may use Lax-Milgram Theorem, Corollary 4.5.5 and Remark 4.5.7, to conclude that there exists a one-to-one, bounded linear mapping $T_2 : X \rightarrow X^*$ such that $B(v, T_2(\tilde{f})) = \tilde{f}(v)$ for all $v \in X$.

We have that $F = T_2 \circ T_1$. ■

Let X be a closed subspace of $H^{k,2}(\Omega)$ containing $H_0^{k,2}(\Omega)$ and let f be an element of $L^2(\Omega)$ as in the previous theorem. The formal adjoint operator associated to

$$L(\mathbf{x}, D)u = \sum_{|\alpha|, |\beta| \leq k} (-1)^\alpha D^\alpha (b_{\alpha,\beta}(\mathbf{x}) D^\beta u)$$

is

$$L^*(\mathbf{x}, D)v = \sum_{|\alpha|, |\beta| \leq k} (-1)^\beta D^\beta (\overline{b_{\alpha,\beta}(\mathbf{x})} D^\alpha v) .$$

A variational formulation of $L(\mathbf{x}, D)u = f$ is to find $u \in X$ such that

$$B(v, u) = \langle v, f \rangle_2 \quad (12.5.4)$$

for all $v \in X$, where B is defined in the previous theorem. Similarly, A variational formulation of $L^*(\mathbf{x}, D)v = f$ is to find $v \in X$ such that

$$B^*(u, v) = \sum_{|\alpha|, |\beta| \leq k} \left\langle D^\beta u, \overline{b_{\alpha,\beta}(\mathbf{x})} D^\alpha v \right\rangle_2 = \langle u, f \rangle_2 \quad (12.5.5)$$

for all $u \in X$. Since

$$B^*(u, v) = \overline{B(v, u)} \quad (12.5.6)$$

for all $u, v \in X$, B^* is a bounded sesquilinear form. The variational problem (12.5.5) is the **adjoint (variational) problem** to the variational problem (12.5.4). The variational

²Note that we also have that T_1 is an isometry into X . Take $v \in X$ such that $\langle v, f \rangle_2 > 0$. then $T_1((\|f\|_2 / \langle v, f \rangle_2) v) = \|f\|_2$.

formulation (12.5.4) is **self adjoint** if $B^* = B$. This implies that $L(\mathbf{x}, D) = L^*(\mathbf{x}, D)$. To prove this, consider $B^*(v, u) = B(v, u)$ for $u, v \in \mathcal{D}(\Omega) \subset X$ and use integration by parts.

Moreover, from (12.5.6), we get that B is coercive if and only if B^* is coercive. We can even take the same constant C in the definition of a coercive sesquilinear form applied to B and B^* . Similarly, B is weakly coercive if and only if B^* is weakly coercive. Again, we can take the same constants C and λ in the definition of a weakly coercive sesquilinear form applied to B and B^* .

If B is coercive, it follows from Theorem 12.5.1 that there exists a bounded linear mapping $F_B : L^2(\Omega) \rightarrow X$ such that $B(v, F_B(f)) = \langle v, f \rangle_2$ for all $v \in X$. Similarly, since B^* is also coercive, there exists a bounded linear mapping $F_{B^*} : L^2(\Omega) \rightarrow X$ such that $B^*(u, F_{B^*}(f)) = \langle u, f \rangle_2$ for all $u \in X$. Let $i : X \rightarrow L^2(\Omega) = H^{0,2}(\Omega)$ be the inclusion mapping. We have from Rellich's theorem, Theorem 11.7.10, that i is a compact operator. It follows from Proposition 4.4.22 that $K_1 = i \circ F_B : L^2(\Omega) \rightarrow L^2(\Omega)$ and $K_2 = i \circ F_{B^*} : L^2(\Omega) \rightarrow L^2(\Omega)$ are compact operators. The following sequence of equalities proves that $K_1^* = (i \circ F_B)^* = i \circ F_{B^*} = K_2$.

$$\begin{aligned} \langle (i \circ F_B)(g), f \rangle_2 &= \langle F_B(g), f \rangle_2 = B^*(F_B(g), F_{B^*}(f)) = \overline{B(F_{B^*}(f), F_B(g))} = \overline{\langle F_{B^*}(f), g \rangle_2} \\ &= \langle g, F_{B^*}(f) \rangle_2 = \langle g, (i \circ F_{B^*})(f) \rangle_2 \end{aligned}$$

for all $f, g \in L^2(\Omega)$.

Theorem 12.5.2

Let Ω be an open subset of $\mathbb{R}^n$ and let X be a closed subspace of $H^{k,2}(\Omega)$ containing $H_0^{k,2}(\Omega)$. Consider the variational formulation (12.5.2) of (12.5.3), where B is weakly coercive. Let

$$V = \{u \in X : B(v, u) = 0 \text{ for all } v \in X\}$$

and

$$W = \{u \in X : B(u, v) = 0 \text{ for all } v \in X\} .$$

Then,

1. $\dim V = \dim W < \infty$.
2. Given $f \in L^2(\Omega)$, there exists a solution of (12.5.2) if and only if $\langle v, f \rangle_2 = 0$ for all $v \in W$. The solution is unique up to the addition of an element of V .
3. Given $f \in L^2(\Omega)$, if $V = W = \{0\}$, then there exists a unique solution of (12.5.2).

Proof.

Since B is weakly coercive, there exist $C > 0$ and $\lambda > 0$ such that $\operatorname{Re} B(u, u) \geq C\|u\|_{k,2,\Omega}^2 - \lambda\|u\|_{2,\Omega}^2$ for all $u \in X$. The sesquilinear form

$$\tilde{B}(v, u) = B(v, u) + \lambda \langle v, u \rangle_2 \quad , \quad u, v \in X \quad , \quad (12.5.7)$$

is therefore coercive on X because $\operatorname{Re} \tilde{B}(u, u) \geq C\|u\|_{k,2,\Omega}^2$ for all $u \in X$. According to the discussion before the statement of the theorem with B replaced by $\tilde{B}$, there exists a compact

operator $K_1 : L^2(\Omega) \rightarrow L^2(\Omega)$ such that $\mathcal{R}(K_1) \subset X$ and

$$\tilde{B}(v, K_1(f)) = \langle v, f \rangle_2$$

for all $v \in X$ and $f \in L^2(\Omega)$.

Given $f \in L^2(\Omega)$, we have that $u \in X$ satisfies

$$\begin{aligned} B(v, u) = \langle v, f \rangle_2 \quad , \quad v \in X &\Leftrightarrow \tilde{B}(v, u) = \langle v, f \rangle_2 + \lambda \langle v, u \rangle_2 \quad , \quad v \in X \\ &\Leftrightarrow \tilde{B}(v, u) = \langle v, f + \lambda u \rangle_2 \quad , \quad v \in X \\ &\Leftrightarrow u = K_1(f + \lambda u) \\ &\Leftrightarrow \lambda^{-1}u - K_1(u) = \lambda^{-1}K_1(f) . \end{aligned} \quad (12.5.8)$$

With $f = 0$, we get $u \in V$ if and only if $u \in L^2(\Omega)$ satisfies $\lambda^{-1}u = K_1(u)$ because $u = \lambda K_1(u) \in \mathcal{R}(K_1) \subset X$. Hence

$$V = \{u \in L^2(\Omega) : K_1(u) = \lambda^{-1}u\} = \mathcal{N}(K_1 - \lambda^{-1}\text{Id}) .$$

A similar reasoning with B^* instead of B shows that there exists a compact operator $K_2 : L^2(\Omega) \rightarrow L^2(\Omega)$ with $\mathcal{R}(K_2) \subset X$ such that

$$W = \{u \in L^2(\Omega) : K_2(u) = \lambda^{-1}u\} = \mathcal{N}(K_2 - \lambda^{-1}\text{Id}) .$$

From the discussion before the statement of this theorem, we have that $K_2 = K_1^*$. Item (1) of the theorem is a consequence of item (3) of Theorem 4.4.24.

We have shown in (12.5.8) that $u \in X$ satisfies (12.5.2) if and only if $K_1(f) = u - \lambda K_1(u) = -\lambda(K_1(u) - \lambda^{-1}u)$. Therefore, (12.5.2) has a solution if and only if $K_1(f) \in \mathcal{R}(K_1 - \lambda^{-1}\text{Id})$. Since $\mathcal{R}(K_1 - \lambda^{-1}\text{Id})$ is closed in $L^2(\Omega)$ according to item (2) of Theorem 4.4.24, we get from Theorems 4.4.41 that $\mathcal{R}(K_1 - \lambda^{-1}\text{Id}) = \mathcal{N}(K_1^* - \lambda^{-1}\text{Id})^\perp = \mathcal{N}(K_2 - \lambda^{-1}\text{Id})^\perp = W^\perp$ in $L^2(\Omega)$. Thus, $u \in X$ satisfies (12.5.2) if and only if $K_1(f) \in W^\perp$. Finally, since

$$\langle K_1(f), w \rangle_2 = \langle f, K_1^*(w) \rangle_2 = \lambda^{-1} \langle f, w \rangle_2$$

for all $w \in W$ and $f \in L^2(\Omega)$, we have $K_1(f) \in W^\perp$ if and only if $f \in W^\perp$. This is item (2) of the theorem. Item (3) is a consequence of (2). ■

Remark 12.5.3

In our previous discussion, we have assumed that the boundary conditions of the Dirichlet problem (12.2.3) were homogeneous. More precisely, we have assumed that $\frac{\partial^j u}{\partial \nu^j} = 0$ on $\partial\Omega$ for $0 \leq j < k$. However, more general boundary conditions can be considered.

The variational approach that we have developed before can also be used to solve Dirichlet problems with non-homogeneous boundary conditions. Consider

$$\begin{aligned} L(\mathbf{x}, D)u(\mathbf{x}) &= \sum_{|\alpha|, |\beta| \leq k} (-1)^{|\alpha|} D^\alpha (b_{\alpha, \beta}(\mathbf{x}) D^\beta u(\mathbf{x})) = f(\mathbf{x}) \quad , \quad \mathbf{x} \in \Omega , \\ \frac{\partial^j u}{\partial \nu^j}(\mathbf{x}) &= g_j(\mathbf{x}) \quad , \quad \mathbf{x} \in \partial\Omega \text{ and } 0 \leq j < k , \end{aligned} \quad (12.5.9)$$

where the functions $g_j : \partial\Omega \rightarrow \mathbb{C}$ are nice functions. For instance, $g_j \in C^\infty(\partial\Omega)$ for all j .

First, find $g \in H^{2k,2}(\mathbb{R}^n)$ such that $\frac{\partial^j g}{\partial \nu^j} = g_j$ on $\partial\Omega$ for $0 \leq j < k$. According to Theorem 12.3.3, such a function g can be found if $g_j \in C^\infty(\partial\Omega)$ for all j . Then, u is a solution of (12.5.9) if and only if $v = u - g$ is a solution

$$\begin{aligned} L(\mathbf{x}, D)v(\mathbf{x}) &= f(\mathbf{x}) - L(\mathbf{x}, D)g(\mathbf{x}) \quad , \quad \mathbf{x} \in \Omega \quad , \\ \frac{\partial^j v}{\partial \nu^j}(\mathbf{x}) &= 0 \quad , \quad \mathbf{x} \in \partial\Omega \text{ and } 0 \leq j < k \quad . \end{aligned}$$

♠

The existence and uniqueness of the solution for the following examples is a direct consequence of Theorem 12.5.1. However, for the sake of illustrating with concrete examples the proof of Theorem 12.5.1 and to present the variational approach to study elliptic partial differential equations, we deduce the variational problem associated to the elliptic partial differential equation given in each example and use the Lax-Milgram Theorem to prove the existence and uniqueness of solutions.

Example 12.5.4

Let Ω be a bounded subset of $\mathbb{R}^n$. Find the weak solution of the Dirichlet Problem

$$\begin{aligned} -\Delta u + u &= f \quad \text{in } \Omega \\ u &= 0 \quad \text{on } \partial\Omega \end{aligned} \tag{12.5.10}$$

where $f \in L^2(\Omega)$.

Variational Problem: The weak solution is the solution $u \in H_0^{1,2}(\Omega)$ of the equation

$$\int_{\Omega} (\nabla u \cdot \nabla v + uv - fv) \, d\mathbf{x} = 0 \tag{12.5.11}$$

for all $v \in H_0^{1,2}(\Omega)$.

We prove that a classical solution is a weak solution. If $u \in C^2(\overline{\Omega})$ is a classical solution of (12.5.10), integrating (12.5.10) multiplied by $v \in C_c^1(\Omega)$ over Ω gives

$$\int_{\Omega} (\nabla u \cdot \nabla v + uv - fv) \, d\mathbf{x} = 0$$

for all $v \in C_c^1(\Omega)$. The integral was computed using Green's identity (9.1.1). Since $C_c^1(\Omega)$ is dense in $H_0^{1,2}(\Omega)$, we have that (12.5.11) is true³. Moreover, it follows from Proposition 11.8.14 that $u \in H_0^{1,2}(\Omega)$ because $u \in C^2(\overline{\Omega})$ and $u(\mathbf{x}) = 0$ for all $\mathbf{x} \in \partial\Omega$.

Existence and uniqueness of the weak solution:

$$\begin{aligned} B : H_0^{1,2}(\Omega) \times H_0^{1,2}(\Omega) &\rightarrow \mathbb{R} \\ (u, v) &\mapsto \int_{\Omega} (\nabla u \cdot \nabla v + uv) \, d\mathbf{x} = \int_{\Omega} (\nabla u, u) \cdot (\nabla v, v) \, d\mathbf{x} \end{aligned}$$

³See Example 12.1.1 for the details of a similar case.

is a coercive, symmetric bounded bilinear mapping. That B is bounded is proved as in Example 12.1.1 with two applications of Schwarz inequality, one in $\mathbb{R}^n$ and the other in $L^2(\Omega)$; namely,

$$\begin{aligned} |B(u, v)| &\leq \int_{\Omega} |(\nabla u, u) \cdot (\nabla v, v)| \, d\mathbf{x} \leq \int_{\Omega} \|(\nabla u, u)\|_2 \|(\nabla v, v)\|_2 \, d\mathbf{x} \\ &\leq \left(\int_{\Omega} \|(\nabla u, u)\|_2^2 \, d\mathbf{x} \right)^{1/2} \left(\int_{\Omega} \|(\nabla v, v)\|_2^2 \, d\mathbf{x} \right)^{1/2} \\ &= \left(\int_{\Omega} \left(\sum_{j=1}^n \left| \frac{\partial u}{\partial x_j} \right|^2 + |u|^2 \right) \, d\mathbf{x} \right)^{1/2} \left(\int_{\Omega} \left(\sum_{j=1}^n \left| \frac{\partial v}{\partial x_j} \right|^2 + |v|^2 \right) \, d\mathbf{x} \right)^{1/2} = \|u\|_{1,2} \|v\|_{1,2} \end{aligned}$$

for all $u, v \in H_0^{1,2}(\Omega)$. Since $B(u, u) = \|u\|_{1,2}^2$ for $u \in H_0^{1,2}(\Omega)$, it is obvious that B is coercive. It is also obvious that B is symmetric. The function $f \in L^2(\Omega)$ yields a bounded functional on $H_0^{1,2}(\Omega)$ defined by

$$\tilde{f}(u) = \langle u, f \rangle_2 = \int_{\Omega} f u \, d\mathbf{x}$$

for all $u \in H_0^{1,2}(\Omega)$. We may thus apply Lax-Milgram Theorem to (12.5.11) to conclude that there exists a unique weak solution $u \in H_0^{1,2}(\Omega)$ given by

$$\frac{1}{2} \int_{\Omega} (\nabla u \cdot \nabla u + u u) \, d\mathbf{x} - \int_{\Omega} f u \, d\mathbf{x} = \min_{v \in H_0^{1,2}(\Omega)} \left\{ \frac{1}{2} \int_{\Omega} (\nabla v \cdot \nabla v + v v) \, d\mathbf{x} - \int_{\Omega} f v \, d\mathbf{x} \right\}.$$

Regularity of the solution: As a consequence of Theorem 12.6.18, $u \in H^2(\Omega)$.

Classical solution: If the weak solution u is in $C^2(\overline{\Omega})$ and assuming that Ω has a boundary of class C^1 , it follows from Proposition 11.8.14 that $u(\mathbf{x}) = 0$ for all $\mathbf{x} \in \partial\Omega$ because $u \in H_0^{1,2}(\Omega) \cap C^2(\overline{\Omega})$. Using Green's identity (9.1.1), we can show that (12.5.11) becomes

$$\int_{\Omega} (\Delta u + u - f) v \, d\mathbf{x} = 0$$

for all $v \in C_c^1(\Omega)$. Since $C_c^1(\Omega)$ is dense in $C_c(\Omega)$, we get that $\Delta u + u - f = 0$ almost everywhere. If $f \in C(\Omega)$, we have $-\Delta u + u - f = 0$ everywhere on Ω since $u \in C^2(\Omega)$. Thus u is the classical solution of (12.5.10). ♣

Example 12.5.5

Let Ω be a bounded subset of $\mathbb{R}^n$ with a boundary of class C^2 . Find the weak solution of the Neumann Problem

$$\begin{aligned} -\Delta u + u &= f \quad \text{in } \Omega \\ \frac{\partial u}{\partial \nu} &= 0 \quad \text{on } \partial\Omega \end{aligned} \tag{12.5.12}$$

where $f \in L^2(\Omega)$.

Variational Problem: The weak solution is the solution $u \in H^{1,2}(\Omega)$ of the equation

$$\int_{\Omega} (\nabla u \cdot \nabla v + u v - f v) \, d\mathbf{x} = 0 \tag{12.5.13}$$

for all $v \in H^{1,2}(\Omega)$.

We prove that a classical solution is a weak solution. Suppose that $u \in C^2(\overline{\Omega})$ is a classical solution of (12.5.12). We then have that

$$\int_{\Omega} (-\Delta u + u - f) v \, d\mathbf{x} = 0 \quad (12.5.14)$$

for all $v \in C^1(\overline{\Omega})$. We get from Green's identity (9.1.1) that

$$\int_{\partial\Omega} v \frac{\partial u}{\partial \nu} \, dS_{\mathbf{x}} = \int_{\Omega} (v \Delta u + \nabla u \cdot \nabla v) \, d\mathbf{x}$$

for all $v \in C^1(\overline{\Omega})$. Since $\frac{\partial u}{\partial \nu} = 0$ on $\partial\Omega$, we get

$$\int_{\Omega} v \Delta u \, d\mathbf{x} = - \int_{\Omega} \nabla u \cdot \nabla v \, d\mathbf{x}$$

for all $v \in C^1(\overline{\Omega})$. Hence, (12.5.14) becomes

$$\int_{\Omega} (\nabla u \cdot \nabla v) + u v - f v \, d\mathbf{x} = 0$$

for all $v \in C^1(\overline{\Omega})$. Since $C^1(\overline{\Omega}) \supset \mathcal{D}(\Omega)$ is dense in $H^{1,2}(\Omega)$, we have that (12.5.13) is true.

Existence and uniqueness of the weak solution: The bilinear form

$$\begin{aligned} B : H^{1,2}(\Omega) \times H^{1,2}(\Omega) &\rightarrow \mathbb{R} \\ (u, v) &\mapsto \int_{\Omega} (\nabla u \cdot \nabla v + u v) \, d\mathbf{x} \end{aligned}$$

is coercive, symmetric and bounded. The proof is identical to the proof in Example 12.5.4 that the bilinear form $B : H_0^{1,2}(\Omega) \times H_0^{1,2}(\Omega) \rightarrow \mathbb{R}$ is coercive, symmetric and bounded. Moreover, we also have that the functional on $H^{1,2}(\Omega)$ defined by

$$\tilde{f}(u) = \langle u, f \rangle_2 = \int_{\Omega} f u \, d\mathbf{x}$$

for all $u \in H^{1,2}(\Omega)$ is a bounded mapping. We may use Lax-Milgram Theorem to (12.5.13) to conclude that there exists a unique weak solution $u \in H^{1,2}(\Omega)$ given by

$$\frac{1}{2} \int_{\Omega} (\nabla u \cdot \nabla u + u u) \, d\mathbf{x} - \int_{\Omega} f u \, d\mathbf{x} = \min_{v \in H^{1,2}(\Omega)} \left\{ \frac{1}{2} \int_{\Omega} (\nabla v \cdot \nabla v + v v) \, d\mathbf{x} - \int_{\Omega} f v \, d\mathbf{x} \right\}.$$

Regularity of the solution: As a consequence of Theorem 12.6.18, $u \in H^2(\Omega)$.

Classical solution: If the weak solution u is in $C^2(\overline{\Omega})$, we get from (12.5.13) and Green's identity (9.1.1) that

$$\int_{\partial\Omega} v \frac{\partial u}{\partial \nu} \, dS_{\mathbf{x}} + \int_{\Omega} (-v \Delta u + u v - f v) \, d\mathbf{x} = 0 \quad (12.5.15)$$

for all $v \in C^1(\overline{\Omega})$.

Since $C_c^1(\Omega) \subset C^1(\overline{\Omega})$, we get from (12.5.15) that

$$\int_{\Omega} (-v \Delta u + u v - f v) \, d\mathbf{x} = 0$$

for all $v \in C_c^1(\Omega)$ because $v = 0$ on $\partial\Omega$ (we assume that v is extended to $\overline{\Omega}$). Since $C_c^1(\Omega) \supset \mathcal{D}(\Omega)$ is dense in $L^2(\Omega)$, we get that $-\Delta u + u - f = 0$ almost everywhere. If $f \in C(\Omega)$, we have $-\Delta u + u - f = 0$ everywhere on Ω since $u \in C^2(\Omega)$. Back to (12.5.15), this implies that

$$\int_{\partial\Omega} v \frac{\partial u}{\partial \nu} \, dS_{\mathbf{x}} = 0$$

for all $v \in C^1(\overline{\Omega})$. Since $C^1(\overline{\Omega})|_{\partial\Omega} \supset \mathcal{D}(\partial\Omega)$ is dense in $L^2(\partial\Omega)$, we get $\frac{\partial u}{\partial \nu} = 0$ almost everywhere on $\partial\Omega$ (using the measure along $\partial\Omega$). Since $u \in C^2(\overline{\Omega})$, this implies that $\frac{\partial u}{\partial \nu}(\mathbf{x}) = 0$ for all $\mathbf{x} \in \partial\Omega$. ♣

For the next example, we will use a slightly different procedure than the procedure introduced in Remark 12.5.3. Both procedures are equivalent.

Example 12.5.6

Let Ω be a bounded subset of $\mathbb{R}^n$. Find the weak solution of the Dirichlet Problem

$$\begin{aligned} -\Delta u + u &= f & \text{in } \Omega \\ u &= g & \text{on } \partial\Omega \end{aligned} \tag{12.5.16}$$

where $f \in L^2(\Omega)$ and $g \in C^\infty(\partial\Omega)$.

Variational Problem: Choose $\tilde{g} \in H^{1,2}(\Omega) \cap C(\overline{\Omega})$ such that $\tilde{g} = g$ on $\partial\Omega$. Let

$$K = \{v \in H^{1,2}(\Omega) : v - \tilde{g} \in H_0^{1,2}(\Omega)\}.$$

K is a closed and convex subset of $H^{1,2}(\Omega)$. That K be independent of the choice of $\tilde{g}$ is a consequence of Proposition 11.8.14. K depends only of g .

The weak solution is the solution $u \in K \subset H^{1,2}(\Omega)$ of the equation

$$B(u, v) \equiv \int_{\Omega} (\nabla u \cdot \nabla v + u v - f v) \, d\mathbf{x} = 0 \tag{12.5.17}$$

for all $v \in H^{1,2}(\Omega)$.

We prove that a classical solution is a weak solution. As usual, if $u \in C^2(\overline{\Omega})$ is a classical solution of (12.5.16), then integrating (12.5.16) multiplied by $v \in C_c^1(\Omega)$ over Ω gives

$$\int_{\Omega} (\nabla u \cdot \nabla v + u v - f v) \, d\mathbf{x} = 0$$

for all $v \in C_c^1(\Omega)$. As in the previous example, Green's identity (9.1.1) may be used to compute the integral. Since $C_c^1(\Omega) \supset \mathcal{D}(\Omega)$ is dense in $H_0^{1,2}(\Omega)$, we have that (12.5.17) is

true. Moreover, it follows from Proposition 11.8.14 that $u - \tilde{g} \in H_0^{1,2}(\Omega)$ because $u - \tilde{g} \in H^{1,2}(\Omega) \cap C(\overline{\Omega})$ and $u - \tilde{g} = 0$ on $\partial\Omega$. Thus $u \in K$.

Existence and uniqueness of the weak solution: As we have proved in previous examples, the bilinear form

$$B : H^{1,2}(\Omega) \times H^{1,2}(\Omega) \rightarrow \mathbb{R}$$

$$(u, v) \mapsto \int_{\Omega} (\nabla u \cdot \nabla v + uv) \, d\mathbf{x}$$

is coercive, symmetric and bounded. Moreover,

$$\tilde{f}(u) = \langle u, f \rangle_2 = \int_{\Omega} f(x)u(x) \, d\mathbf{x}$$

for all $u \in H^{1,2}(\Omega)$ is a bounded linear functional.

Before using Stampacchia Theorem, Theorem 4.5.4, we need to reformat the variational equation. We have that $u \in K$ is a weak solution if and only if

$$B(v - u, u) = \int_{\Omega} (\nabla(v - u) \cdot \nabla u + u(v - u)) \, d\mathbf{x} \geq \int_{\Omega} f(v - u) \, d\mathbf{x} = \langle v - u, f \rangle_2 \quad (12.5.18)$$

for all $v \in K$. If u is a weak solution, we get from (12.5.17) that (12.5.18) is true because $u - v \in H_0^1(\Omega)$ for all $v \in K$. In fact we have equality in (12.5.18). Conversely, if $u \in K$ satisfies (12.5.18), we substitute $v = u \pm w$ with $w \in H_0^1(\Omega)$ in (12.5.18) to get respectively

$$B(w, u) = \int_{\Omega} (\nabla w \cdot \nabla u + wu) \, d\mathbf{x} \geq \int_{\Omega} f w \, d\mathbf{x} = \langle w, f \rangle_2$$

and

$$-B(w, u) = - \int_{\Omega} (\nabla w \cdot \nabla u + wu) \, d\mathbf{x} \geq - \int_{\Omega} f w \, d\mathbf{x} = - \langle w, f \rangle_2$$

for all $w \in H_0^1(\Omega)$. Thus (12.5.17) is true.

We may now apply Stampacchia Theorem with $H = H^{1,2}(\Omega)$ to conclude that there exists a unique $u \in K$ satisfying (12.5.18). Therefore, there exists a unique weak solution. This solution u is given by

$$\frac{1}{2} \int_{\Omega} (\nabla u \cdot \nabla u + uu) \, d\mathbf{x} - \int_{\Omega} f u \, d\mathbf{x} = \min_{v \in K} \left\{ \frac{1}{2} \int_{\Omega} (\nabla v \cdot \nabla v + vv) \, d\mathbf{x} - \int_{\Omega} f v \, d\mathbf{x} \right\}.$$

Regularity of the solution: As a consequence of Theorem 12.6.18, $u \in H^2(\Omega)$.

Classical solution: If $f \in C(\Omega)$ and the weak solution u is in $C^2(\overline{\Omega})$, we can prove as we did many times before that $-\Delta u + u - f = 0$ everywhere on Ω . Moreover, it follows from Proposition 11.8.14 that $u - \tilde{g} = 0$ on $\partial\Omega$ because $u \in K \cap C^2(\overline{\Omega})$ implies that $u - \tilde{g} \in H_0^{1,2}(\Omega) \cap C(\overline{\Omega})$. $\clubsuit$

If the sesquilinear form in (12.5.2) is weakly coercive and symmetric, we get the following additional properties.

Theorem 12.5.7

Let Ω be an open subset of $\mathbb{R}^n$ and let X be a closed subspace of $H^{k,2}(\Omega)$ containing $H_0^{k,2}(\Omega)$. Consider the variational formulation (12.5.2) of (12.5.3), where B is coercive or weakly coercive, and symmetric.

1. There exists an orthonormal basis $\{u_j\}_{j=1}^\infty$ of $L^2(\Omega)$ consisting of eigenvectors of the sesquilinear form B defined in (12.5.2). We define $u_j \in L^2(\Omega)$ as an eigenvector associated to the eigenvalue λ_j if $u_j \neq 0$ and $B(v, u_j) = \langle v, \lambda_j u_j \rangle_2$ for all $v \in X$.
2. If λ_j is an eigenvalue of B , then λ_j is real and $\lambda_j > -\lambda$, where λ is the constant in the definition of a weakly coercive sesquilinear form if B is weakly coercive, or $\lambda = 0$ if B is coercive.
3. The eigenvalues $\{\lambda_j\}_{j=1}^\infty$ can be arranged in an increasing sequence converging to $+\infty$.
4. If the coefficients $b_{\alpha,\beta}$ of $L(\mathbf{x}, D)$ are all of class $C^\infty(\Omega)$, then the eigenvectors u_j are all in $C^\infty(\Omega)$.

Proof.

The following proof is valid for B weakly coercive (case $\lambda > 0$ below) or B coercive (case $\lambda = 0$). The restriction $\lambda > 0$ in the proof of Theorem 12.5.2 is not needed here because we do not have to play with λ^{-1} .

Let $\tilde{B}$ be the sesquilinear form defined in (12.5.7) of Theorem 12.5.2. Let K_1 and K_2 be the compact operators defined in the paragraph preceding Theorem 12.5.2) with B replaced by $\tilde{B}$ and B^* replaced by $\tilde{B}^*$ respectively.

Since B is symmetric, $\tilde{B}$ is symmetric. It follows from its definition that K_1 is a symmetric operator because $F_{\tilde{B}} = F_{\tilde{B}^*}$ in the definition of K_1 and K_2 given in the paragraph preceding Theorem 12.5.2.

K_1 is also one-to-one since it is the composition $i \circ F_{\tilde{B}}$, where $F_{\tilde{B}} : L^2(\Omega) \rightarrow X$ given by Theorem 12.5.1 is one-to-one and $i : X \rightarrow L^2(\Omega)$ is the inclusion.

It follows from Theorem 4.4.44 (and the two paragraphs preceding this theorem) that there exists an orthonormal basis of eigenvectors $\{u_j\}_{j=1}^\infty \subset L^2(\Omega)$ of K_1 such that the eigenvalue β_j associated to u_j is real and the sequence $\{\beta_j\}_{j=1}^\infty$ is a decreasing sequence converging to 0 if there are infinitely many distinct eigenvalues. The β_j may have the same value for some values of j if $\dim \mathcal{N}(K_1 - \beta_j \text{Id}) > 1$; namely, if the eigenspace associated to β_j has dimension larger than 1.

The eigenvalues are positive numbers because

$$\begin{aligned} \beta_j \|u_j\|_2^2 &= \langle \beta_j u_j, u_j \rangle_2 = \langle K_1(u_j), u_j \rangle_2 = \tilde{B}(F_{\tilde{B}}(u_j), F_{\tilde{B}}(u_j)) \\ &= \text{Re } \tilde{B}(F_{\tilde{B}}(u_j), F_{\tilde{B}}(u_j)) \geq C \|F_{\tilde{B}}(u_j)\|_{k,2}^2 > 0, \end{aligned}$$

where we have used $\tilde{B}(v, v) \in \mathbb{R}$ for all $v \in X$ because B is a symmetric operator, and where C is the constant from the definition of a coercive or weakly coercive operator satisfied by B .

From

$$\tilde{B}(v, \beta_j u_j) = \tilde{B}(v, (i \circ K_1)(u_j)) = \tilde{B}(v, F_{\tilde{B}}(u_j)) = \langle v, u_j \rangle_2$$

for all $v \in X$, we get $\tilde{B}(v, u_j) = \langle v, \beta_j^{-1} u_j \rangle_2$ for all $v \in X$.

Let $\lambda_j = \beta_j^{-1} - \lambda$ for $j > 0$. Then

$$\begin{aligned} B(v, u_j) &= \tilde{B}(v, u_j) - \lambda \langle v, u_j \rangle_2 = \langle v, \beta_j^{-1} u_j \rangle_2 - \lambda \langle v, u_j \rangle_2 \\ &= \langle v, (\beta_j^{-1} - \lambda) u_j \rangle_2 = \langle v, \lambda_j u_j \rangle_2 \end{aligned}$$

for all $v \in X$. Thus, $\lambda_j > -\lambda$ for all j and $\lim_{j \rightarrow \infty} \lambda_j = \infty$. Note that there are infinitely many distinct eigenvalues because the eigenspaces $\mathcal{N}(K_1 - \beta_j \text{Id})$ for $j > 0$ are finite dimensional and $L^2(\Omega)$ is infinite dimensional.

Finally, to prove item (4), we note that an eigenvector u_j is a solution in the sense of distribution of the partial differential equation $(L(\mathbf{x}, D) - \lambda_j \text{Id}) u_j = 0$ because

$$\langle v, L(\mathbf{x}, D) u_j - \lambda_j u_j \rangle_2 = B(v, u_j) - \langle v, \lambda_j u_j \rangle_2 = 0$$

for all $v \in \mathcal{D}(\Omega) \subset H_0^{k,2}(\Omega) \subset \cdot$. Since $L(\mathbf{x}, D) - \lambda_j \text{Id}$ is an elliptic operator with coefficients of class C^∞ on Ω , it follows from Corollary 12.6.17 that we will see in the next section that $u_j \in C^\infty(\Omega)$. ■

We now apply our results to the original Dirichlet and Neumann problems.

Theorem 12.5.8

Let Ω be a bounded open subset of $\mathbb{R}^n$ with a boundary of class C^2 . If the operator $L(\mathbf{x}, D)$ in the Dirichlet problem (12.2.3) is strongly elliptic on $\overline{\Omega}$, and $L(\mathbf{x}, D) = L^*(\mathbf{x}, D)$, we get the following results.

1. There exists an orthonormal basis $\{u_j\}_{j=1}^\infty$ of $L^2(\Omega)$ consisting of eigenvectors of $L(\mathbf{x}, D)$.
2. $u_j \in C^\infty(\overline{\Omega})$ for all j .
3. If $\nu(\mathbf{x})$ denotes the outward unit normal to $\partial\Omega$ at $\mathbf{x} \in \partial\Omega$, then $\frac{\partial^i u_j}{\partial \nu^i} = 0$ on $\partial\Omega$ for $0 \leq i < k$.
4. The eigenvalues are real, isolated and can be arranged in a increasing sequence in absolute value that converges to $+\infty$.

Moreover, if the sesquilinear form B of the variational formulation (12.2.4) of the Dirichlet problem (12.2.3) is also coercive, we get the following result.

5. The eigenvalues of $L(\mathbf{x}, D)$ are positive.

Proof.

Let $B(v, u) = \langle v, f \rangle_2$ for all $u, v \in H_0^{k,2}(\Omega)$ be a variational formulation of the Dirichlet problem (12.2.3) with $f \in L^2(\Omega)$, where B is defined in (12.5.2) with $X = H_0^{k,2}(\Omega)$. A variational formulation of the adjoint equation $L^*(\mathbf{x}, D)v = f$ is

$$B^*(u, v) = \langle u, f \rangle_2 \quad , \quad u, v \in H_0^{k,2}(\Omega) \quad , \quad (12.5.19)$$

where B^* is defined in (12.5.5) with $X = H_0^{k,2}(\Omega)$. Since $L^*(\mathbf{x}, D) = L(\mathbf{x}, D)$, (12.5.19) is another variational formulation of the Dirichlet problem (12.2.3). Let

$$D(u, v) = \frac{1}{2} (B(u, v) + B^*(u, v)) \quad , \quad u, v \in H_0^{k,2}(\Omega) \quad .$$

Then

$$D(v, u) = \langle v, f \rangle_2 \quad , \quad u, v \in H_0^{k,2}(\Omega) \quad ,$$

is still another variational formulation of the Dirichlet problem (12.2.3). The sesquilinear form D is symmetric. Moreover, it follows from Gårding's inequality, Theorem 12.4.5, that D is weakly coercive on $H_0^{k,2}(\Omega)$.

The conclusions of the theorem are direct consequences of the conclusions of Theorem 12.5.7 except for items (2) and (3). For item (2), we have from Theorem 12.5.7 that $u_j \in C^\infty(\Omega)$ for all j . For all j , the continuous extension of u_j and its partial derivatives to $\bar{\Omega}$ is proved in Section 12.6.2 later. As for item (3), this result follows from Proposition 11.8.14. Since $u_j \in H_0^{k,2}(\Omega) \cap C^\infty(\bar{\Omega})$, Proposition 11.8.14 implies that $D^\alpha u_j(\mathbf{x}) = 0$ for all $\mathbf{x} \in \partial\Omega$ and $|\alpha| < k$. ■

Corollary 12.5.9

Let Ω be a bounded open subset of $\mathbb{R}^n$ with a boundary of class C^2 . There exists an orthonormal basis $\{u_j\}_{j=1}^\infty$ of $L^2(\Omega)$ consisting of eigenvectors of $-\Delta$ such that $u_j \in C^\infty(\bar{\Omega})$ and $u_j = 0$ on $\partial\Omega$ for all j . The eigenvalues are positive and can be arranged to form an increasing sequence in absolute value that converges to $+\infty$.

Proof.

The operator $L(\mathbf{x}, D) \equiv -\Delta$ is strongly elliptic and $L^*(\mathbf{x}, D) = L(\mathbf{x}, D)$. We may then apply the previous theorem. Note that the variational formulation of $L(\mathbf{x}, D)u = -\Delta u = f \in L^2(\Omega)$ is

$$B(v, u) = \sum_{j=1}^n \left\langle \frac{\partial v}{\partial x_j}, \frac{\partial u}{\partial x_j} \right\rangle_2 = \langle v, f \rangle_2 \quad , \quad u, v \in H_0^{1,2}(\Omega) \quad .$$

The sesquilinear form B is symmetric and coercive on $H_0^{1,2}(\Omega)$. ■

Example 12.5.10

The conclusions of this example can all be directly deduced from Theorem 12.5.8, however we deduce the conclusions using basic results in functional analysis. These results are those used

to proof Theorem 12.5.8. We hope that this example may provide a better understanding of Theorem 12.5.8 and its proof.

Let $I =]0, 1[$. Suppose that $p \in C^1(\bar{I})$ and that there exists $\alpha > 0$ such that $p(x) \geq \alpha$ for all $x \in I$. Moreover, suppose that $q \in C(\bar{I})$.

Consider the problem

$$\begin{aligned} -(pu')' + qu &= \lambda u \quad \text{on } I \\ u(0) &= u(1) = 0 \end{aligned} \tag{12.5.20}$$

We show that there exist a sequence $\{\lambda_j\}_{j=1}^\infty$ of real numbers and a sequence $\{u_j\}_{j=1}^\infty$ of functions in $L^2(I)$ such that

1. u_j is a weak solution of (12.5.20) when $\lambda = \lambda_j$.
2. $\{u_j\}_{j=1}^\infty$ is a basis of $L^2(I)$.
3. $\lambda_j \leq \lambda_{j+1}$ for all j and $\lambda_j \rightarrow \infty$ as $j \rightarrow \infty$.
4. $u_j \in C^2(\bar{I})$ for all j , and the u_j 's are classical solutions of (12.5.20) for $\lambda = \lambda_j$.

Note that we may have $\lambda_{j_1} = \lambda_{j_2}$ with $j_1 < j_2$ if there are $j_2 - j_1 + 1$ linearly independent weak solutions of (12.5.20) for $\lambda = \lambda_{j_1} = \lambda_{j_2}$.

The variational problem associated to (12.5.20) is

$$\int_I (p u' v' + q u v) \, dx = \lambda \int_I u v \, dx \tag{12.5.21}$$

for all $u, v \in H_0^1(I)$. This is a motivation to consider the variational problem

$$\int_I (p u' v' + q u v) \, dx = \int_I f v \, dx \tag{12.5.22}$$

for all $u, v \in H_0^1(I)$. We have seen in Example 12.1.2 that for each $f \in L^2(I)$ there exists a unique solution $u \in H^2(I) \cap H_0^1(I)$ of this variational problem. We define an operator $T : L^2(I) \rightarrow L^2(I)$ by $T(f) = u$, where $u \in H^2(I) \cap H_0^1(I)$ is the weak solution of the variational problem (12.5.22). In this context, $1/\lambda_j$ is an eigenvalue of T and u_j is an eigenvector associated to $1/\lambda_j$ because u_j and λ_j satisfy $T(\lambda_j u_j) = u_j$.

Without loss of generality, we may assume that $q(x) \geq 0$ for all $x \in I$. If it is not so, let $C = \min_{x \in \bar{I}} q(x)$ and replace q by $q + C$. The values λ_j are then all shifted to $\lambda_j + C$.

We now prove that T is a self-adjoint, compact operator. We first prove that $T : L^2(I) \rightarrow H_0^1(I)$ is continuous. Suppose that $T(f) = u$. From (12.5.22), we have

$$\alpha \|u'\|_2^2 = \int_I \alpha (u')^2 \, dx \leq \int_I p (u')^2 \, dx \leq \int_I f u \, dx \leq \|f\|_2 \|u\|_2 \tag{12.5.23}$$

because $p(x) \geq \alpha$ and $q(x) \geq 0$ for all x . It follows from Poincaré's Inequality, Theorem 11.7.11, that there exists $C > 0$ independent of u such that

$$\|u\|_{1,2} \leq C \|u'\|_2. \tag{12.5.24}$$

We may assume that $C > 1$. Thus, $\|u\|_2^2 + \|u'\|_2^2 \leq C^2 \|u'\|_2^2$ yields $\|u\|_2 \leq \sqrt{C^2 - 1} \|u'\|_2$. We then have from (12.5.23) that $\alpha \|u'\|_2^2 \leq \sqrt{C^2 - 1} \|u'\|_2 \|f\|_2$ and therefore $\|u'\|_2 \leq \frac{\sqrt{C^2 - 1}}{\alpha} \|f\|_2$. We finally get from (12.5.24) that

$$\|T(f)\|_{1,2} = \|u\|_{1,2} \leq \frac{C\sqrt{C^2 - 1}}{\alpha} \|f\|_2.$$

Since the injection of $H^1(I)$ into $C(\bar{I}) \subset L^2(I)$ is compact according to Corollary 11.7.9 to the Sobolev Lemma, we get that $T : L^2(I) \rightarrow L^2(I)$ is compact because the composition of a continuous operator with a compact operator is a compact operator.

To prove that T is self-adjoint, we need to show that

$$\int_I T(f) g \, dx = \int_I f T(g) \, dx \quad (12.5.25)$$

for all $f, g \in L^2(I)$. Let $u = T(f)$ and $v = T(g)$. We then have that

$$\int_I f v \, dx = \int_I (p u' v' + q u v) \, dx = \int_I g u \, dx.$$

This is (12.5.25).

Since T is a self-adjoint compact operator, the spectrum of T contains only real eigenvalues, there is at most a countable number of eigenvalues, and the eigenspace associate to an eigenvalue is of finite dimension. Since $L^2(I)$ is of infinite dimension, there is therefore an infinite number of distinct eigenvalues. Moreover, we may arrange the real eigenvalues $\{\mu_j\}_{j=1}^\infty$ such that $|\mu_j| \geq |\mu_{j+1}|$, and we have $|\mu_j| \rightarrow 0$ as $j \rightarrow \infty$.

If we take a basis of eigenfunctions for each eigenspace and combine them, we get a basis of eigenfunctions $\{u_j\}_{j=1}^\infty$ for $L^2(I)$ (see Theorem 4.4.44 and the paragraphs before this theorem). We assume that $\mu_j = \mu_{j+1} = \dots = \mu_{j+k}$ if the eigenfunctions $u_j, u_{j+1}, \dots, u_{j+k}$ are $k+1$ linearly independent eigenfunctions associated to μ_j .

Since $\mathcal{N}(T) = \{0\}$, the unique solution of (12.5.22) with $f = 0$ is $u = 0$; namely, 0 is not an eigenvalue. Thus $\mu_j \neq 0$ for all j . Moreover, we have

$$\int_I T(f) f \, dx = \int_I u f \, dx = \int_I (p(u')^2 + q u^2) \, dx \geq 0$$

for all $f \in L^2(I)$. Thus

$$\mu_j \int_I u_j^2 \, dx = \int_I T(u_j) u_j \, dx \geq 0$$

implies that $\mu_j \geq 0$ for all j . Hence, we have $\mu_j \geq \mu_{j+1} > 0$ for all j and $\mu_j \rightarrow 0$ as $j \rightarrow \infty$.

Except for the last item about the differentiability of the u_j 's, we have proved all the other claims of the example with $\lambda_j = 1/\mu_j$ for all j .

It reminds to prove that $u_j \in C^2(\bar{I})$ for all j , and they are classical solutions of (12.5.20) when $\lambda = \lambda_j$.

We have shown in Example 12.1.2 that $u = T(f) \in C^1(\bar{I})$ if $f \in L^2(I)$. For the variational problem (12.5.21), we then have that $u_j \in C^1(\bar{I})$ and thus $f = \lambda_j u_j \in C^1(\bar{I})$. Again in Example 12.1.2, we have shown that if $f \in C(I)$, then $u \in C^2(\bar{I})$. Because $f = \lambda_j u_j \in C^1(\bar{I})$ in the variational problem (12.5.21), we finally obtain that $u_j \in C^2(\bar{I})$.

The justification given at the end of Example 12.1.2 proves that the weak solution $u_j \in C^2(\bar{I}) \cap H_0^1(I)$ of (12.5.20) is in fact a classical solution of (12.5.20) with $\lambda = \lambda_j$. ♣

To study the Neumann's problem (12.2.6), we need the following lemma.

Lemma 12.5.11

Suppose that the coefficients in the differential operator $L(\mathbf{x}, D)$ defined in (12.2.6) are bounded functions on Ω and, moreover, the coefficients $b_{i,j}$ are all real-valued functions. Let $B : H^{1,2}(\Omega) \times H^{1,2}(\Omega) \rightarrow \mathbb{R}$ be the sesquilinear form defined in (12.2.8) and associated to (12.2.6). If $L(\mathbf{x}, D)$ is strongly elliptic on Ω , then B is weakly coercive over $H^{1,2}(\Omega)$, and over any closed subspaces of $H^{1,2}(\Omega)$.

Proof.

Let $c_{i,j} = (b_{i,j} + b_{j,i})/2$. Then, for each $\mathbf{x} \in \Omega$,

$$R(\mathbf{x}, \boldsymbol{\xi}) = \sum_{i,j=1}^n c_{i,j}(\mathbf{x}) \xi_i \xi_j$$

for all $\boldsymbol{\xi} \in \mathbb{R}^n$ is a symmetric bilinear form. Moreover, from the definition of a strongly elliptic operator, there exists a constant C_1 independent of $\mathbf{x} \in \Omega$ such that

$$R(\mathbf{x}, \boldsymbol{\xi}) = Q(\mathbf{x}, \boldsymbol{\xi}) \equiv \sum_{i,j=1}^n b_{i,j}(\mathbf{x}) \xi_i \xi_j \geq C_1 \|\boldsymbol{\xi}\|^2$$

for all $\boldsymbol{\xi} \in \mathbb{R}^n$. Recall that Q is the principal symbol associated to (12.2.6). The matrix C with entries $c_{i,j}$ for $1 \leq i, j \leq n$ is therefore a symmetric positive definite $n \times n$ matrix. Hence,

$$\operatorname{Re} \sum_{i,j=1}^n b_{i,j}(\mathbf{x}) \xi_i \bar{\xi}_j = \sum_{i,j=1}^n c_{i,j}(\mathbf{x}) \xi_i \bar{\xi}_j \geq C_1 \|\boldsymbol{\xi}\|^2 \quad (12.5.26)$$

for all $\boldsymbol{\xi} \in \mathbb{C}^n$. If $\boldsymbol{\xi} = \nabla u(\mathbf{x})$, then

$$\operatorname{Re} \sum_{i,j=1}^n b_{i,j}(\mathbf{x}) D_{x_i} u(\mathbf{x}) \overline{D_{x_j} u(\mathbf{x})} \geq C_1 \|\nabla u(\mathbf{x})\|^2 = C_1 \sum_{j=1}^n |D_{x_j} u(\mathbf{x})|^2$$

for all $\mathbf{x} \in \Omega$. It follows that

$$\operatorname{Re} \sum_{i,j=1}^n \int_{\Omega} b_{i,j}(\mathbf{x}) D_{x_i} u(\mathbf{x}) \overline{D_{x_j} u(\mathbf{x})} d\mathbf{x} \geq C_1 \sum_{j=1}^n \int_{\Omega} |D_{x_j} u(\mathbf{x})|^2 d\mathbf{x} = C_1 (\|u\|_{1,2,\Omega}^2 - \|u\|_{2,\Omega}^2) . \quad (12.5.27)$$

Moreover, let C_2 be a constant such that $|b_i(\mathbf{x})| \leq C_2$ for all $\mathbf{x} \in \Omega$ and $0 \leq i \leq n$. Using Schwarz inequality, we find that

$$|\langle u, b_i D_{x_i} u \rangle_2| \leq \|u\|_{2,\Omega} \|b_i D_{x_i} u\|_{2,\Omega} \leq C_2 \|u\|_{2,\Omega} \|u\|_{1,2,\Omega} \quad (12.5.28)$$

for $0 < i \leq n$ and

$$|\langle u, b_0 u \rangle_2| \leq \|u\|_{2,\Omega} \|b_0 u\|_{2,\Omega} \leq C_2 \|u\|_{2,\Omega} \|u\|_{1,2,\Omega} . \quad (12.5.29)$$

Hence, it follows from (12.5.27), (12.5.28) and (12.5.29) that

$$\operatorname{Re} B(u, u) \geq C_1 (\|u\|_{1,2,\Omega}^2 - \|u\|_{2,\Omega}^2) - (n+1)C_2 \|u\|_{2,\Omega} \|u\|_{1,2,\Omega}$$

for all $u \in H^{1,2}(\Omega)$. From $ab \leq (a^2 + b^2)/2$ with $a = \sqrt{C_1} \|u\|_{1,2,\Omega}$ and $b = \frac{(n+1)C_2}{\sqrt{C_1}} \|u\|_{2,\Omega}$, we get

$$(n+1)C_2 \|u\|_{2,\Omega} \|u\|_{1,2,\Omega} \leq \frac{1}{2} \left(C_1 \|u\|_{1,2,\Omega}^2 + \frac{(n+1)^2 C_2^2}{C_1} \|u\|_{2,\Omega}^2 \right)$$

for all $u \in H^{1,2}(\Omega)$. Therefore,

$$\begin{aligned} \operatorname{Re} B(u, u) &\geq C_1 (\|u\|_{1,2,\Omega}^2 - \|u\|_{2,\Omega}^2) - \frac{1}{2} \left(C_1 \|u\|_{1,2,\Omega}^2 + \frac{(n+1)^2 C_2^2}{C_1} \|u\|_{2,\Omega}^2 \right) \\ &= \frac{C_1}{2} \|u\|_{1,2,\Omega}^2 - \left(C_1 + \frac{(n+1)^2 C_2^2}{2C_1} \right) \|u\|_{2,\Omega}^2 \geq C \|u\|_{1,2,\Omega}^2 - \lambda \|u\|_{2,\Omega}^2 \end{aligned}$$

for all $u \in H^{1,2}(\Omega)$, where $C = \frac{C_1}{2}$ and $\lambda = C_1 + \frac{(n+1)^2 C_2^2}{2C_1}$. This proves that B is weakly coercive on $H^{1,2}(\Omega)$. $\blacksquare$

Before stating our existence result for the Neumann problem, we need to recall that the adjoint of the operator

$$L(\mathbf{x}, D)u = - \sum_{i,j=1}^n \frac{\partial}{\partial x_i} \left(b_{i,j}(\mathbf{x}) \frac{\partial u}{\partial x_j} \right) + b_0(\mathbf{x})u$$

given in (12.2.6) is the operator

$$L^*(\mathbf{x}, D)u = - \sum_{i,j=1}^n \frac{\partial}{\partial x_j} \left(b_{i,j}(\mathbf{x}) \frac{\partial u}{\partial x_i} \right) + b_0(\mathbf{x})u$$

where we have assumed that the functions $b_{i,j}$ were real-valued functions of class $C^\infty(\Omega)$. If the matrix B with entries $b_{i,j}$ for $1 \leq i, j \leq n$ is symmetric, then

$$\sum_{i,j=1}^n \frac{\partial}{\partial x_i} \left(b_{i,j} \frac{\partial u}{\partial x_j} \right) = \sum_{i,j=1}^n \frac{\partial}{\partial x_i} \left(b_{j,i} \frac{\partial u}{\partial x_j} \right) = \sum_{i,j=1}^n \frac{\partial}{\partial x_j} \left(b_{i,j} \frac{\partial u}{\partial x_i} \right)$$

for $u \in C^2(\Omega)$, where we have use the symmetry of $b_{i,j}$ for the first equality and interchanged i and j for the second equality. We get that $L(\mathbf{x}, D) = L^*(\mathbf{x}, D)$.

Remark 12.5.12

If we expand

$$L(\mathbf{x}, D)u = - \sum_{i,j=1}^n \frac{\partial}{\partial x_i} \left(b_{i,j}(\mathbf{x}) \frac{\partial u}{\partial x_j} \right) + b_0(\mathbf{x})u ,$$

we get

$$L(\mathbf{x}, D)u = - \sum_{i,j=1}^n b_{i,j}(\mathbf{x}) \frac{\partial^2 u}{\partial x_i \partial x_j} - \sum_{i=1}^n \left(\sum_{j=1}^n \frac{\partial b_{j,i}}{\partial x_j}(\mathbf{x}) \right) \frac{\partial u}{\partial x_i} + b_0(\mathbf{x})u .$$

If we compare with

$$L(\mathbf{x}, D)u = - \sum_{i,j=1}^n a_{i,j}(\mathbf{x}) \frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^n a_i(\mathbf{x}) \frac{\partial u}{\partial x_i} + a_0(\mathbf{x})u ,$$

we get $a_{i,j} = b_{i,j}$, $a_i = - \sum_{j=1}^n \frac{\partial b_{j,i}}{\partial x_j}$ and $b_0 = a_0$. The condition $a_i = - \sum_{j=1}^n \frac{\partial a_{j,i}}{\partial x_j}$ for $1 \leq i \leq n$ is the condition for $L(\mathbf{x}, D)u$ in standard form to be equal to its adjoint

$$L^*(\mathbf{x}, D)u = - \sum_{i,j=1}^n a_{i,j}(\mathbf{x}) \frac{\partial^2 u}{\partial x_i \partial x_j} - \sum_{i=1}^n \frac{\partial}{\partial x_i} (a_i(\mathbf{x})u) + a_0(\mathbf{x})u$$

if we assume that the matrix with entries $a_{i,j}$ is symmetric. We leave it to the reader to verify this statement. ♠

Theorem 12.5.13

Let Ω be a bounded open subset of $\mathbb{R}^n$ with a boundary of class C^2 . Suppose that the operator $L(\mathbf{x}, D)$ in the Neumann problem (12.2.6) is elliptic on $\overline{\Omega}$, that $L(\mathbf{x}, D) = L^*(\mathbf{x}, D)$, that the coefficients $b_{i,j}$ of $L(\mathbf{x}, D)$ are real-valued functions of class $C^\infty(\Omega)$, and that the $n \times n$ matrix B with entries $b_{i,j}$ for $1 \leq i, j \leq n$ is symmetric. We have the following results.

1. There exists an orthonormal basis $\{u_j\}_{j=1}^\infty$ of $L^2(\Omega)$ consisting of eigenvectors of $L(\mathbf{x}, D)$.
2. $u_j \in C^\infty(\overline{\Omega})$ for all j .
3. If $\nu(\mathbf{x})$ denotes the outward unit normal to $\partial\Omega$ at $\mathbf{x} \in \partial\Omega$, then

$$\sum_{i,j=1}^n b_{i,j}(\mathbf{x}) \frac{\partial u_k}{\partial x_j}(\mathbf{x}) \nu_i(\mathbf{x}) = 0$$

on $\partial\Omega$ for all k .

4. The eigenvalues are real, isolated and can be arranged to form an increasing sequence in absolute value that converges to $+\infty$.
5. If $b_0 \geq 0$ on Ω , the eigenvalues are non-negative. If in addition $b_0 > 0$ on an open subset of Ω , the eigenvalues are positive.

Proof.

Recall from Remark 2.4.13 that $L(\mathbf{x}, D)$ elliptic on a compact set implies that $L(\mathbf{x}, D)$ is

strongly elliptic on that set. We then have from Lemma 12.5.11 that the bilinear form B defined in (12.2.8) is weakly coercive over $H^{1,2}(\Omega)$.

Hence, items (1) and (4) comes from Theorem 12.5.7. Item (2) is a regularity results and will be proved in Section 12.6.2.

Item (3) comes from (12.2.11), where u is replaced by u_k and f by $\lambda_k u_k$ with λ_k the eigenvalue associated to u_k .

To prove item (5), we refer to (12.5.26) in the proof of Lemma 12.5.11. Since the matrix B is symmetric, $c_{i,j} = b_{i,j}$ in the proof of Lemma 12.5.11. Thus (12.5.26) becomes

$$\sum_{i,j=1}^n b_{i,j}(\mathbf{x}) \xi_i \bar{\xi}_j \geq C_1 \|\boldsymbol{\xi}\|^2$$

for all $\boldsymbol{\xi} \in \mathbb{C}^n$ for some constant C_1 . As we have proved before the statement of the theorem, we have

$$\begin{aligned} B(u, u) &= \sum_{i,j=1}^n \int_{\Omega} b_{i,j}(\mathbf{x}) \frac{\partial u}{\partial x_i} \frac{\partial \bar{u}}{\partial x_j} d\mathbf{x} + \int_{\Omega} b_0 |u|^2 d\mathbf{x} \\ &\geq C_1 \int_{\Omega} \sum_{i=1}^n |D_{x_i} u|^2 d\mathbf{x} + \int_{\Omega} b_0 u^2 d\mathbf{x} \geq \int_{\Omega} b_0 |u|^2 d\mathbf{x} \end{aligned}$$

for $u \in H^{1,2}(\Omega)$. Hence, if λ_j is an eigenvalue associated to the eigenvector u_j , then $\lambda_j \in \mathbb{R}$ by item (4) and

$$\lambda_j \|u_j\|_2^2 = \langle u_j, \lambda_j u_j \rangle_2 = \langle u_j, F_B(u_j) \rangle_2 = B(u_j, u_j) \geq \int_{\Omega} b_0 |u_j|^2 d\mathbf{x}.$$

Since u_j and b_0 are at least continuous function on Ω , we have that $\lambda_j \geq 0$ if $b_0 \geq 0$ on Ω , and $\lambda_j > 0$ if in addition of $b_0 \geq 0$ on Ω we have that $b_0 > 0$ on an open subset of Ω . ■

12.5.1 Eigenvalue Problems for the Laplacian Operator (continued)

We can now look back at the eigenvalue problems for the Laplacian operator that we studied in Section 9.11 and provide a complete justification for the eigenvalues and eigenfunctions that we found.

As before, let Ω be an open and bounded subset of $\mathbb{R}^n$ with a sufficiently smooth boundary. We seek λ such that there exists a non-trivial function $u \in H_0^1(\Omega)$ satisfying $-\Delta u = \lambda u$ on Ω .

The **variational formulation** of the eigenvalue problem $-\Delta u = \lambda u$ is to find $u \in H_0^1(\Omega)$ such that

$$B(u, v) \equiv \int_{\Omega} \nabla u(\mathbf{x}) \cdot \nabla v(\mathbf{x}) d\mathbf{x} = \lambda \int_{\Omega} u(\mathbf{x}) v(\mathbf{x}) d\mathbf{x} \quad (12.5.30)$$

for all $v \in H_0^1(\Omega)$.

To justify this formulation, suppose that $u \in C(\overline{\Omega}) \cap C^1(\Omega)$ is a non-trivial classical solution of $-\Delta u = \lambda u$ for a given λ . From the Green's identity (9.1.1), we have

$$\begin{aligned} \int_{\Omega} \nabla u(\mathbf{x}) \cdot \nabla v(\mathbf{x}) \, d\mathbf{x} &= \int_{\partial\Omega} v(\mathbf{x}) \frac{\partial u}{\partial \nu}(\mathbf{x}) \, dS_{\mathbf{x}} - \int_{\Omega} v(\mathbf{x}) \Delta u(\mathbf{x}) \, d\mathbf{x} \\ &= - \int_{\Omega} v(\mathbf{x}) \Delta u(\mathbf{x}) \, d\mathbf{x} = \lambda \int_{\Omega} v(\mathbf{x}) u(\mathbf{x}) \, d\mathbf{x} \end{aligned}$$

for all functions $v \in \mathcal{D}(\Omega)$. Thus, (12.5.30) follows by density of $\mathcal{D}(\Omega)$ in $H_0^1(\Omega)$. Recall that $\frac{\partial u}{\partial \nu}(\mathbf{x})$ is the directional derivative of u at $\mathbf{x} \in \partial\Omega$ in the direction of the outward unit normal $\nu(\mathbf{x})$ to the surface $\partial\Omega$ at $\mathbf{x}$.

Since $-\Delta$ is strongly elliptic and the sesquilinear form B of the variational formulation (12.5.30) is also coercive, we may use Theorem 12.5.8 to get a complete description of the eigenvalues and eigenfunctions of $-\Delta$ with their properties.

It is interesting to prove the orthogonality of eigenfunctions associated to different eigenvalues. Suppose that λ_1 and λ_2 are two distinct eigenvalues of $-\Delta$, $u_1 \in L^2(\Omega)$ is an eigenfunction associate to λ_1 and $u_2 \in L^2(\Omega)$ is an eigenfunction associate to λ_2 . Then, it follows from (12.5.30) that

$$\lambda_2 \int_{\Omega} u_1(\mathbf{x}) u_2(\mathbf{x}) \, d\mathbf{x} = \int_{\Omega} \nabla u_1(\mathbf{x}) \cdot \nabla u_2(\mathbf{x}) \, d\mathbf{x} = \lambda_1 \int_{\Omega} u_1(\mathbf{x}) u_2(\mathbf{x}) \, d\mathbf{x}.$$

This gives

$$(\lambda_2 - \lambda_1) \int_{\Omega} u_1(\mathbf{x}) u_2(\mathbf{x}) \, d\mathbf{x} = 0.$$

Since $\lambda_2 - \lambda_1 \neq 0$, this shows that u_1 and u_2 are orthogonal with respect to the scalar product $\langle f, g \rangle = \int_{\Omega} f g \, d\mathbf{x}$ in $L^2(\Omega)$.

Let $\{u_j\}_{j=1}^{\infty}$ be an orthogonal basis of eigenfunctions for $-\Delta$. We assume that u_j is an eigenfunction associated to the eigenvalue λ_j and that this eigenvalue is repeated k times in $\lambda_1 \leq \lambda_2 \leq \dots$ if the dimension of the eigenspace is k .

Example 12.5.14

Fourier series can be used to solve the Poisson equation $-\Delta u = f$ with $u = 0$ on $\partial\Omega$ if $f \in L^2(\Omega)$. Suppose that $f = \sum_{j=1}^{\infty} a_j u_j$ is the Fourier series expansion of f with $a_j = \frac{\langle f, u_j \rangle_2}{\langle u_j, u_j \rangle_2}$.

If $u = \sum_{j=1}^{\infty} b_j u_j$ is the solution the Poisson equation, then a substitution of these two series in $-\Delta u = f$ yields $\lambda_j b_j = a_j$ for all j . Thus $b_j = a_j / \lambda_j$ for $j > 0$. Recall that the series above converges with respect to the L^2 -norm on $L^2(\Omega)$ and that $-\Delta u = f$ in the variational sense.

♣

Example 12.5.15

We may express the Green function on Ω in terms of the basis of eigenfunctions. Suppose that $G : \Omega \times \overline{\Omega} \rightarrow \mathbb{R}$ is the Green function on Ω . According to Theorem 9.5.5, the solution u_j of $\Delta u = -\lambda_j u$ with $u = 0$ on $\partial\Omega$ satisfies

$$u_j(\mathbf{x}) = -\lambda_j \int_{\Omega} G(\mathbf{x}, \mathbf{y}) u_j(\mathbf{y}) \, d\mathbf{y}$$

for all $\mathbf{x} \in \Omega$. We have assumed that $u_j \in L^1(\Omega)$.

The following discussion could be justified rigorously. For $\mathbf{x} \in \Omega$ fixed, the Fourier series of the function $\mathbf{y} \mapsto G(\mathbf{x}, \mathbf{y})$ for $\mathbf{y} \in \Omega$ is

$$G(\mathbf{x}, \cdot) = \sum_{j=1}^{\infty} a_j(\mathbf{x}) u_j(\cdot) ,$$

where this series converges with respect to the standard L^2 -norm on $L^2(\Omega)$ and

$$a_j(\mathbf{x}) = \left(\int_{\Omega} |u_j(\mathbf{y})|^2 d\mathbf{y} \right)^{-1} \int_{\Omega} G(\mathbf{x}, \mathbf{y}) u_j(\mathbf{y}) d\mathbf{y} = -\frac{u_j(\mathbf{x})}{\lambda_j \langle u_j, u_j \rangle} .$$

Therefore,

$$G(\mathbf{x}, \mathbf{y}) = - \sum_{j=1}^{\infty} \frac{u_j(\mathbf{y}) u_j(\mathbf{x})}{\lambda_j \langle u_j, u_j \rangle} ,$$

where this series converges with respect to the standard L^2 -norm on $L^2(\Omega \times \Omega)$. ♣

12.6 Regularity of Weak Solutions

Let Ω be an open subset of $\mathbb{R}^n$. In this section, we consider the elliptic operator

$$L(\mathbf{x}, D) = \sum_{|\alpha| \leq k} a_{\alpha}(\mathbf{x}) D^{\alpha} \tag{12.6.1}$$

with $a_{\alpha} \in C^{\infty}(\Omega)$ for all α . Some additional constraints may be imposed on the a_{α} in some occasions.

The goal of this section is to determine in which space the solution of an elliptic partial differential equation belongs to. The analysis is divided in two sections, the regularity of the solution in the interior of Ω (local regularity) and the regularity of the solution at the boundary of Ω (including on the boundary).

Before talking about the regularity of solutions, we need to present some results and tools that will be useful for this purpose.

Lemma 12.6.1

Given $s \in \mathbb{R}$ and $t > n/2$, there exists a constant C such that

$$\|\phi f\|_{s,\rho} \leq (2\pi)^{n/2} \sup_{\mathbf{x} \in \mathbb{R}^n} |\phi(\mathbf{x})| \|f\|_{s,\rho} + C \|\phi\|_{|s-1|+1+t,\rho} \|f\|_{s-1,\rho}$$

for all $\phi \in \mathcal{S}(\mathbb{R}^n)$ and $f \in H^{s,2}(\mathbb{R}^n)$.

Proof.

i) We first prove that there exists C such that

$$\left(\int_{\mathbb{R}^n} \left| \int_{\mathbb{R}^n} ((1 + \|\mathbf{x}\|_2^2)^{s/2} - (1 + \|\mathbf{y}\|_2^2)^{s/2}) \hat{\phi}(\mathbf{x} - \mathbf{y}) \hat{f}(\mathbf{y}) d\mathbf{y} \right|^2 d\mathbf{x} \right)^{1/2} \leq C \|\phi\|_{|s-1|+1+t,\rho} \|f\|_{s-1,\rho} \quad (12.6.2)$$

for all $\phi \in \mathcal{S}(\mathbb{R}^n)$ and $f \in H^{s-1,2}(\mathbb{R}^n)$. Choose $v \in L^2(\mathbb{R}^n)$ such that $\hat{v}(\mathbf{y}) = (1 + \|\mathbf{y}\|_2^2)^{(s-1)/2} \hat{f}(\mathbf{y})$ for $\mathbf{y} \in \mathbb{R}^n$. Such a v exists according to Plancherel theorem, Theorem 6.6.6, because $f \in H^{s-1,2}(\mathbb{R}^n)$ implies that $(1 + \|\mathbf{y}\|_2^2)^{(s-1)/2} \hat{f} \in L^2(\mathbb{R}^n)$ by definition. We may restate (12.6.2) as

$$\left(\int_{\mathbb{R}^n} \left| \int_{\mathbb{R}^n} ((1 + \|\mathbf{x}\|_2^2)^{s/2} - (1 + \|\mathbf{y}\|_2^2)^{s/2}) (1 + \|\mathbf{y}\|_2^2)^{(1-s)/2} \hat{\phi}(\mathbf{x} - \mathbf{y}) \hat{v}(\mathbf{y}) d\mathbf{y} \right|^2 d\mathbf{x} \right)^{1/2} \leq C \|\phi\|_{|s-1|+1+t,\rho} \|v\|_{0,\rho} \quad (12.6.3)$$

for all $\phi \in \mathcal{S}(\mathbb{R}^n)$ and $v \in H^{0,2}(\mathbb{R}^n) = L^2(\mathbb{R}^n)$.

We have that

$$\int_{\mathbb{R}^n} ((1 + \|\mathbf{x}\|_2^2)^{s/2} - (1 + \|\mathbf{y}\|_2^2)^{s/2}) (1 + \|\mathbf{y}\|_2^2)^{(1-s)/2} \hat{\phi}(\mathbf{x} - \mathbf{y}) \hat{v}(\mathbf{y}) d\mathbf{y} = \int_{\mathbb{R}^n} K(\mathbf{x}, \mathbf{y}) \hat{v}(\mathbf{y}) d\mathbf{y}$$

with

$$K(\mathbf{x}, \mathbf{y}) = ((1 + \|\mathbf{x}\|_2^2)^{s/2} - (1 + \|\mathbf{y}\|_2^2)^{s/2}) (1 + \|\mathbf{y}\|_2^2)^{(1-s)/2} \hat{\phi}(\mathbf{x} - \mathbf{y}) .$$

Our goal is to apply the Generalized Young's Inequality, Theorem 6.5.1, to show that

$$\left\| \int_{\mathbb{R}^n} K(\mathbf{x}, \mathbf{y}) \hat{v}(\mathbf{y}) d\mathbf{y} \right\|_2 \leq C_1 \|\hat{v}\|_2 = C_1 \|v\|_{s-1,\rho}$$

for a constant C_1 such that $\sup_{\mathbf{x} \in \mathbb{R}^n} \int_{\mathbb{R}^n} |K(\mathbf{x}, \mathbf{y})| d\mathbf{y} \leq C_1$ and $\sup_{\mathbf{y} \in \mathbb{R}^n} \int_{\mathbb{R}^n} |K(\mathbf{x}, \mathbf{y})| d\mathbf{x} \leq C_1$. This will give us (12.6.3) after we show that we can take $C_1 = C \|\phi\|_{|s-1|+1+t,\rho}$ for some C .

Since $\frac{\partial}{\partial w}(1 + w^2)^{s/2} = sw(1 + w^2)^{(s-2)/2}$, we get from the Mean Value Theorem that

$$(1 + \|\mathbf{x}\|_2^2)^{s/2} - (1 + \|\mathbf{y}\|_2^2)^{s/2} = sw(1 + w^2)^{(s-2)/2} (\|\mathbf{x}\|_2 - \|\mathbf{y}\|_2)$$

for some w between $\|\mathbf{x}\|_2$ and $\|\mathbf{y}\|_2$. Hence,

$$\begin{aligned} |(1 + \|\mathbf{x}\|_2^2)^{s/2} - (1 + \|\mathbf{y}\|_2^2)^{s/2}| &\leq |s| \sup_w |w(1 + w^2)^{(s-2)/2}| |\|\mathbf{x}\|_2 - \|\mathbf{y}\|_2| \\ &\leq |s| \sup_w |(1 + w^2)^{1/2} (1 + w^2)^{(s-2)/2}| |\|\mathbf{x}\|_2 - \|\mathbf{y}\|_2| \leq |s| \sup_w |(1 + w^2)^{(s-1)/2}| \|\mathbf{x} - \mathbf{y}\|_2 , \end{aligned}$$

where the supremum is taken over all w between $\|\mathbf{x}\|_2$ and $\|\mathbf{y}\|_2$. Since $w \rightarrow (1 + w^2)^{(s-1)/2}$ is increasing on $[0, \infty[$ for $s > 1$ or decreasing on $[0, \infty[$ for $s < 1$, we get that

$$|(1 + \|\mathbf{x}\|_2^2)^{s/2} - (1 + \|\mathbf{y}\|_2^2)^{s/2}| \leq |s| ((1 + \|\mathbf{x}\|_2^2)^{(s-1)/2} + (1 + \|\mathbf{y}\|_2^2)^{(s-1)/2}) \|\mathbf{x} - \mathbf{y}\|_2 ,$$

It follows from this relation and Lemma 11.1.21 that

$$\begin{aligned}
 |K(\mathbf{x}, \mathbf{y})| &\leq |s| \left(\frac{(1 + \|\mathbf{x}\|_2^2)^{(s-1)/2}}{(1 + \|\mathbf{y}\|_2^2)^{(s-1)/2}} + 1 \right) \|\mathbf{x} - \mathbf{y}\|_2 |\hat{\phi}(\mathbf{x} - \mathbf{y})| \\
 &\leq |s| \left(2^{|s-1|/2} (1 + \|\mathbf{x} - \mathbf{y}\|_2^2)^{|s-1|/2} + 1 \right) \|\mathbf{x} - \mathbf{y}\|_2 |\hat{\phi}(\mathbf{x} - \mathbf{y})| \\
 &\leq 2^{1+|s-1|/2} |s| (1 + \|\mathbf{x} - \mathbf{y}\|_2^2)^{|s-1|/2} (1 + \|\mathbf{x} - \mathbf{y}\|_2^2)^{1/2} |\hat{\phi}(\mathbf{x} - \mathbf{y})| \\
 &= 2^{1+|s-1|/2} |s| (1 + \|\mathbf{x} - \mathbf{y}\|_2^2)^{(|s-1|+1)/2} |\hat{\phi}(\mathbf{x} - \mathbf{y})|.
 \end{aligned}$$

Hence, from Schwarz inequality, we get

$$\begin{aligned}
 \int_{\mathbb{R}^n} |K(\mathbf{x}, \mathbf{y})| \, d\mathbf{y} &\leq 2^{1+|s-1|/2} |s| \int_{\mathbb{R}^n} (1 + \|\mathbf{x} - \mathbf{y}\|_2^2)^{(|s-1|+1)/2} |\hat{\phi}(\mathbf{x} - \mathbf{y})| \, d\mathbf{y} \\
 &\leq 2^{1+|s-1|/2} |s| \left(\int_{\mathbb{R}^n} ((1 + \|\mathbf{x} - \mathbf{y}\|_2^2)^{(|s-1|+1+t)/2} |\hat{\phi}(\mathbf{x} - \mathbf{y})|)^2 \, d\mathbf{y} \right)^{1/2} \\
 &\quad \left(\int_{\mathbb{R}^n} ((1 + \|\mathbf{x} - \mathbf{y}\|_2^2)^{-t/2})^2 \, d\mathbf{y} \right)^{1/2} = C \|\phi\|_{|s-1|+1+t, \rho}
 \end{aligned}$$

for all $\mathbf{x} \in \mathbb{R}^n$ with $C = 2^{1+|s-1|/2} |s| \left(\int_{\mathbb{R}^n} (1 + \|\mathbf{x} - \mathbf{y}\|_2^2)^{-t} \, d\mathbf{y} \right)^{1/2}$.

A very similar reasoning shows that $\sup_{\mathbf{y} \in \mathbb{R}^n} \int_{\mathbb{R}^n} |K(\mathbf{x}, \mathbf{y})| \, d\mathbf{x} \leq C \|\phi\|_{|s-1|+1+t, \rho}$.

ii) Using triangle inequality on $L^2(\mathbb{R}^n)$, we find that

$$\begin{aligned}
 \|\phi f\|_{s, \rho} &= \left(\int_{\mathbb{R}^n} |(\phi f)^\wedge(\mathbf{x})|^2 (1 + \|\mathbf{x}\|_2^2)^s \, d\mathbf{x} \right)^{1/2} \\
 &= \left(\int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} \hat{\phi}(\mathbf{x} - \mathbf{y}) \hat{f}(\mathbf{y}) \, d\mathbf{y} \right)^2 (1 + \|\mathbf{x}\|_2^2)^s \, d\mathbf{x} \right)^{1/2} \\
 &= \left(\int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} ((1 + \|\mathbf{x}\|_2^2)^{s/2} - (1 + \|\mathbf{y}\|_2^2)^{s/2}) \hat{\phi}(\mathbf{x} - \mathbf{y}) \hat{f}(\mathbf{y}) \, d\mathbf{y} \right. \right. \\
 &\quad \left. \left. + \int_{\mathbb{R}^n} (1 + \|\mathbf{y}\|_2^2)^{s/2} \hat{\phi}(\mathbf{x} - \mathbf{y}) \hat{f}(\mathbf{y}) \, d\mathbf{y} \right)^2 \, d\mathbf{x} \right)^{1/2} \\
 &\leq \left(\int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} ((1 + \|\mathbf{x}\|_2^2)^{s/2} - (1 + \|\mathbf{y}\|_2^2)^{s/2}) \hat{\phi}(\mathbf{x} - \mathbf{y}) \hat{f}(\mathbf{y}) \, d\mathbf{y} \right)^2 \, d\mathbf{x} \right)^{1/2} \\
 &\quad + \left(\int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} (1 + \|\mathbf{y}\|_2^2)^{s/2} \hat{\phi}(\mathbf{x} - \mathbf{y}) \hat{f}(\mathbf{y}) \, d\mathbf{y} \right)^2 \, d\mathbf{x} \right)^{1/2}.
 \end{aligned}$$

From (i), we have that the first integral on the right hand side is bounded by $C \|\phi\|_{|s-1|+1+t, \rho} \|f\|_{s-1, \rho}$.

For the second integral, let $g \in L^2(\mathbb{R}^n)$ be such that $\hat{g} = (1 + \|\mathbf{y}\|_2^2)^{s/2} \hat{f}$. Such a g exists according to Plancherel theorem, Theorem 6.6.6, because $f \in H^{s,2}(\mathbb{R}^n)$ implies that $(1 + \|\mathbf{y}\|_2^2)^{s/2} \hat{f} \in L^2(\mathbb{R}^n)$ by definition. We have that

$$\left(\int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} (1 + \|\mathbf{y}\|_2^2)^{s/2} \hat{\phi}(\mathbf{x} - \mathbf{y}) \hat{f}(\mathbf{y}) \, d\mathbf{y} \right)^2 \, d\mathbf{x} \right)^{1/2} = \left(\int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} \hat{\phi}(\mathbf{x} - \mathbf{y}) \hat{g}(\mathbf{y}) \, d\mathbf{y} \right)^2 \, d\mathbf{x} \right)^{1/2}$$

$$\begin{aligned}
&= (2\pi)^{n/2} \left(\int_{\mathbb{R}^n} |\hat{\phi} * \hat{g}|^2 d\mathbf{x} \right)^{1/2} = (2\pi)^{n/2} \left(\int_{\mathbb{R}^n} |(\phi g)^\wedge|^2 d\mathbf{x} \right)^{1/2} = (2\pi)^{n/2} \|\phi g\|_{0,\rho} \\
&= (2\pi)^{n/2} \|\phi g\|_2 \leq (2\pi)^{n/2} \sup_{\mathbf{x} \in \mathbb{R}^n} |\phi(\mathbf{x})| \|g\|_2 = (2\pi)^{n/2} \sup_{\mathbf{x} \in \mathbb{R}^n} |\phi(\mathbf{x})| \|g\|_{0,\rho} \\
&= (2\pi)^{n/2} \sup_{\mathbf{x} \in \mathbb{R}^n} |\phi(\mathbf{x})| \left(\int_{\mathbb{R}^n} |\hat{g}|^2 d\mathbf{x} \right)^{1/2} = (2\pi)^{n/2} \sup_{\mathbf{x} \in \mathbb{R}^n} |\phi(\mathbf{x})| \left(\int_{\mathbb{R}^n} |(1 + \|\mathbf{y}\|_2^2)^{s/2} \hat{f}|^2 d\mathbf{x} \right)^{1/2} \\
&= (2\pi)^{n/2} \sup_{\mathbf{x} \in \mathbb{R}^n} |\phi(\mathbf{x})| \|f\|_{s,\rho},
\end{aligned}$$

where we have used Plancherel theorem for the fifth and sixth equality. $\blacksquare$

The proofs of regularity of solutions require the **method of translations**, also known as the **method of difference quotient**, that we present before turning all our attention to the goal of this section.

For $f \in L^2(\mathbb{R}^n)$ and $h \in \mathbb{R} \setminus \{0\}$, let

$$(\Delta_{he_j} f)(\mathbf{x}) = \frac{1}{h} (f(\mathbf{x} + he_j) - f(\mathbf{x})) \quad (12.6.4)$$

and, recursively, $\Delta_{he_j}^q f = \Delta_{he_j} (\Delta_{he_j}^{q-1} f)$ for $q > 1$. We also define $\Delta_{he_j}^0$ as $\Delta_{he_j}^0 f = f$ for $h \in \mathbb{R}$. For $\mathbf{h} \in \mathbb{R}^n$ and $\alpha \in \mathbb{N}^n$ with $h_j \neq 0$ for $\alpha_j \neq 0$, let

$$\Delta_{\mathbf{h}}^\alpha f = \Delta_{h_n \mathbf{e}_n}^{\alpha_n} \Delta_{h_{n-1} \mathbf{e}_{n-1}}^{\alpha_{n-1}} \cdots \Delta_{h_1 \mathbf{e}_1}^{\alpha_1} f.$$

$\Delta_{\mathbf{h}}^\alpha$ is independent of the order of the difference quotients $\Delta_{h_i \mathbf{e}_i}^{\alpha_i}$ chosen. We also have that $\Delta_{\mathbf{h}}^\alpha (\Delta_{\mathbf{h}}^\beta f) = \Delta_{\mathbf{h}}^\alpha (\Delta_{\mathbf{h}}^\beta f)$ for all multi-indices α and β in $\mathbb{N}^n$. If $f \in H^k(\mathbb{R}^n)$, it is useful to note that $D^\alpha (\Delta_{\mathbf{h}}^\beta f) = \Delta_{\mathbf{h}}^\beta (D^\alpha f)$ for all multi-indices α and β in $\mathbb{N}^n$ with $|\alpha| \leq k$, where the derivative is in the sense of distribution. The proofs of these statements is left to the reader.

The **adjoint** of $\Delta_{\mathbf{h}}^\alpha$ is $(-1)^{|\alpha|} \Delta_{-\mathbf{h}}^\alpha$. The corollary to the next result justifies this definition. The next result will occasionally be useful later.

Proposition 12.6.2

Suppose that $k \in \mathbb{N}$ and that $f, g \in L^2(\Omega)$, where Ω is an open subset of $\mathbb{R}^n$. Moreover, suppose that there exists δ such that $\text{supp } f + B_{k\delta}(\mathbf{0}) \subset \Omega$ and $\text{supp } g + B_{k\delta}(\mathbf{0}) \subset \Omega$, then

$$\int_{\Omega} (\Delta_{\mathbf{h}}^\alpha f) g d\mathbf{x} = (-1)^{|\alpha|} \int_{\Omega} f (\Delta_{-\mathbf{h}}^\alpha g) d\mathbf{x}$$

for all multi-indices α such that $|\alpha| \leq k$, and all $\mathbf{h} \in \mathbb{R}^n$ such that $\|\mathbf{h}\| < \delta$.

Proof.

i) We first consider $\alpha = \alpha_j \mathbf{e}_j$ with $\alpha_j \leq k$ and $\|\mathbf{y}\| < \delta$. We have that

$$\begin{aligned}
\int_{\Omega} (\Delta_{\mathbf{h}}^\alpha f(\mathbf{x})) g(\mathbf{x}) d\mathbf{x} &= \int_{\Omega} (\Delta_{h_j \mathbf{e}_j}^{\alpha_j} f(\mathbf{x})) g(\mathbf{x}) d\mathbf{x} \\
&= \int_{\Omega} \frac{1}{h_j^{\alpha_j}} \left(\sum_{i=0}^{\alpha_j} (-i)^i \binom{\alpha_j}{i} f(\mathbf{x} + (\alpha_j - i) h_j \mathbf{e}_j) \right) g(\mathbf{x}) d\mathbf{x}
\end{aligned}$$

$$= \frac{1}{h_j^{\alpha_j}} \sum_{i=0}^{\alpha_j} (-i)^i \binom{\alpha_j}{i} \int_{\Omega} f(\mathbf{x} + (\alpha_j - i)h_j \mathbf{e}_j) g(\mathbf{x}) d\mathbf{x} .$$

Using the substitution $\mathbf{y} = \mathbf{x} + (\alpha_j - i)h_j \mathbf{e}_j$, we get

$$\begin{aligned} \int_{\Omega} (\Delta_{\mathbf{h}}^{\alpha} f(\mathbf{x})) g(\mathbf{x}) d\mathbf{x} &= \frac{1}{h_j^{\alpha_j}} \sum_{i=0}^{\alpha_j} (-i)^i \binom{\alpha_j}{i} \int_{\Omega} f(\mathbf{y}) g(\mathbf{y} - (\alpha_j - i)h_j \mathbf{e}_j) d\mathbf{y} \\ &= (-1)^{\alpha_j} \int_{\Omega} f(\mathbf{y}) \left(\frac{1}{(-h_j)^{\alpha_j}} \sum_{i=0}^{\alpha_j} (-i)^i \binom{\alpha_j}{i} g(\mathbf{y} - (\alpha_j - i)h_j \mathbf{e}_j) \right) d\mathbf{y} \\ &= (-1)^{\alpha_j} \int_{\Omega} f(\mathbf{y}) (\Delta_{-h_j \mathbf{e}_j}^{\alpha_j} g(\mathbf{y})) d\mathbf{y} . \end{aligned}$$

Note that the previous change of variable is justified by the fact that $\text{supp } f - (\alpha_j - i)h_j \mathbf{e}_j \subset \Omega$ and $\text{supp } g + (\alpha_j - i)h_j \mathbf{e}_j \subset \Omega$ for all i .

ii) For the general case, we proceed recursively using (i).

$$\begin{aligned} \int_{\Omega} (\Delta_{\mathbf{h}}^{\alpha} f(\mathbf{x})) g(\mathbf{x}) d\mathbf{x} &= \int_{\Omega} (\Delta_{h_n \mathbf{e}_n}^{\alpha_n} \Delta_{h_{n-1} \mathbf{e}_{n-1}}^{\alpha_{n-1}} \cdots \Delta_{h_1 \mathbf{e}_1}^{\alpha_1} f(\mathbf{x})) g(\mathbf{x}) d\mathbf{x} \\ &= (-1)^{\alpha_n} \int_{\Omega} (\Delta_{h_{n-1} \mathbf{e}_{n-1}}^{\alpha_{n-1}} \cdots \Delta_{h_1 \mathbf{e}_1}^{\alpha_1} f(\mathbf{x})) (\Delta_{-h_n \mathbf{e}_n}^{\alpha_n} g(\mathbf{x})) d\mathbf{x} \\ &\quad \dots \\ &= (-1)^{\alpha_n + \alpha_{n-1} + \cdots + \alpha_1} \int_{\Omega} f(\mathbf{x}) (\Delta_{-h_1 \mathbf{e}_1}^{\alpha_1} \cdots \Delta_{-h_{n-1} \mathbf{e}_{n-1}}^{\alpha_{n-1}} \Delta_{-h_n \mathbf{e}_n}^{\alpha_n} g(\mathbf{x})) d\mathbf{x} \\ &= (-1)^{|\alpha|} \int_{\Omega} f(\mathbf{x}) (\Delta_{-\mathbf{h}}^{\alpha} g(\mathbf{x})) d\mathbf{x} . \quad \blacksquare \end{aligned}$$

Corollary 12.6.3

Suppose that $f, g \in L^2(\mathbb{R}^n)$. Then

$$\int_{\mathbb{R}^n} (\Delta_{\mathbf{h}}^{\alpha} f) g d\mathbf{x} = (-1)^{|\alpha|} \int_{\mathbb{R}^n} f (\Delta_{-\mathbf{h}}^{\alpha} g) d\mathbf{x}$$

for all multi-indices α and all $\mathbf{h} \in \mathbb{R}^n$.

Proof.

The conditions $\text{supp } f + B_{k\delta}(\mathbf{0}) \subset \mathbb{R}^n$ and $\text{supp } g + B_{k\delta}(\mathbf{0}) \subset \mathbb{R}^n$ of the previous proposition are satisfied for all $\delta > 0$ and $k \in \mathbb{N}$. ■

Proposition 12.6.4

Given $f \in H^{s,2}(\mathbb{R}^n)$ and $\alpha \in \mathbb{N}^n$, we have that

$$\|D^{\alpha} f\|_{s,\rho} = \limsup_{\mathbf{h} \rightarrow \mathbf{0}} \|\Delta_{\mathbf{h}}^{\alpha} f\|_{s,\rho} , \quad (12.6.5)$$

where the possibility of both sides to be infinite is not ruled out. It follows that

$D^\alpha f \in H^{s,2}(\mathbb{R}^n)$ if and only if $\limsup_{\mathbf{h} \rightarrow \mathbf{0}} \|\Delta_{\mathbf{h}}^\alpha f\|_{s,\rho}$ is finite.

Moreover, $\|\Delta_{\mathbf{h}}^\alpha f\|_{s,\rho} \leq \|D^\alpha f\|_{s,\rho}$ for all $f \in H^{s,2}(\mathbb{R}^n)$, $\mathbf{h} \in \mathbb{R}^n$ and $\alpha \in \mathbb{N}^n$ with $h_j \neq 0$ for $\alpha_j \neq 0$.

Proof.

We first recall a definition and a result that will be quite useful in this proof. As we have defined in Section 6.5, $f_{\mathbf{r}}(\mathbf{x}) = f(\mathbf{x} - \mathbf{r})$ for all $\mathbf{x}, \mathbf{r} \in \mathbb{R}^n$. From Proposition 6.6.3, we have that $(f_{\mathbf{r}})^\wedge(\mathbf{y}) = e^{-i\mathbf{r} \cdot \mathbf{y}} \hat{f}(\mathbf{y})$.

We prove by induction that

$$(\Delta_{\mathbf{h}}^\alpha f)^\wedge(\mathbf{y}) = \prod_{j=1}^m \left(\frac{(2i)^{\alpha_j}}{h_j^{\alpha_j}} e^{ih_j y_j \alpha_j / 2} \sin^{\alpha_j} \left(\frac{h_j y_j}{2} \right) \right) \hat{f}(\mathbf{y}) \quad (12.6.6)$$

for $\mathbf{h} = (h_1, \dots, h_m, 0, \dots, 0)$ and $\alpha = (\alpha_1, \dots, \alpha_m, 0, \dots, 0)$.

For $\mathbf{h} = (h_1, 0, \dots, 0)$ and $\alpha = (\alpha_1, 0, \dots, 0)$, we have that

$$\begin{aligned} (\Delta_{\mathbf{h}}^\alpha f)^\wedge(\mathbf{y}) &= (\Delta_{h_1 \mathbf{e}_1}^{\alpha_1} f)^\wedge(\mathbf{y}) = \frac{1}{h_1^{\alpha_1}} \sum_{j_1=0}^{\alpha_1} (-1)^{j_1} \binom{\alpha_1}{j_1} (f_{-(\alpha_1-j_1)h_1 \mathbf{e}_1})^\wedge(\mathbf{y}) \\ &= \frac{1}{h_1^{\alpha_1}} \sum_{j_1=0}^{\alpha_1} (-1)^{j_1} \binom{\alpha_1}{j_1} e^{i(\alpha_1-j_1)h_1 y_1} \hat{f}(\mathbf{y}) = \frac{1}{h_1^{\alpha_1}} (e^{ih_1 y_1} - 1)^{\alpha_1} \hat{f}(\mathbf{y}) \\ &= \frac{(2i)^{\alpha_1}}{h_1^{\alpha_1}} e^{ih_1 y_1 \alpha_1 / 2} \sin^{\alpha_1} \left(\frac{h_1 y_1}{2} \right) \hat{f}(\mathbf{y}). \end{aligned}$$

So (12.6.6) is true for $m = 1$.

Suppose that (12.6.6) is true for $m < n$. If $\mathbf{h} = (h_1, \dots, h_{m+1}, 0, \dots, 0)$ and $\alpha = (\alpha_1, \dots, \alpha_{m+1}, 0, \dots, 0)$, let $\tilde{\mathbf{h}} = (h_1, \dots, h_m, 0, \dots, 0)$ and $\tilde{\alpha} = (\alpha_1, \dots, \alpha_m, 0, \dots, 0)$. Then,

$$\begin{aligned} (\Delta_{\mathbf{h}}^\alpha f)^\wedge(\mathbf{y}) &= (\Delta_{\tilde{\mathbf{h}}}^{\tilde{\alpha}} (\Delta_{h_{m+1} \mathbf{e}_{m+1}}^{\alpha_{m+1}} f))^\wedge(\mathbf{y}) = \prod_{j=1}^m \left(\frac{(2i)^{\alpha_j}}{h_j^{\alpha_j}} e^{ih_j y_j \alpha_j / 2} \sin^{\alpha_j} \left(\frac{h_j y_j}{2} \right) \right) (\Delta_{h_{m+1} \mathbf{e}_{m+1}}^{\alpha_{m+1}} f)^\wedge(\mathbf{y}) \\ &= \prod_{j=1}^m \left(\frac{(2i)^{\alpha_j}}{h_j^{\alpha_j}} e^{ih_j y_j \alpha_j / 2} \sin^{\alpha_j} \left(\frac{h_j y_j}{2} \right) \right) \left(\frac{(2i)^{\alpha_{m+1}}}{h_{m+1}^{\alpha_{m+1}}} e^{ih_{m+1} y_{m+1} \alpha_{m+1} / 2} \sin^{\alpha_{m+1}} \left(\frac{h_{m+1} y_{m+1}}{2} \right) \hat{f}(\mathbf{y}) \right) \\ &= \prod_{j=1}^{m+1} \left(\frac{(2i)^{\alpha_j}}{h_j^{\alpha_j}} e^{ih_j y_j \alpha_j / 2} \sin^{\alpha_j} \left(\frac{h_j y_j}{2} \right) \right) \hat{f}(\mathbf{y}), \end{aligned}$$

where we have used the hypothesis of induction for the second equality. This is (12.6.6) with m replaced by $m + 1$. Thus (12.6.6) is true for all $m \leq n$.

Since $\left| \frac{2}{h_j} \sin \left(\frac{h_j y_j}{2} \right) \right| \leq |y_j|$ for all j , we get from (12.6.6) that

$$|(\Delta_{\mathbf{h}}^\alpha f)^\wedge(\mathbf{y})| \leq \prod_{j=1}^n |y_j|^{\alpha_j} |\hat{f}(\mathbf{y})| = \left| \prod_{j=1}^n y_j^{\alpha_j} \hat{f}(\mathbf{y}) \right| = |\mathbf{y}^\alpha \hat{f}(\mathbf{y})| \quad (12.6.7)$$

for all $\mathbf{y} \in \mathbb{R}^n$. Moreover, since $\lim_{h_j \rightarrow 0} \frac{2}{h_j} \sin\left(\frac{h_j y_j}{2}\right) = y_j$ for all j , we get from (12.6.6) that

$$\lim_{\mathbf{h} \rightarrow \mathbf{0}} (\Delta_{\mathbf{h}}^{\alpha} f)^{\wedge}(\mathbf{y}) = i^{|\alpha|} \prod_{j=1}^n y_j^{\alpha_j} \hat{f}(\mathbf{y}) = (i\mathbf{y})^{\alpha} \hat{f}(\mathbf{y}) \quad (12.6.8)$$

for all $\mathbf{y} \in \mathbb{R}^n$.

i) If $D^{\alpha} f \in H^{s,2}(\mathbb{R}^n)$, then it follows from (12.6.7) that

$$\begin{aligned} (1 + \|\mathbf{y}\|_2^2)^s |(\Delta_{\mathbf{h}}^{\alpha} f)^{\wedge}(\mathbf{y})|^2 &\leq (1 + \|\mathbf{y}\|_2^2)^s |\mathbf{y}^{\alpha} \hat{f}(\mathbf{y})|^2 \\ &= (1 + \|\mathbf{y}\|_2^2)^s |(D^{\alpha} f)^{\wedge}(\mathbf{y})|^2 \in L(\mathbb{R}^n) \end{aligned} \quad (12.6.9)$$

for all $\mathbf{y} \in \mathbb{R}^n$, and from (12.6.8) that

$$\lim_{\mathbf{h} \rightarrow \mathbf{0}} (1 + \|\mathbf{y}\|_2^2)^s |(\Delta_{\mathbf{h}}^{\alpha} f)^{\wedge}(\mathbf{y})|^2 = (1 + \|\mathbf{y}\|_2^2)^s |(i\mathbf{y})^{\alpha} \hat{f}(\mathbf{y})|^2 = (1 + \|\mathbf{y}\|_2^2)^s |(D^{\alpha} f)^{\wedge}(\mathbf{y})|^2$$

for all $\mathbf{y} \in \mathbb{R}^n$. Hence, we get from the Lebesgue Dominate Convergence Theorem that

$$\lim_{\mathbf{h} \rightarrow \mathbf{0}} \|\Delta_{\mathbf{h}}^{\alpha} f\|_{s,\rho} = \|D^{\alpha} f\|_{s,\rho}.$$

ii) Let us suppose that $\limsup_{\mathbf{h} \rightarrow \mathbf{0}} \|\Delta_{\mathbf{h}}^{\alpha} f\|_{s,\rho}$ is finite but $\|D^{\alpha} f\|_{s,\rho} = \infty$; namely,

$$\int_{\mathbb{R}^n} |\mathbf{y}^{\alpha} \hat{f}(\mathbf{y})|^2 (1 + \|\mathbf{y}\|_2^2)^s d\mathbf{y} = \infty \quad (\text{see item (1) in Remark 12.6.5 below}).$$

Therefore, given $M > 0$, there exists $R > 0$ such that

$$\int_{\|\mathbf{y}\| \leq R} |\mathbf{y}^{\alpha} \hat{f}(\mathbf{y})|^2 (1 + \|\mathbf{y}\|_2^2)^s d\mathbf{y} > M.$$

Since $\lim_{p \rightarrow 0} \left| \frac{\sin(p)}{p} \right| = 1$, there exists $P > 0$ such that $\left| \frac{\sin(p)}{p} \right| \geq \frac{1}{2}$ for $|p| < P$. Therefore, $\left| \frac{2}{h_j y_j} \sin\left(\frac{h_j y_j}{2}\right) \right| \geq \frac{1}{2}$ for $\left| \frac{h_j y_j}{2} \right| < P$. It follows that $\left| \frac{2}{h_j y_j} \sin\left(\frac{h_j y_j}{2}\right) \right| \geq \frac{1}{2}$ for all y_j such that $|y_j| \leq R$ if $|h_j| < 2P/R$.

Therefore, we have for $\|\mathbf{h}\|_2 < 2P/R$ that

$$\begin{aligned} \|\Delta_{\mathbf{h}}^{\alpha} f\|_{s,\rho}^2 &= \int_{\mathbb{R}^n} \prod_{j=1}^n \left| \frac{(2i)^{\alpha_j}}{h_j^{\alpha_j}} e^{ih_j y_j \alpha_j / 2} \sin^{\alpha_j} \left(\frac{h_j y_j}{2} \right) \right|^2 |\hat{f}(\mathbf{y})|^2 (1 + \|\mathbf{y}\|_2^2)^s d\mathbf{y} \\ &= \int_{\mathbb{R}^n} \prod_{j=1}^n \left| \frac{2^{\alpha_j}}{h_j^{\alpha_j}} \sin^{\alpha_j} \left(\frac{h_j y_j}{2} \right) \right|^2 |\hat{f}(\mathbf{y})|^2 (1 + \|\mathbf{y}\|_2^2)^s d\mathbf{y} \\ &\geq \int_{\|\mathbf{y}\| \leq R} \prod_{j=1}^n \left| \frac{2^{\alpha_j}}{h_j^{\alpha_j}} \sin^{\alpha_j} \left(\frac{h_j y_j}{2} \right) \right|^2 |\hat{f}(\mathbf{y})|^2 (1 + \|\mathbf{y}\|_2^2)^s d\mathbf{y} \\ &\geq \int_{\|\mathbf{y}\| \leq R} \prod_{j=1}^n \left(\frac{|y_j|^{\alpha_j}}{2^{\alpha_j}} \right)^2 |\hat{f}(\mathbf{y})|^2 (1 + \|\mathbf{y}\|_2^2)^s d\mathbf{y} \end{aligned}$$

$$= \frac{1}{2^{2|\alpha|}} \int_{\|\mathbf{y}\| \leq R} |\mathbf{y}^\alpha \hat{f}(\mathbf{y})|^2 (1 + \|\mathbf{y}\|_2^2)^s d\mathbf{y} > \frac{M}{2^{2|\alpha|}}.$$

Since this is true for all $M > 0$, we get that $\limsup_{\mathbf{h} \rightarrow \mathbf{0}} \|\Delta_{\mathbf{h}}^\alpha f\|_{s,\rho} = \infty$. This contradicts our hypothesis, therefore we must have that $\int_{\mathbb{R}^n} |\mathbf{y}^\alpha \hat{f}(\mathbf{y})|^2 (1 + \|\mathbf{y}\|_2^2)^s d\mathbf{y} < \infty$ and so $D^\alpha f \in H^{s,2}(\mathbb{R}^n)$ (see item (2) in Remark 12.6.5 below).

iii) The last conclusion of the proposition follows from (12.6.9). ■

Remark 12.6.5

1. If $f \in H^{s,2}(\mathbb{R}^n)$, then f is a temperate distribution. It follows from Proposition 6.6.12 that $(D^\alpha f)^\wedge = (i\mathbf{y})^\alpha \hat{f}$ in the sense of distribution. So $(D^\alpha f)^\wedge$ is also represented by a measurable function. However, there is no guarantee that $(1 + \|\mathbf{y}\|_2^2)^{s/2} (D^\alpha f)^\wedge$ is in $L^2(\mathbb{R}^n)$; namely, that $\|D^\alpha f\|_{s,\rho}$ is finite.
2. If $\int_{\mathbb{R}^n} |\mathbf{y}^\alpha \hat{f}(\mathbf{y})|^2 (1 + \|\mathbf{y}\|_2^2)^s d\mathbf{y} < \infty$ and $s \geq 0$, then $(D^\alpha f)^\wedge = (i\mathbf{y})^\alpha \hat{f} \in L^2(\mathbb{R}^n)$ in the sense of distributions. It follows from Plancherel theorem, Theorem 6.6.6, that there exists $g \in L^2(\mathbb{R}^n)$ such that $(D^\alpha f)^\wedge = \hat{g}$ in the sense of distributions. As we have shown in the proof of Theorem 11.1.3, this implies that $D^\alpha f = g \in L^2(\mathbb{R}^n)$ in the sense of distributions.
3. From the proof of the previous proposition, we have that $\limsup_{\mathbf{h} \rightarrow \mathbf{0}} \|\Delta_{\mathbf{h}}^\alpha f\|_{s,\rho} < \infty$ implies that $D^\alpha f \in H^{s,2}(\mathbb{R}^n)$, and that we have in fact $\lim_{\mathbf{h} \rightarrow \mathbf{0}} \|\Delta_{\mathbf{h}}^\alpha f\|_{s,\rho} = \|D^\alpha f\|_{s,\rho}$. ♠

Given $\phi \in \mathcal{S}(\mathbb{R}^n)$, we define the linear operator $[\Delta_{\mathbf{h}}^\alpha, \phi]$ on $H^{s,2}(\mathbb{R}^n)$ as

$$[\Delta_{\mathbf{h}}^\alpha, \phi]f = \Delta_{\mathbf{h}}^\alpha(\phi f) - \phi \Delta_{\mathbf{h}}^\alpha f$$

for $f \in H^{s,2}(\mathbb{R}^n)$.

We will also refer below to the following order relation between multi-indices. Two multi-indices α and β satisfy $\alpha \leq \beta$ if $\alpha_j \leq \beta_j$ for all j , and $\alpha < \beta$ if $\alpha_j \leq \beta_j$ for all j and there exists at least one j such that $\alpha_j < \beta_j$.

Proposition 12.6.6

Suppose that $\phi \in \mathcal{S}(\mathbb{R}^n)$, $\alpha \in \mathbb{N}^n$, $\mathbf{h} \in \mathbb{R}^n$ and $s \in \mathbb{R}$. Then,

$$\begin{aligned} [\Delta_{\mathbf{h}}^\alpha, \phi] : H^{s,2}(\mathbb{R}^n) &\rightarrow H^{s-|\alpha|+1,2}(\mathbb{R}^n) \\ f &\mapsto \Delta_{\mathbf{h}}^\alpha(\phi f) - \phi \Delta_{\mathbf{h}}^\alpha f \end{aligned}$$

is a bounded operator and the bound is independent of $\mathbf{h}$.

Proof.

The case $|\alpha| = 0$ is trivial.

i) We consider the case $|\alpha| = 1$. We have that

$$\begin{aligned} [\Delta_{h_j e_j}, \phi]f &= \Delta_{h_j e_j}(\phi f) - \phi(\Delta_{h_j e_j} f) = \frac{1}{h_j} (\phi_{-h_j e_j} f_{-h_j e_j} - \phi f) - \frac{1}{h_j} \phi(f_{-h_j e_j} - f) \\ &= \frac{1}{h_j} (\phi_{-h_j e_j} f_{-h_j e_j} - \phi f_{-h_j e_j}) = (\Delta_{h_j e_j} \phi) f_{-h_j e_j}. \end{aligned}$$

Moreover, $f \rightarrow f_{-h_j e_j}$ is an isometry from $H^{s,2}(\mathbb{R}^n)$ onto itself. We leave the proof of this claim as an exercise for the reader. Recall that $(f_{-h_j e_j})^\wedge(\mathbf{y}) = e^{ih_j y_j} \hat{f}(\mathbf{y})$.

Since $\Delta_{h_j e_j} \phi \in \mathcal{S}(\mathbb{R}^n)$, we have from Proposition 11.1.22 that

$$\|(\Delta_{h_j e_j} \phi) f_{-h_j e_j}\|_{s,\rho} \leq C_{j,h_j} \|f_{-h_j e_j}\|_{s,\rho} = C_{j,h_j} \|f\|_{s,\rho}$$

for a constant C_{j,h_j} . We claim that the constant C_{j,h_j} can be taken to be independent of j and h_j . Recall that, in the proof of Proposition 11.1.22, C_{j,h_j} satisfies

$$\sup_{\mathbf{z} \in \mathbb{R}^n} \int_{\mathbb{R}^n} |K_{j,h_j}(\mathbf{y}, \mathbf{z})| d\rho(\mathbf{y}) < C_{j,h_j} \quad \text{and} \quad \sup_{\mathbf{y} \in \mathbb{R}^n} \int_{\mathbb{R}^n} |K_{j,h_j}(\mathbf{y}, \mathbf{z})| d\rho(\mathbf{z}) < C_{j,h_j}, \quad (12.6.10)$$

where

$$K_{j,h_j}(\mathbf{y}, \mathbf{z}) = (2\pi)^{-n/2} (1 + \|\mathbf{z}\|^2)^{-s} (\Delta_{h_j e_j} \phi)^\wedge(\mathbf{y} - \mathbf{z}).$$

As we have shown in the proof of Proposition 12.6.4,

$$\begin{aligned} |(\Delta_{h_j e_j} \phi)^\wedge(\mathbf{z})| &= \left| \frac{1}{h_j} ((\phi_{-h_j e_j})^\wedge(\mathbf{z}) - \hat{\phi}(\mathbf{z})) \right| = \left| \frac{1}{h_j} (e^{ih_j z_j} \hat{\phi}(\mathbf{z}) - \hat{\phi}(\mathbf{z})) \right| \\ &= \left| \frac{1}{h_j} (e^{ih_j z_j} - 1) \right| |\hat{\phi}(\mathbf{z})| = \left| \frac{2i}{h_j} e^{ih_j z_j/2} \sin\left(\frac{h_j z_j}{2}\right) \right| |\hat{\phi}(\mathbf{z})| \leq |z_j| |\hat{\phi}(\mathbf{z})| \leq |\mathbf{z}| |\hat{\phi}(\mathbf{z})| \end{aligned}$$

for all $\mathbf{z} \in \mathbb{R}^n$, $h_j \in \mathbb{R}$ and j . Following the steps of the proof of Proposition 11.1.22, we have that

$$\begin{aligned} \int_{\mathbb{R}^n} |K_{j,h_j}(\mathbf{y}, \mathbf{z})| d\rho(\mathbf{y}) &= (2\pi)^{-n/2} \int_{\mathbb{R}^n} (1 + \|\mathbf{z}\|^2)^{-s} (1 + \|\mathbf{y}\|^2)^s |(\Delta_{h_j e_j} \phi)^\wedge(\mathbf{y} - \mathbf{z})| d\mathbf{y} \\ &\leq 2^s (2\pi)^{-n/2} \int_{\mathbb{R}^n} (1 + \|\mathbf{y} - \mathbf{z}\|^2)^s |(\Delta_{h_j e_j} \phi)^\wedge(\mathbf{y} - \mathbf{z})| d\mathbf{y} \\ &\leq 2^s (2\pi)^{-n/2} \int_{\mathbb{R}^n} (1 + \|\mathbf{y} - \mathbf{z}\|^2)^s |\mathbf{y} - \mathbf{z}| |\hat{\phi}(\mathbf{y} - \mathbf{z})| d\mathbf{y} \\ &\leq 2^s (2\pi)^{-n/2} \int_{\mathbb{R}^n} (1 + \|\mathbf{w}\|^2)^s |\mathbf{w}| |\hat{\phi}(\mathbf{w})| d\mathbf{w} \equiv R_{s,1} < \infty \end{aligned}$$

for all $\mathbf{z} \in \mathbb{R}^n$, $h_j \in \mathbb{R}$ and j because $\hat{\phi} \in \mathcal{S}(\mathbb{R}^n)$ by Proposition 6.6.11. Likewise,

$$\begin{aligned} \int_{\mathbb{R}^n} |K_{j,h_j}(\mathbf{y}, \mathbf{z})| d\rho(\mathbf{z}) &= (2\pi)^{-n/2} \int_{\mathbb{R}^n} |(\Delta_{h_j e_j} \phi)^\wedge(\mathbf{y} - \mathbf{z})| d\mathbf{z} \\ &\leq (2\pi)^{-n/2} \int_{\mathbb{R}^n} |\mathbf{y} - \mathbf{z}| |\hat{\phi}(\mathbf{y} - \mathbf{z})| d\mathbf{z} \leq (2\pi)^{-n/2} \int_{\mathbb{R}^n} |\mathbf{w}| |\hat{\phi}(\mathbf{w})| d\mathbf{w} \equiv R_{s,2} < \infty \end{aligned}$$

for all $\mathbf{y} \in \mathbb{R}^n$, $h_j \in \mathbb{R}$ and j because $\hat{\phi} \in \mathcal{S}(\mathbb{R}^n)$ by Proposition 6.6.11.

We have that (12.6.10) is satisfied for all h_j and j with C_{j,h_j} replaced by $R_s = \max\{R_{s,1}, R_{s,2}\}$. Hence

$$\|[\Delta_{h_j e_j}, \phi] f\|_{s,\rho} \leq R_s \|f\|_{s,\rho} \quad (12.6.11)$$

for all $h_j \in \mathbb{R}$ and j . This proves the proposition for $|\alpha| = 1$.

ii) We now consider the case where $|\alpha| > 1$. We have that $[\Delta_{\mathbf{h}}^{\alpha}, \phi]f$ is a finite sum of terms of the form $\Delta_{\mathbf{h}}^{\mu} [\Delta_{h_j e_j}, \phi] \Delta_{\mathbf{h}}^{\eta} f$, where $|\mu| + 1 + |\eta| = |\alpha|$; to be precise, $\eta + \mu + e_j = \alpha$ and $1 \leq j \leq n$. For instance,

$$\begin{aligned} [\Delta_{h_1 e_1 + h_2 e_2}^2, \phi] f &= [\Delta_{h_2 e_2} \Delta_{h_1 e_1}, \phi] f = \Delta_{h_2 e_2} \Delta_{h_1 e_1} (\phi f) - \phi \Delta_{h_2 e_2} \Delta_{h_1 e_1} f \\ &= \Delta_{h_2 e_2} \Delta_{h_1 e_1} (\phi f) - \Delta_{h_2 e_2} (\phi \Delta_{h_1 e_1} f) + \Delta_{h_2 e_2} (\phi \Delta_{h_1 e_1} f) - \phi \Delta_{h_2 e_2} \Delta_{h_1 e_1} f \\ &= \Delta_{h_2 e_2} [\Delta_{h_1 e_1}, \phi] f + [\Delta_{h_2 e_2}, \phi] (\Delta_{h_1 e_1} f) . \end{aligned} \quad (12.6.12)$$

This in fact proves the claim for $|\alpha| = 2$. We leave it to the reader to prove by induction on $|\alpha|$ that the claim is true for all α . Note that (12.6.12) can be generalized to

$$[\Delta_{\mathbf{h}}^{\beta} \Delta_{\mathbf{h}}^{\nu}, \phi] f = \Delta_{\mathbf{h}}^{\beta} [\Delta_{\mathbf{h}}^{\nu}, \phi] f + [\Delta_{\mathbf{h}}^{\beta}, \phi] (\Delta_{\mathbf{h}}^{\nu} f)$$

for any multi-indices β and ν .

We also have from Propositions 11.1.5 and 12.6.4 that

$$\|\Delta_{\mathbf{h}}^{\beta} g\|_{t,\rho} \leq \|D^{\beta} g\|_{t,\rho} \leq \|g\|_{t+|\beta|,\rho} \quad (12.6.13)$$

for all $t \in \mathbb{R}$, $g \in H^{t,2}(\mathbb{R}^n)$, $\mathbf{h} \in \mathbb{R}^n$ and $\beta \in \mathbb{N}^n$ with $h_j \neq 0$ for $\beta_j \neq 0$. Hence,

$$\begin{aligned} \|\Delta_{\mathbf{h}}^{\mu} [\Delta_{h_j e_j}, \phi] \Delta_{\mathbf{h}}^{\eta} f\|_{s-|\alpha|+1,\rho} &\leq \|[\Delta_{h_j e_j}, \phi] \Delta_{\mathbf{h}}^{\eta} f\|_{s-|\alpha|+|\mu|+1,\rho} \leq R_{s-|\alpha|+|\mu|+1} \|\Delta_{\mathbf{h}}^{\eta} f\|_{s-|\alpha|+|\mu|+1,\rho} \\ &\leq R_{s-|\alpha|+|\mu|+1} \|f\|_{s-|\alpha|+|\mu|+|\eta|+1,\rho} = R_{s-|\alpha|+|\mu|+1} \|f\|_{s,\rho} . \end{aligned}$$

We finally get that $\|[\Delta_{\mathbf{h}}^{\alpha}, \phi] f\|_{s-|\alpha|+1,\rho} \leq C \|f\|_{s,\rho}$, where $C = \sum_{|\mu|+|\nu|+1=|\alpha|} R_{s-|\alpha|+|\mu|+1}$. ■

Corollary 12.6.7

Suppose that $L(\mathbf{x}, D)$ is the elliptic operator defined in (12.6.1) with $a_{\alpha} \in \mathcal{S}(\mathbb{R}^n)$ for all α . Moreover, suppose that $\beta \in \mathbb{N}^n$, $\mathbf{h} \in \mathbb{R}^n$ and $s \in \mathbb{R}$. Then the operator

$$[\Delta_{\mathbf{h}}^{\beta}, L(\mathbf{x}, D)] = \Delta_{\mathbf{h}}^{\beta} \circ L(\mathbf{x}, D) - L(\mathbf{x}, D) \circ \Delta_{\mathbf{h}}^{\beta}$$

is a bounded operator from $H^{s,2}(\mathbb{R}^n)$ into $H^{s-k-|\beta|+1,2}(\mathbb{R}^n)$ with the bound independent of h .

Proof.

Since D^α commutes with Δ_h^β , we get that

$$\begin{aligned} [\Delta_h^\beta, L]f(x) &= \Delta_h^\beta \left(\sum_{|\alpha| \leq k} a_\alpha(x) D^\alpha f(x) \right) - \sum_{|\alpha| \leq k} a_\alpha(x) D^\alpha (\Delta_h^\beta f(x)) \\ &= \sum_{|\alpha| \leq k} (\Delta_h^\beta (a_\alpha(x) D^\alpha f(x)) - a_\alpha(x) \Delta_h^\beta (D^\alpha f(x))) = \sum_{|\alpha| \leq k} [\Delta_h^\beta, a_\alpha] D^\alpha f(x) \end{aligned}$$

for all $f \in H^{s,2}(\mathbb{R}^n)$. It follows from Propositions 11.1.5 and 12.6.6 that

$$\begin{aligned} \|[\Delta_h^\beta, L]f\|_{s-k-|\beta|+1,\rho} &\leq \sum_{|\alpha| \leq k} \|[\Delta_h^\beta, a_\alpha] D^\alpha f\|_{s-k-|\beta|+1,\rho} \leq \sum_{|\alpha| \leq k} C_{a_\alpha, \beta} \|D^\alpha f\|_{s-k,\rho} \\ &\leq \sum_{|\alpha| \leq k} C_{a_\alpha, \beta} \|f\|_{s-k+|\alpha|,\rho} \leq \left(\sum_{|\alpha| \leq k} C_{a_\alpha, \beta} \right) \|f\|_{s,\rho}, \end{aligned}$$

where the constants $C_{a_\alpha, \beta}$ are independent of h . ■

Using Theorem 11.6.3, we may prove the following corollary to Proposition 12.6.4. Recall that $\mathbb{R}_+^n = \{(x_1, x_2, \dots, x_n) \in \mathbb{R}^n : x_n > 0\}$ and $\mathbb{R}_-^n = \{(x_1, x_2, \dots, x_n) \in \mathbb{R}^n : x_n < 0\}$.

Corollary 12.6.8

Let $\Omega_r = \mathbb{R}_+^n \cap B_r(\mathbf{0})$ for some $r > 0$. If the support of $f \in H^{k,2}(\mathbb{R}^n)$ is a subset of $\Omega_{r,0} = (\mathbb{R}^n \setminus \mathbb{R}_-^n) \cap B_r(\mathbf{0})$ (Figure 12.1) and $\alpha \in \mathbb{N}^n$ is a multi-index with $\alpha_n = 0$, then $D^\alpha f \in H^{k,2}(\Omega_r)$ if and only if

$$A_{f,\alpha} \equiv \limsup_{h \rightarrow 0} \|\Delta_h^\alpha f\|_{k,\Omega_r} < \infty.$$

In particular, if $D^\alpha f \in H^{k,2}(\Omega_r)$, then there exists a constant C , independent of f and r , such that $\|D^\alpha f\|_{k,\Omega_r} \leq C A_{f,\alpha}$.

Proof.

According to Theorem 11.6.3, there exists a strong k -extension operator E on $\mathbb{R}_+^n$ and a strong $(k - |\alpha|)$ -extension operator E_α on $\mathbb{R}_+^n$ such that $D^\alpha(E(g)) = E_\alpha(D^\alpha g)$ for all $g \in H^{k,2}(\mathbb{R}_+^n)$. Since $\alpha_n = 0$, a quick look at the proof of Theorem 11.6.3 shows that $E_\alpha = E$. Moreover, we also see that $E(\Delta_h^\alpha g) = \Delta_h^\alpha E(g)$ for all $g \in H^{k,2}(\mathbb{R}_+^n)$ because $\alpha_n = 0$.

We have that $E : H^{k,2}(\mathbb{R}_+^n) \rightarrow H^{k,2}(\mathbb{R}^n)$ is a bounded operator and $E(g) = g$ on $\mathbb{R}_+^n$ for all $g \in H^{k,2}(\mathbb{R}_+^n)$. So, there exists a positive constant C_1 such that $\|E(g)\|_{k,2,\mathbb{R}^n} \leq C_1 \|g\|_{k,2,\mathbb{R}_+^n}$ for all $g \in H^{k,2}(\mathbb{R}_+^n)$.

Since the norms $\|\cdot\|_{k,2,\mathbb{R}^n}$ and $\|\cdot\|_{k,\rho}$ are equivalent, there exist positive constants C_2 and C_3 such that

$$C_3 \|g\|_{k,2,\mathbb{R}^n} \leq \|g\|_{k,\rho} \leq C_2 \|g\|_{k,2,\mathbb{R}^n}$$

for all $g \in H^{k,2}(\mathbb{R}^n)$.

i) Suppose that $D^\alpha f \in H^{k,2}(\Omega_r)$. Since $\text{supp } f \subset \Omega_{r,0}$, we have that $\|D^\alpha f\|_{k,2,\mathbb{R}_+^n} = \|D^\alpha f\|_{k,2,\Omega_r} < \infty$. Hence,

$$\|D^\alpha aE(f)\|_{k,\rho} = \|E(D^\alpha f)\|_{k,\rho} \leq C_2 \|E(D^\alpha f)\|_{k,2,\mathbb{R}^n} \leq C_1 C_2 \|D^\alpha f\|_{k,2,\mathbb{R}_+^n} < \infty .$$

Thus, according to Proposition 12.6.4,

$$\limsup_{\mathbf{h} \rightarrow \mathbf{0}} \|\Delta_{\mathbf{h}}^\alpha E(f)\|_{k,\rho} = \|D^\alpha E(f)\|_{k,\rho} < \infty .$$

It follows that

$$\begin{aligned} A_{f,\alpha} &\equiv \limsup_{\mathbf{h} \rightarrow \mathbf{0}} \|\Delta_{\mathbf{h}}^\alpha f\|_{k,2,\Omega_r} = \limsup_{\mathbf{h} \rightarrow \mathbf{0}} \|E(\Delta_{\mathbf{h}}^\alpha f)\|_{k,2,\Omega_r} \leq \limsup_{\mathbf{h} \rightarrow \mathbf{0}} \|E(\Delta_{\mathbf{h}}^\alpha f)\|_{k,2,\mathbb{R}^n} \\ &\leq C_3^{-1} \limsup_{\mathbf{h} \rightarrow \mathbf{0}} \|E(\Delta_{\mathbf{h}}^\alpha f)\|_{k,\rho} < \infty . \end{aligned}$$

ii) Suppose $A_{f,\alpha} < \infty$. Since $\text{supp } f$ is a compact subset of the bounded set $\Omega_{r,0}$, there exists $\delta > 0$ such that $|\alpha|\mathbf{h} + \text{supp } f \subset \Omega_{r,0}$ for $\|\mathbf{h}\| < \delta$ with $h_n = 0$. Thus, $\text{supp } \Delta_{\mathbf{h}}^\alpha f \subset \Omega_{r,0}$ for $\|\mathbf{h}\| < \delta$ without constraints on h_n since we assume that $\alpha_n = 0$. Hence,

$$\begin{aligned} \limsup_{\mathbf{h} \rightarrow \mathbf{0}} \|\Delta_{\mathbf{h}}^\alpha E(f)\|_{k,\rho} &= \limsup_{\mathbf{h} \rightarrow \mathbf{0}} \|E(\Delta_{\mathbf{h}}^\alpha f)\|_{k,\rho} \leq C_2 \limsup_{\mathbf{h} \rightarrow \mathbf{0}} \|E(\Delta_{\mathbf{h}}^\alpha f)\|_{k,2,\mathbb{R}^n} \\ &\leq C_1 C_2 \limsup_{\mathbf{h} \rightarrow \mathbf{0}} \|\Delta_{\mathbf{h}}^\alpha f\|_{k,2,\mathbb{R}_+^n} = C_1 C_2 \limsup_{\mathbf{h} \rightarrow \mathbf{0}} \|\Delta_{\mathbf{h}}^\alpha f\|_{k,2,\Omega_r} = C_1 C_2 A_{f,\alpha} < \infty , \end{aligned} \quad (12.6.14)$$

where we have used the fact that $\text{supp } \Delta_{\mathbf{h}}^\alpha f \subset \Omega_{r,0}$ for $\|\mathbf{h}\| < \delta$ for the second to last equality in the previous expression. It follows from Proposition 12.6.4 that $D^\alpha E(f) \in H^{k,2}(\mathbb{R}^n)$. Since $\Omega_r \subset \mathbb{R}^n$, we have that $D^\alpha E(f) \in H^{k,2}(\Omega_r)$. Since $E(f) = f$ on $\mathbb{R}_+^n \supset \Omega_r$, we have that $D^\alpha f \in H^{k,2}(\Omega_r)$.

iii) It also follows from Proposition 12.6.4 that $\|D^\alpha E(f)\|_{k,\rho} = \limsup_{\mathbf{h} \rightarrow \mathbf{0}} \|\Delta_{\mathbf{h}}^\alpha E(f)\|_{k,\rho}$. Hence, we get from (12.6.14) that

$$\begin{aligned} \|D^\alpha f\|_{k,2,\Omega_r} &= \|D^\alpha E(f)\|_{k,2,\mathbb{R}_+^n} \leq \|D^\alpha E(f)\|_{k,2,\mathbb{R}^n} \leq C_3^{-1} \|D^\alpha E(f)\|_{k,\rho} \\ &= C_3^{-1} \limsup_{\mathbf{h} \rightarrow \mathbf{0}} \|\Delta_{\mathbf{h}}^\alpha E(f)\|_{k,\rho} \leq C_1 C_2 C_3^{-1} A_{f,\alpha} = C A_{f,\alpha} , \end{aligned}$$

where $C = C_1 C_2 C_3^{-1}$. Note that C is independent of f and r because C_1 depends only on E , and C_2 and C_3 depend only on the equivalence between the norms $\|\cdot\|_{k,2,\mathbb{R}^n}$ and $\|\cdot\|_{k,\rho}$. ■

The corollary to the next proposition will be needed to study the regularity of solutions at the boundary in Section 12.6.2.

Proposition 12.6.9

Suppose that $\phi \in \mathcal{S}(\mathbb{R}^n)$, $\alpha \in \mathbb{N}^n$, $\mathbf{h} \in \mathbb{R}^n$ and $s \in \mathbb{R}$. Then, there exists a positive

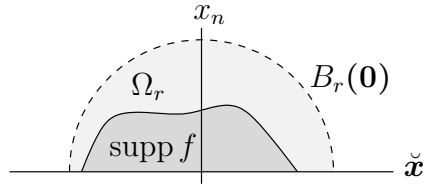


Figure 12.1: Support of f in Corollary 12.6.8. We have set $\check{\mathbf{x}} = (x_1, x_2, \dots, x_{n-1})$.

constant C independent of $\mathbf{h}$ such that

$$\|[\Delta_{\mathbf{h}}^{\alpha}, \phi]f\|_{s,\rho} \leq C \sum_{\beta < \alpha} \|D^{\beta} f\|_{s,\rho} \quad (12.6.15)$$

for all $f \in H^{s,2}(\mathbb{R}^n)$. The possibility that the right hand side be ∞ is not ruled out.

Proof.

The case $|\alpha| = 0$ is trivial. The case $|\alpha| = 1$ follows word for word from (i) in the proof of Proposition 12.6.6. Namely, we get that

$$\|[\Delta_{h_j e_j}, \phi]f\|_{s,2,\rho} \leq R_{\phi,s} \|f\|_{s,2,\rho}$$

for all $h_j \in \mathbb{R}$ and j , where $R_{\phi,s}$ is a positive constant that depends on $s \in \mathbb{R}$ and $\phi \in \mathcal{D}(\mathbb{R}^n)$ only. We get (12.6.15) for $|\alpha| = 1$ if $C = R_{\phi,s}$.

We now consider the case $|\alpha| > 1$. As was mentioned in (ii) of the prove of Proposition 12.6.6, we have that $[\Delta_{\mathbf{h}}^{\alpha}, \phi]f$ is a finite sum of terms of the form $\Delta_{\mathbf{h}}^{\eta} [\Delta_{h_j e_j}, \phi] \Delta_{\mathbf{h}}^{\mu} f$, where $\eta + \mu + e_j = \alpha$ and $1 \leq j \leq n$.

We have from Proposition 12.6.4 that $\|\Delta_{\mathbf{h}}^{\beta} g\|_{t,\rho} \leq \|D^{\beta} g\|_{t,\rho}$ for all $t \in \mathbb{R}$, $g \in H^{t,2}(\mathbb{R}^n)$, $\mathbf{h} \in \mathbb{R}^n$ and $\beta \in \mathbb{N}^n$ with $h_j \neq 0$ for $\beta_j \neq 0$.

We also have that

$$\begin{aligned} D^{\eta} [\Delta_{\mathbf{y}}^{\beta}, \phi]g &= D^{\eta} (\Delta_{\mathbf{y}}^{\beta}(\phi g) - \phi \Delta_{\mathbf{y}}^{\beta} g) = \Delta_{\mathbf{y}}^{\beta} D^{\eta}(\phi g) - D^{\eta}(\phi \Delta_{\mathbf{y}}^{\beta} g) \\ &= \Delta_{\mathbf{y}}^{\beta} \left(\sum_{\zeta + \xi = \eta} \binom{\eta}{\xi} D^{\zeta} \phi D^{\xi} g \right) - \left(\sum_{\zeta + \xi = \eta} \binom{\eta}{\xi} D^{\zeta} \phi D^{\xi} (\Delta_{\mathbf{y}}^{\beta} g) \right) \\ &= \sum_{\zeta + \xi = \eta} \binom{\eta}{\xi} (\Delta_{\mathbf{y}}^{\beta} (D^{\zeta} \phi D^{\xi} g) - D^{\zeta} \phi (\Delta_{\mathbf{y}}^{\beta} D^{\xi} g)) = \sum_{\zeta + \xi = \eta} \binom{\eta}{\xi} [\Delta_{\mathbf{y}}^{\beta}, D^{\zeta} \phi] D^{\xi} g \end{aligned}$$

for all multi-indices β and η in $\mathbb{N}^n$, and $g \in H^{|\eta|}(\mathbb{R}^n)$. The previous equality is in the sense of distributions.

The inequality in (12.6.15) is obviously true if $D^{\beta} f \notin H^{s,2}(\mathbb{R}^n)$ for some β . The right hand side of (12.6.15) is then ∞ . So, we may assume that $D^{\beta} f \in H^{s,2}(\mathbb{R}^n)$ for all $\beta < \alpha$.

Hence,

$$\begin{aligned}
\|\Delta_h^\eta [\Delta_{h_j e_j}, \phi] \Delta_h^\mu f\|_{s,\rho} &\leq \|D^\eta [\Delta_{h_j e_j}, \phi] \Delta_h^\mu f\|_{s,\rho} \\
&\leq \sum_{\xi+\zeta=\eta} \binom{\eta}{\xi} \|\Delta_{h_j e_j}, D^\zeta \phi\|_{s,\rho} \|D^\xi (\Delta_h^\mu f)\|_{s,\rho} \leq \sum_{\xi+\zeta=\eta} \binom{\eta}{\xi} R_{D^\zeta \phi, s} \|D^\xi (\Delta_h^\mu f)\|_{s,\rho} \\
&= \sum_{\xi+\zeta=\eta} \binom{\eta}{\xi} R_{D^\zeta \phi, s} \|\Delta_h^\mu (D^\xi f)\|_{s,\rho} \leq \sum_{\xi+\zeta=\eta} \binom{\eta}{\xi} R_{D^\zeta \phi, s} \|(D^{\xi+\mu} f)\|_{s,\rho},
\end{aligned}$$

where the last inequality comes from Proposition 12.6.4. We then get that

$$\begin{aligned}
\|[\Delta_{h_j e_j}, \psi] f\|_{s,\rho} &\leq \sum_{j=1}^n \sum_{\eta+\mu+e_j \leq \alpha} \|\Delta_h^\eta [\Delta_{h_j e_j}, \phi] \Delta_h^\mu f\|_{s,\rho} \\
&\leq \sum_{j=1}^n \sum_{\eta+\mu+e_j \leq \alpha} \sum_{\xi+\zeta=\eta} \binom{\eta}{\xi} R_{D^\zeta \phi, s} \|(D^{\xi+\mu} f)\|_{s,\rho} \\
&= \sum_{j=1}^n \sum_{\substack{\mu \leq \alpha - e_j \\ \text{if } \alpha_j > 0}} \sum_{\xi \leq \eta = \alpha - e_j - \mu} \binom{\eta}{\xi} R_{D^{\eta-\xi} \phi, s} \|(D^{\xi+\mu} f)\|_{s,\rho}.
\end{aligned}$$

We have that $\xi + \mu \leq \alpha - e_j < \alpha$. Thus, we get (12.6.15) if C is the maximum of $\binom{\eta}{\xi} R_{D^{\eta-\xi} \phi, s}$ for $\mu \leq \alpha - e_j$ if $\alpha_j > 0$ and $\xi \leq \eta = \alpha - e_j - \mu$. ■

Corollary 12.6.10

Let $\Omega_r = \mathbb{R}_+^n \cap B_r(\mathbf{0})$ for some $r > 0$. Suppose that $\phi \in \mathcal{S}(\mathbb{R}^n)$ and that α is a multi-index with $k = |\alpha|$ and $\alpha_n = 0$. There exists a constant C such that, for all $0 < \delta < 1$,

$$\|[\Delta_h^\alpha, \phi] f\|_{k,2,\Omega_r} \leq C \sum_{\beta < \alpha} \|D^\beta f\|_{k,2,\Omega_r} \quad (12.6.16)$$

for $f \in H^{k,2}(\Omega_r)$ with $\text{supp } f \subset \Omega_{\delta r,0} = (\mathbb{R}^n \setminus \mathbb{R}_+^n) \cap B_{\delta r}(\mathbf{0})$ and $\|h\| < (1 - \delta)r$.

Proof.

According to Theorem 11.6.3, there exists a strong k -extension operator E on $\mathbb{R}_+^n$ and a strong $(k - |\beta|)$ -extension operator E_β on $\mathbb{R}_+^n$ such that $D^\beta(E(g)) = E_\beta(D^\beta g)$ for all $g \in H^{k,2}(\mathbb{R}_+^n)$, where $|\beta| \leq k$. Since $\alpha_n = 0$, a quick look at the proof of Theorem 11.6.3 shows that $E_\beta = E$. Moreover, we also see that $E(\Delta_h^\alpha f) = \Delta_h^\alpha E(f)$ because $\alpha_n = 0$.

Since E is a strong k -extension operator E on $\mathbb{R}_+^n$, we have that $E : H^{k,2}(\mathbb{R}_+^n) \rightarrow H^{k,2}(\mathbb{R}^n)$ is a bounded operator and $E(g) = g$ on $\mathbb{R}_+^n$ for all $g \in H^{k,2}(\mathbb{R}_+^n)$. So, there exists a positive constant C_1 such that $\|E(g)\|_{k,2,\mathbb{R}^n} \leq C_1 \|g\|_{k,2,\mathbb{R}_+^n}$ for all $g \in H^{k,2}(\mathbb{R}_+^n)$.

As we have done several times in previous proof, we use the equivalence of the norms $\|\cdot\|_{k,2,\mathbb{R}^n}$ and $\|\cdot\|_{k,\rho}$. Suppose that

$$C_3 \|g\|_{k,2,\mathbb{R}^n} \leq \|g\|_{k,\rho} \leq C_2 \|g\|_{k,2,\mathbb{R}^n}$$

for all $g \in H^{k,2}(\mathbb{R}^n)$ and some positive constants C_2 and C_3 .

From Proposition 12.6.9, there exists a positive constant C_0 independent of $\mathbf{h}$ such that

$$\|[\Delta_{\mathbf{h}}^{\alpha}, \phi]g\|_{k,\rho} \leq C_0 \sum_{\beta < \alpha} \|D^{\beta} g\|_{k,\rho}$$

for all $g \in H^{k,2}(\mathbb{R}^n)$.

The inequality in (12.6.16) is obviously true if $D^{\beta} f \notin H^{k,2}(\Omega_r)$ for some β . The right hand side of (12.6.16) is then ∞ . So, we may assume that $D^{\beta} f \in H^{k,2}(\Omega_r)$ for all $\beta < \alpha$. Since $\text{supp } f$ is a compact subset of the bounded set $\Omega_{\delta r,0}$, we have that

$$\|E(D^{\beta} f)\|_{k,\rho} \leq C_2 \|E(D^{\beta} f)\|_{k,2,\mathbb{R}^n} \leq C_1 C_2 \|D^{\beta} f\|_{k,2,\mathbb{R}_+^n} = C_1 C_2 \|D^{\beta} f\|_{k,2,\Omega_r} < \infty$$

for all $\beta < \alpha$. Thus, $E(D^{\beta} f) \in H^{k,2}(\mathbb{R}^n)$ for all $\beta < \alpha$.

Moreover, since $\text{supp } f$ is a compact subset of the bounded set $\Omega_{\delta r,0}$, we have that $|\alpha|\mathbf{h} + \text{supp } f \subset \Omega_{r,0}$ for $\|\mathbf{h}\| < (1-\delta)r$ with $h_n = 0$. Thus, $\text{supp } \Delta_{\mathbf{h}}^{\alpha} f \subset \Omega_{r,0}$ for $\|\mathbf{h}\| < (1-\delta)r$ without constraints on h_n since we assume that $\alpha_n = 0$. Hence, $\text{supp}[\Delta_{\mathbf{h}}^{\alpha}, \phi]f \subset \Omega_{r,0}$.

Hence, for $\|\mathbf{h}\| < (1-\delta)r$, we have that

$$\begin{aligned} \|[\Delta_{\mathbf{h}}^{\alpha}, \phi]f\|_{k,2,\Omega_r} &= \|[\Delta_{\mathbf{h}}^{\alpha}, \phi]E(f)\|_{k,2,\mathbb{R}_+^n} \leq \|[\Delta_{\mathbf{h}}^{\alpha}, \phi]E(f)\|_{k,2,\mathbb{R}^n} \leq C_3^{-1} \|[\Delta_{\mathbf{h}}^{\alpha}, \phi]E(f)\|_{k,\rho} \\ &\leq C_0 C_3^{-1} \sum_{\beta < \alpha} \|D^{\beta} E(f)\|_{k,\rho} \leq C_0 C_2 C_3^{-1} \sum_{\beta < \alpha} \|D^{\beta} E(f)\|_{k,2,\mathbb{R}^n} = C_0 C_2 C_3^{-1} \sum_{\beta < \alpha} \|E(D^{\beta} f)\|_{k,2,\mathbb{R}^n} \\ &\leq C_0 C_1 C_2 C_3^{-1} \sum_{\beta < \alpha} \|D^{\beta} f\|_{k,2,\mathbb{R}_+^n} = C_0 C_1 C_2 C_3^{-1} \sum_{\beta < \alpha} \|D^{\beta} f\|_{k,2,\Omega_r}. \end{aligned}$$

This gives (12.6.16) with $C = C_0 C_1 C_2 C_3^{-1}$. ■

Remark 12.6.11

The proof of Corollary 12.6.10 can be slightly modified to show that there exists a constant C such that, for all $0 < \delta < 1$,

$$\|[\Delta_{\mathbf{h}}^{\alpha}, \phi]f\|_{k,\Omega_r} \leq C \|f\|_{k+|\alpha|-1,\Omega_r}$$

for $f \in H^{k+|\alpha|-1,2}(\Omega_r)$ with $\text{supp } f \subset \Omega_{\delta r,0} = (\mathbb{R}^n \setminus \mathbb{R}_-^n) \cap B_{\delta r}(\mathbf{0})$ and $\|\mathbf{h}\| < (1-\delta)r$. Note that this conclusion is not as strong as the conclusion of Corollary 12.6.10 because $\beta < \alpha$ is more restrictive than $|\beta| < |\alpha|$ as we have in the definition of $\|\cdot\|_{k+|\alpha|-1,2,\Omega_r}$.

According to Theorem 11.6.5, there exists a total extension operator E on $\mathbb{R}_+^n$. From a quick look at the definition of E given in the sketch of the proof of Theorem 11.6.5, we see that $E(\Delta_{\mathbf{h}}^{\alpha} f) = \Delta_{\mathbf{h}}^{\alpha} E(f)$ because $\alpha_n = 0$.

Since E is a total extension operator E on $\mathbb{R}_+^n$, we have for every $q \in \mathbb{N}$ that $E : H^{q,2}(\mathbb{R}_+^n) \rightarrow H^{q,2}(\mathbb{R}^n)$ is a bounded operator and $E(g) = g$ on $\mathbb{R}_+^n$ for all $g \in H^{q,2}(\mathbb{R}_+^n)$. It can be shown that there exists a positive constant C_1 independent of q such that $\|E(g)\|_{q,2,\mathbb{R}^n} \leq C_1 \|g\|_{q,2,\mathbb{R}_+^n}$ for all $g \in H^{q,2}(\mathbb{R}_+^n)$.

To prove our claim, we will need the equivalence of the norms $\|\cdot\|_{q,2,\mathbb{R}^n}$ and $\|\cdot\|_{q,\rho}$ for $q = k$ and $q = k + |\alpha| - 1$. Namely,

$$C_{q,3} \|g\|_{q,2,\mathbb{R}^n} \leq \|g\|_{k,\rho} \leq C_{q,2} \|g\|_{q,2,\mathbb{R}^n}$$

for all $g \in H^{q,2}(\mathbb{R}^n)$ and some positive constants $C_{q,2}$ and $C_{q,3}$.

From Proposition 12.6.6, there exists a positive constant C_0 independent of $\mathbf{h}$ such that

$$\|[\Delta_{\mathbf{h}}^{\alpha}, \phi]g\|_{k,\rho} \leq C_0 \|g\|_{k+|\alpha|-1,\rho}$$

for all $g \in H^{k+|\alpha|-1,2}(\mathbb{R}^n)$.

As we stated in the proof of Corollary 12.6.10, we have that $|\alpha|\mathbf{h} + \text{supp } f \subset \Omega_{r,0}$ for $\|\mathbf{h}\| < (1-\delta)r$ with $h_n = 0$. Thus, $\text{supp } \Delta_{\mathbf{h}}^{\alpha} f \subset \Omega_{r,0}$ for $\|\mathbf{h}\| < \delta$ without constraints on h_n since we assume that $\alpha_n = 0$. Hence, $\text{supp}[\Delta_{\mathbf{h}}^{\alpha}, \phi]f \subset \Omega_{r,0}$ and

$$\begin{aligned} \|[\Delta_{\mathbf{h}}^{\alpha}, \phi]f\|_{k,2,\Omega_r} &\leq \|[\Delta_{\mathbf{h}}^{\alpha}, \phi]E(f)\|_{k,2,\mathbb{R}^n} \leq C_{k,3}^{-1} \|[\Delta_{\mathbf{h}}^{\alpha}, \phi]E(f)\|_{k,\rho} \leq C_0 C_{k,3}^{-1} \|E(f)\|_{k+|\alpha|-1,\rho} \\ &\leq C_0 C_1 C_{k+|\alpha|-1,3}^{-1} \|f\|_{k+|\alpha|-1,2,\mathbb{R}_+^n} = C_0 C_1 C_{k+|\alpha|-1,3}^{-1} \|f\|_{k+|\alpha|-1,2,\Omega_r} = C \|f\|_{k+|\alpha|-1,2,\Omega_r}, \end{aligned}$$

where $C = C_0 C_1 C_{k+|\alpha|-1,3}^{-1}$. ♠

12.6.1 Local Regularity

This subject is also called **internal regularity** by some authors. However, local regularity seems more appropriate because the results in this section address the regularity of weak solutions of elliptic equations in local part of the domain Ω of the equations.

To proof the main result of this section, we need a few preliminary results.

Proposition 12.6.12

Suppose that $\Omega \subset \mathbb{R}^n$ is a bounded open set and that $L(\mathbf{x}, D)$ defined in (12.6.1) is elliptic in a open neighbourhood of $\bar{\Omega}$. For every $s \in \mathbb{R}$, there exists a constant $C > 0$ such that

$$\|\underline{u}\|_{s,\rho} \leq C \left(\|L(\mathbf{x}, D)\underline{u}\|_{s-k,\rho} + \|\underline{u}\|_{s-1,\rho} \right) \quad (12.6.17)$$

for $u \in H_0^{s,2}(\Omega)$, where as usual

$$\underline{u}(\mathbf{x}) = \begin{cases} u(\mathbf{x}) & \text{if } \mathbf{x} \in \Omega \\ 0 & \text{if } \mathbf{x} \in \mathbb{R}^n \setminus \Omega \end{cases}$$

In particular, if $s \in \mathbb{N}$, there exists a constant $\tilde{C}$ such that

$$\|u\|_{s,2,\Omega} \leq \tilde{C} \left(\|L(\mathbf{x}, D)u\|_{s-k,2,\Omega} + \|u\|_{s-1,2,\Omega} \right) \quad (12.6.18)$$

for $u \in H_0^{s,2}(\Omega)$.

Proof.

i) We first consider the elliptic operator $L_1(\mathbf{x}, D) = \sum_{|\alpha|=k} a_\alpha D^\alpha$, where we assume that the coefficients a_α are constant. We have that

$$(L_1(\mathbf{x}, D)u)^\wedge(\mathbf{y}) = (-i)^k \sum_{|\alpha|=k} a_\alpha \mathbf{y}^\alpha \hat{u}(\mathbf{y})$$

for all $u \in H^{s,2}(\mathbb{R}^n)$. Since $L_1(\mathbf{x}, D)$ is elliptic, there exists a constant $C_0 > 0$ such that

$$\left| \sum_{|\alpha|=k} a_\alpha \mathbf{y}^\alpha \right| \geq C_0 \|\mathbf{y}\|^k$$

for all $\mathbf{y} \in \mathbb{R}^n$. thus,

$$\left| (L_1(\mathbf{x}, D)u)^\wedge(\mathbf{y}) \right| = \left| \sum_{|\alpha|=k} a_\alpha \mathbf{y}^\alpha \right| |\hat{u}(\mathbf{y})| \geq C_0 \|\mathbf{y}\|^k |\hat{u}(\mathbf{y})|$$

and so

$$\begin{aligned} (1 + \|\mathbf{y}\|_2^2)^s |\hat{u}(\mathbf{y})|^2 &\leq (1 + \|\mathbf{y}\|_2^2)^{s-k} (1 + \|\mathbf{y}\|_2^2)^k |\hat{u}(\mathbf{y})|^2 \leq 2^k (1 + \|\mathbf{y}\|_2^2)^{s-k} (1 + \|\mathbf{y}\|_2^{2k}) |\hat{u}(\mathbf{y})|^2 \\ &\leq 2^k (1 + \|\mathbf{y}\|_2^2)^{s-k} |\hat{u}(\mathbf{y})|^2 + \frac{2^k}{C_0^2} (1 + \|\mathbf{y}\|_2^2)^{s-k} |(L_1(\mathbf{x}, D)u)^\wedge(\mathbf{y})|^2. \end{aligned}$$

We get that

$$\begin{aligned} \|u\|_{s,\rho}^2 &= \int_{\mathbb{R}^n} (1 + \|\mathbf{y}\|_2^2)^s |\hat{u}(\mathbf{y})|^2 d\mathbf{y} \\ &\leq 2^k \int_{\mathbb{R}^n} (1 + \|\mathbf{y}\|_2^2)^{s-k} |\hat{u}(\mathbf{y})|^2 d\mathbf{y} + \frac{2^k}{C_0^2} \int_{\mathbb{R}^n} (1 + \|\mathbf{y}\|_2^2)^{s-k} |(L_1(\mathbf{x}, D)u)^\wedge(\mathbf{y})|^2 d\mathbf{y} \\ &= 2^k \|u\|_{s-k,\rho}^2 + \frac{2^k}{C_0^2} \|L_1(\mathbf{x}, D)u\|_{s-k,\rho}^2 \leq 2^k \|u\|_{s-1,\rho}^2 + \frac{2^k}{C_0^2} \|L_1(\mathbf{x}, D)u\|_{s-k,\rho}^2 \\ &\leq C_1^2 \left(\|u\|_{s-1,\rho} + \|L_1(\mathbf{x}, D)u\|_{s-k,\rho} \right)^2, \end{aligned}$$

where $C_1^2 = 2^k \max\{1, 1/C_0^2\}$. We get

$$\|u\|_{s,\rho} \leq C_1 \left(\|u\|_{s-1,\rho} + \|L_1(\mathbf{x}, D)u\|_{s-k,\rho} \right). \quad (12.6.19)$$

ii) We now consider the elliptic operator $L_2(\mathbf{x}, D) = \sum_{|\alpha|=k} a_\alpha(\mathbf{x}) D^\alpha$, where we assume that $a_\alpha \in \mathcal{D}(\mathbb{R}^n)$. We may make this assumption because we assume that the a_α are of class C^∞ in an open neighbourhood U of $\overline{\Omega}$. Therefore, we may choose $\phi \in \mathcal{D}(\mathbb{R}^n)$ such that $\text{supp } \phi \subset U$ and $\phi(\mathbf{x}) = 1$ for all $\mathbf{x} \in \overline{\Omega}$, and set

$$\tilde{a}_\alpha(\mathbf{x}) = \begin{cases} \phi(\mathbf{x}) a_\alpha(\mathbf{x}) & \text{if } \mathbf{x} \in U \\ 0 & \text{if } \mathbf{x} \in \mathbb{R}^n \setminus U \end{cases}$$

to get functions $\tilde{a}_\alpha \in \mathcal{D}(\mathbb{R}^n)$ such that $\tilde{a}_\alpha = a_\alpha$ on Ω . Therefore, the elliptic partial differential equation on Ω is not changed if we replace the coefficients a_α by the coefficients $\tilde{a}_\alpha$. This is what we do. From now on, we drop the tilde to simplify the notation.

Since ∇a_α is continuous and has a compact support, we have that

$$K = \sup_{\mathbf{x} \in \mathbb{R}^n} \|\nabla a_\alpha(\mathbf{x})\|_2 < \infty .$$

Therefore, from the Mean Value Theorem, we get that

$$|a_\alpha(\mathbf{x}_1) - a_\alpha(\mathbf{x}_2)| \leq K \|\mathbf{x}_1 - \mathbf{x}_2\|_2 \quad (12.6.20)$$

for all $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^n$. Since there is a finite number of a_α , we may assume that (12.6.20) is true for all α with $|\alpha| = k$ and all $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^n$.

Given $\mathbf{x}_0 \in \overline{\Omega}$ and $\delta > 0$, choose $\psi \in \mathcal{D}(\mathbb{R}^n)$ such that $\text{supp } \psi \subset B_{2\delta}(\mathbf{0})$, $|\psi(\mathbf{x})| \leq 1$ for all $\mathbf{x}$, and $\psi(\mathbf{x}) = 1$ for all $\mathbf{x} \in B_\delta(\mathbf{0})$. Suppose that $u \in H^{s,2}(\mathbb{R}^n)$ is such that $\text{supp } u \subset B_\delta(\mathbf{x}_0)$. Then it follows from Lemma 12.6.1 that there exists a constant C_3 such that

$$\begin{aligned} \|(a_\alpha - a_\alpha(\mathbf{x}_0)) D^\alpha u\|_{s-k,\rho} &= \|(a_\alpha - a_\alpha(\mathbf{x}_0)) \psi_{\mathbf{x}_0} D^\alpha u\|_{s-k,\rho} \\ &\leq (2\pi)^{n/2} \sup_{\mathbf{x} \in \mathbb{R}^n} |(a_\alpha(\mathbf{x}) - a_\alpha(\mathbf{x}_0)) \psi_{\mathbf{x}_0}| \|D^\alpha u\|_{s-k,\rho} \\ &\quad + C_3 \|(a_\alpha - a_\alpha(\mathbf{x}_0)) \psi_{\mathbf{x}_0}\|_{|s-k-1|+1+t,\rho} \|D^\alpha u\|_{s-k-1,\rho} \\ &\leq C_2 \delta \|D^\alpha u\|_{s-k,\rho} + C_4 \|D^\alpha u\|_{s-k-1,\rho} \leq C_2 \delta \|u\|_{s,\rho} + C_4 \|u\|_{s-1,\rho} , \end{aligned}$$

where $t > n/2$ is arbitrary but fixed, $C_2 = 2(2\pi)^{n/2}K$ and $C_4 \geq C_3 \max_{\mathbf{x}_0 \in \overline{\Omega}} \|(a_\alpha - a_\alpha(\mathbf{x}_0)) \psi_{\mathbf{x}_0}\|_{|s-k-1|+1+t,\rho}$. Such a maximum exists because $\overline{\Omega}$ is compact and $\mathbf{x}_0 \mapsto (a_\alpha - a_\alpha(\mathbf{x}_0)) \psi_{\mathbf{x}_0}$ is a continuous mapping from $\overline{\Omega}$ into $H^{|s-k-1|+1+t}(\mathbb{R}^n)$. Recall that $\psi_{\mathbf{x}_0}(\mathbf{x}) = \psi(\mathbf{x} - \mathbf{x}_0)$ for all $\mathbf{x} \in \mathbb{R}^n$. Hence, C_4 is independent of $\mathbf{x}_0 \in \overline{\Omega}$.

We then have that

$$\|L_2(\mathbf{x}, D)u - L_2(\mathbf{x}_0, D)u\|_{s-k,\rho} \leq \sum_{|\alpha|=k} \|(a_\alpha - a_\alpha(\mathbf{x}_0)) D^\alpha u\|_{s-k,\rho} \leq C_2 \delta N \|u\|_{s,\rho} + C_4 N \|u\|_{s-1,\rho} ,$$

where $N = \sum_{|\alpha|=k} 1$. From (12.6.19) with $L_1(\mathbf{x}, D)$ replaced by $L_2(\mathbf{x}_0, D)$ which has constant coefficients, we get

$$\begin{aligned} \|u\|_{s,\rho} &\leq C_1 \left(\|u\|_{s-1,\rho} + \|L_2(\mathbf{x}_0, D)u\|_{s-k,\rho} \right) \\ &\leq C_1 \left(\|u\|_{s-1,\rho} + \|L_2(\mathbf{x}_0, D)u - L_2(\mathbf{x}, D)u\|_{s-k,\rho} + \|L_2(\mathbf{x}, D)u\|_{s-k,\rho} \right) \\ &\leq C_1 \left(\|u\|_{s-1,\rho} + C_2 \delta N \|u\|_{s,\rho} + C_4 N \|u\|_{s-1,\rho} + \|L_2(\mathbf{x}, D)u\|_{s-k,\rho} \right) \\ &= C_1 (1 + C_4 N) \|u\|_{s-1,\rho} + C_1 \|L_2(\mathbf{x}, D)u\|_{s-k,\rho} + C_1 C_2 \delta N \|u\|_{s,\rho} . \end{aligned}$$

If we choose $0 < \delta < 1/(2C_1 C_2 N)$, we get

$$\|u\|_{s,\rho} \leq C_1 (1 + C_4 N) \|u\|_{s-1,\rho} + C_1 \|L_2(\mathbf{x}, D)u\|_{s-k,\rho} + \frac{1}{2} \|u\|_{s,\rho} .$$

Therefore,

$$\begin{aligned} \|u\|_{s,\rho} &\leq 2C_1(1 + C_4N)\|u\|_{s-1,\rho} + 2C_1\|L_2(\mathbf{x}, D)u\|_{s-k,\rho} \\ &\leq C_5\left(\|u\|_{s-1,\rho} + \|L_2(\mathbf{x}, D)u\|_{s-k,\rho}\right), \end{aligned} \quad (12.6.21)$$

where $C_5 = 2C_1(1 + C_4N)$.

iii) As in (ii), we still consider the elliptic operator $L_2(\mathbf{x}, D) = \sum_{|\alpha|=k} a_\alpha(\mathbf{x}) D^\alpha$, where we assume that $a_\alpha \in \mathcal{D}(\mathbb{R}^n)$. Since $\overline{\Omega}$ is compact, there exists a finite subset $\{x_i\}_{i=1}^I$ of $\overline{\Omega}$ such that $\{B_\delta(x_i)\}_{i=1}^I$ is an open cover of $\overline{\Omega}$ with δ selected in (ii). Let $\{\phi_i\}_{i=1}^I$ be a partition of unity of $\overline{\Omega}$ subordinate to $\{B_\delta(x_i)\}_{i=1}^I$.

Given $u \in H_0^{s,2}(\Omega)$, then $\underline{u} \in H^{s,2}(\mathbb{R}^n)$ according to Proposition 11.2.2 and $\text{supp } \phi_i \underline{u} \subset B_\delta(x_i)$. Therefore,

$$\begin{aligned} \|\underline{u}\|_{s,2} &\leq \sum_{i=1}^I \|\phi_i \underline{u}\|_{s,2} \leq C_5 \sum_{i=1}^I \left(\|\phi_i \underline{u}\|_{s-1,\rho} + \|L_2(\mathbf{x}, D)(\phi_i \underline{u})\|_{s-k,\rho} \right) \\ &\leq C_5 \sum_{i=1}^I \left(\|\phi_i \underline{u}\|_{s-1,\rho} + \|\phi_i L_2(\mathbf{x}, D)\underline{u}\|_{s-k,\rho} \right. \\ &\quad \left. + \|L_2(\mathbf{x}, D)(\phi_i \underline{u}) - \phi_i L_2(\mathbf{x}, D)\underline{u}\|_{s-k,\rho} \right). \end{aligned} \quad (12.6.22)$$

We have from Proposition 11.1.22 that there exists a constant $C_6 > 0$ such that

$$\|\phi_i \underline{u}\|_{s-1,\rho} \leq C_6 \|\underline{u}\|_{s-1,\rho}. \quad (12.6.23)$$

Since there are a finite number of ϕ_i , we may assume that C_6 is large enough for (12.6.23) to be true for all i . Similarly, from Proposition 11.1.22, we have that there exists a constant $C_7 > 0$ such that

$$\|\phi_i L_2(\mathbf{x}, D)\underline{u}\|_{s-k,\rho} \leq C_7 \|L_2(\mathbf{x}, D)\underline{u}\|_{s-k,\rho}. \quad (12.6.24)$$

Again, since there are a finite number of ϕ_i , we may assume that C_7 is large enough for (12.6.24) to be true for all i . Since

$$L_2(\mathbf{x}, D)(\phi_i \underline{u}) - (\phi_i L_2(\mathbf{x}, D)\underline{u}) = \sum_{|\alpha| \leq k-1} \psi_\alpha D^\alpha \underline{u},$$

where $\psi_\alpha \in \mathcal{D}(\mathbb{R}^n)$, we get from Proposition 11.1.5 and Proposition 11.1.22 that

$$\begin{aligned} \|L_2(\mathbf{x}, D)(\phi_i \underline{u}) - \phi_i L_2(\mathbf{x}, D)\underline{u}\|_{s-k,\rho} &= \left\| \sum_{|\alpha| \leq k-1} \psi_\alpha D^\alpha \underline{u} \right\|_{s-k,\rho} \leq \sum_{|\alpha| \leq k-1} \|\psi_\alpha D^\alpha \underline{u}\|_{s-k,\rho} \\ &= \sum_{|\alpha| \leq k-1} R_{\alpha, s-k} \|D^\alpha \underline{u}\|_{s-k,\rho} \leq \sum_{|\alpha| \leq k-1} R_{\alpha, s-k} \|\underline{u}\|_{s-k+|\alpha|,\rho} \leq \left(\sum_{|\alpha| \leq k-1} R_{\alpha, s-k} \right) \|\underline{u}\|_{s-1,\rho} \end{aligned} \quad (12.6.25)$$

for some constants $R_{\alpha, s-k} > 0$. As we did before, since there are a finite number of ϕ_i , we may assume that the $R_{\alpha, s-k} > 0$ are large enough for (12.6.25) to be true for all i . If $C_8 = \sum_{|\alpha| \leq k-1} R_{\alpha, s-k}$, we get that

$$\|L_2(\mathbf{x}, D)(\phi_i \underline{u}) - \phi_i L_2(\mathbf{x}, D)\underline{u}\|_{s-k,\rho} \leq C_8 \|\underline{u}\|_{s-1,\rho} \quad (12.6.26)$$

for all i . Combining (12.6.22), (12.6.23), (12.6.24) and (12.6.26), we obtain

$$\begin{aligned} \|\underline{u}\|_{s,\rho} &\leq C_5 \sum_{i=1}^I \left(C_6 \|\underline{u}\|_{s-1,\rho} + C_7 \|L_2(\mathbf{x}, D)\underline{u}\|_{s-k,\rho} + C_8 \|\underline{u}\|_{s-1,\rho} \right) \\ &\leq C_9 \left(\|\underline{u}\|_{s-1,\rho} + \|L_2(\mathbf{x}, D)\underline{u}\|_{s-k,\rho} \right), \end{aligned} \quad (12.6.27)$$

where $C_9 = IC_5 \max\{C_6 + C_8, C_7\}$.

iv) We now consider the full elliptic operator $L(\mathbf{x}, D)$ given in (12.6.17). We have that $L(\mathbf{x}, D) = L_2(\mathbf{x}, D) + L_3(\mathbf{x}, D)$, where $L_3(\mathbf{x}, D) = \sum_{|\alpha| < k} a_\alpha(\mathbf{x}) D^\alpha$. We have that

$$\|\underline{u}\|_{s,\rho} \leq C_9 \left(\|\underline{u}\|_{s-1,\rho} + \|L_2(\mathbf{x}, D)\underline{u}\|_{s-k,\rho} \right) \quad (12.6.28)$$

from (12.6.27). Proceeding as we just did in (iii) for $\|L_2(\mathbf{x}, D)(\phi_i \underline{u}) - \phi_i L_2(\mathbf{x}, D)\underline{u}\|_{s-k,\rho}$, we find that there exists a constant $C_{10} > 0$ such that

$$\|L_3(\mathbf{x}, D)(\phi_i \underline{u})\|_{s-k,\rho} \leq C_{10} \|\underline{u}\|_{s-1,\rho}$$

for all i . Thus,

$$\|L_3(\mathbf{x}, D)(\underline{u})\|_{s-k,\rho} \leq \sum_{i=1}^I \|L_3(\mathbf{x}, D)(\phi_i \underline{u})\|_{s-k,\rho} \leq C_{11} \|\underline{u}\|_{s-1,\rho}, \quad (12.6.29)$$

where $C_{11} = IC_{10}$. It follows from (12.6.28) and (12.6.29) that

$$\begin{aligned} \|\underline{u}\|_{s,\rho} &\leq C_9 \left(\|\underline{u}\|_{s-1,\rho} + \|L_2(\mathbf{x}, D)\underline{u}\|_{s-k,\rho} \right) = C_9 \left(\|\underline{u}\|_{s-1,\rho} + \|L(\mathbf{x}, D)\underline{u} - L_3(\mathbf{x}, D)\underline{u}\|_{s-k,\rho} \right) \\ &\leq C_9 \left(\|\underline{u}\|_{s-1,\rho} + \|L(\mathbf{x}, D)\underline{u}\|_{s-k,\rho} + \|L_3(\mathbf{x}, D)\underline{u}\|_{s-k,\rho} \right) \\ &\leq C_9 \left(\|\underline{u}\|_{s-1,\rho} + \|L(\mathbf{x}, D)\underline{u}\|_{s-k,\rho} + C_{11} \|\underline{u}\|_{s-1,\rho} \right) \leq C \left(\|\underline{u}\|_{s-1,\rho} + \|L(\mathbf{x}, D)\underline{u}\|_{s-k,\rho} \right) \end{aligned}$$

for $u \in H_0^{s,2}(\Omega)$, where $C = C_9 \max\{1, 1 + C_{11}\}$.

v) (12.6.18) is a consequence of (12.6.17) if we use the equivalence between the norms $\|\cdot\|_{s,\rho}$ and $\|\cdot\|_{s,2}$ for $s \in \mathbb{N}$, and the fact that $\|f\|_{k,2,\Omega} = \|\underline{f}\|_{k,2,\mathbb{R}^n}$ for all $f \in H_0^k(\Omega)$ and $k \in \mathbb{N}$. ■

Corollary 12.6.13

For every non-negative integers t and s such that $t < s - 1$, there exists a positive constant C_t such that $\|u\|_{s,2} \leq C_t (\|L(\mathbf{x}, D)u\|_{s-k,2} + \|u\|_{t,2})$ for all $u \in H_0^{s,2}(\Omega)$.

Proof.

From Lemma 12.4.2 with $\epsilon = \left(\frac{1}{2C}\right)^2$, where C is the constant in (12.6.17), there exists $K > 0$ such that

$$\|u\|_{s-1,2,\Omega} \leq \frac{1}{2C} \|u\|_{s,2,\Omega} + K \|u\|_{t,2,\Omega} \quad (12.6.30)$$

for all $u \in H_0^{s,2}(\Omega)$. To be more precise, Lemma 12.4.2 gives a constant K^2 such that

$$\|\underline{u}\|_{s-1,2,\mathbb{R}^n}^2 \leq \left(\frac{1}{2C}\right)^2 \|\underline{u}\|_{s,2,\mathbb{R}^n}^2 + K^2 \|\underline{u}\|_{t,2,\mathbb{R}^n}^2 \leq \left(\frac{1}{2C} \|\underline{u}\|_{s,2,\mathbb{R}^n} + K \|\underline{u}\|_{t,2,\mathbb{R}^n}\right)^2$$

for all $u \in H_0^{q,2}(\Omega)$, where

$$\underline{u}(\mathbf{x}) = \begin{cases} u(\mathbf{x}) & \text{if } \mathbf{x} \in \Omega \\ 0 & \text{if } \mathbf{x} \in \mathbb{R}^n \setminus \Omega \end{cases}$$

Then we use the fact that $\|f\|_{k,2,\mathbb{R}^n} = \|f\|_{k,2,\Omega}$ for all $f \in H_0^{k,2}(\Omega)$ to get (12.6.30).

If we substitute in (12.6.18), we get

$$\begin{aligned} \|u\|_{s,2} &\leq C (\|L(\mathbf{x}, D)u\|_{s-k,2} + \|u\|_{s-1,2}) \leq C \left(\|L(\mathbf{x}, D)u\|_{s-k,2} + \frac{1}{2C} \|u\|_{s,2} + K \|u\|_{t,2} \right) \\ &= \frac{1}{2} \|u\|_{s,2} + C (\|L(\mathbf{x}, D)u\|_{s-k,2} + K \|u\|_{t,2}) \end{aligned}$$

for all $u \in H_0^{q,2}(\Omega)$. Thus

$$\|u\|_{s,2} \leq 2C (\|L(\mathbf{x}, D)u\|_{s-k,2} + K \|u\|_{t,2}) \leq C_t (\|L(\mathbf{x}, D)u\|_{s-k,2} + \|u\|_{t,2})$$

for all $u \in H_0^{q,2}(\Omega)$, where $C_t = 2C \max\{1, K\}$. Do not forget that K depends on C and t . ■

Lemma 12.6.14

We consider the elliptic operator $L(\mathbf{x}, D)$ defined in (12.6.1). Suppose that $\phi \in \mathcal{D}(\mathbb{R}^n)$ and $s \in \mathbb{R}$. The operator $[L(\mathbf{x}, D), \phi]$ defined by

$$[L(\mathbf{x}, D), \phi](v) = L(\mathbf{x}, D)(\phi v) - \phi L(\mathbf{x}, D)v$$

for $v \in H^{s,2}(\mathbb{R}^n)$ is a bounded mapping from $H^{s,2}(\mathbb{R}^n)$ into $H^{s-k+1,2}(\mathbb{R}^n)$ ⁴.

Proof.

We have that

$$\begin{aligned} [L(\mathbf{x}, D), \phi]v &= L(\mathbf{x}, D)(\phi v) - \phi L(\mathbf{x}, D)v = \sum_{|\alpha| \leq k} a_\alpha(\mathbf{x}) D^\alpha(\phi v) - \phi \sum_{|\alpha| \leq k} a_\alpha(\mathbf{x}) D^\alpha v \\ &= \sum_{|\alpha| \leq k} a_\alpha(\mathbf{x}) \left(\sum_{|\nu|+|\eta|=|\alpha|} b_{\nu,\eta}(D^\nu \phi)(D^\eta v) \right) - \phi \sum_{|\alpha| \leq k} a_\alpha(\mathbf{x}) D^\alpha v \\ &= \sum_{|\eta| \leq k} \left(\sum_{|\nu| \leq k-|\eta|} a_{\nu+\eta}(\mathbf{x}) b_{\nu,\eta}(D^\nu \phi) \right) D^\eta v - \phi \sum_{|\alpha| \leq k} a_\alpha(\mathbf{x}) D^\alpha v \end{aligned}$$

⁴The reader may wonder why we did not regroup all the $[A, B]$ operators that we have introduced so far, and presented them in a beautiful unified theory. We just did not want to introduce more abstraction than we have already introduced.

$$= \sum_{|\boldsymbol{\eta}| \leq k} \left(\sum_{|\boldsymbol{\nu}| \leq k - |\boldsymbol{\eta}|} a_{\boldsymbol{\nu} + \boldsymbol{\eta}}(\mathbf{x}) b_{\boldsymbol{\nu}, \boldsymbol{\eta}}(D^{\boldsymbol{\nu}} \phi) - a_{\boldsymbol{\eta}}(\mathbf{x}) \phi \right) D^{\boldsymbol{\eta}} v = \sum_{|\boldsymbol{\eta}| < k} A_{\boldsymbol{\eta}}(\mathbf{x}) D^{\boldsymbol{\eta}} v$$

for $v \in H^{s,2}(\mathbb{R}^n)$, where the coefficients $A_{\boldsymbol{\eta}}$ are in $\mathcal{D}(\mathbb{R}^n)$. Note that $a_{\boldsymbol{\nu} + \boldsymbol{\eta}}(\mathbf{x}) b_{\boldsymbol{\nu}, \boldsymbol{\eta}}(D^{\boldsymbol{\nu}} \phi) - a_{\boldsymbol{\eta}}(\mathbf{x}) \phi = 0$ for $|\boldsymbol{\eta}| = k$ because $|\boldsymbol{\nu}| = 0$ and $b_{\boldsymbol{\nu}, \boldsymbol{\eta}} = 1$. It follows from Propositions 11.1.5 and 11.1.22 that there exist constants $C_{\boldsymbol{\eta}}$ such that

$$\|A_{\boldsymbol{\eta}}(\mathbf{x}) D^{\boldsymbol{\eta}} v\|_{s-k+1, \rho} \leq C_{\boldsymbol{\eta}} \|D^{\boldsymbol{\eta}} v\|_{s-k+1, \rho} \leq C_{\boldsymbol{\eta}} \|v\|_{s-k+|\boldsymbol{\eta}|+1, \rho} \leq C_{\boldsymbol{\eta}} \|v\|_{s, \rho} \quad (12.6.31)$$

for $v \in H^{s+1,2}(\mathbb{R}^n)$ and $|\boldsymbol{\eta}| < k$. Thus $[L(\mathbf{x}, D), \phi]$ is a bounded mapping from $H^{s,2}(\mathbb{R}^n)$ into $H^{s-k+1,2}(\mathbb{R}^n)$. ■

Lemma 12.6.15

We consider the elliptic operator $L(\mathbf{x}, D)$ defined in (12.6.1). Assume that Ω is a bounded open subsets of $\mathbb{R}^n$ and $s \in \mathbb{R}$. If $u \in H^{s,2}(\mathbb{R}^n)$ and $L(\mathbf{x}, D)u \in H_{loc}^{s-k+1,2}(\Omega)$, then $u \in H_{loc}^{s+1,2}(\Omega)$.

Proof.

According to Proposition 11.1.23, to prove that $u \in H_{loc}^{s+1,2}(\Omega)$, we only need to prove that $\phi u \in H^{s+1,2}(\mathbb{R}^n)$ for all $\phi \in \mathcal{D}(\Omega)$, where as usual $\phi \in \mathcal{D}(\Omega)$ is identified to $\phi \in \mathcal{D}(\mathbb{R}^n)$ with support in Ω . Recall that $\phi u \in \mathcal{D}'(\mathbb{R}^n)$ is defined by $(\phi u)(\psi) = u(\phi \psi)$ for all $\psi \in \mathcal{D}(\mathbb{R}^n)$.

Given $\phi \in \mathcal{D}(\Omega)$, we have from Proposition 11.1.22 that $\phi u \in H^{s,2}(\mathbb{R}^n)$ and from Proposition 11.1.23 that $\phi L(\mathbf{x}, D)u \in H^{s-k+1,2}(\mathbb{R}^n)$. Moreover, we have from Lemma 12.6.14 that $[L(\mathbf{x}, D), \phi]$ is a bounded mapping from $H^{s,2}(\mathbb{R}^n)$ into $H^{s-k+1,2}(\mathbb{R}^n)$. Thus, $[L(\mathbf{x}, D), \phi] u \in H^{s-k+1,2}(\mathbb{R}^n)$. Therefore, $L(\mathbf{x}, D)(\phi u) = \phi L(\mathbf{x}, D)u + [L(\mathbf{x}, D), \phi] u \in H^{s-k+1,2}(\mathbb{R}^n)$.

Let $K = \text{supp } \phi$. Since K is compact and Ω is open, there exists $\delta > 0$ such that $Q = \{\mathbf{x} : \sup_{\mathbf{y} \in K} \|\mathbf{x} - \mathbf{y}\| \leq \delta\}$ is a compact subset of Ω . If $\|\mathbf{h}\| \leq \delta$, then $\Delta_{h_j e_j}(\phi u) \in H^{s,2}(\Omega)$ with $\text{supp } \Delta_{h_j e_j}(\phi u) \subset Q \subset \Omega$ for all j . Hence $\Delta_{h_j e_j}(\phi u) \in H_0^{s,2}(\Omega)$ for all j according to Lemma 11.2.3. Moreover, $\Delta_{h_j e_j}(\phi u) = \Delta_{h_j e_j}(\phi u)$ because $\text{supp } \Delta_{h_j e_j}(\phi u) \subset \Omega$.

Hence, from Proposition 12.6.12, we have for $\|\mathbf{h}\| \leq \delta$ that

$$\begin{aligned} \|\Delta_{h_j e_j}(\phi u)\|_{s, \rho} &\leq C \left(\|L(\mathbf{x}, D)(\Delta_{h_j e_j}(\phi u))\|_{s-k, \rho} + \|\Delta_{h_j e_j}(\phi u)\|_{s-1, \rho} \right) \\ &= C \left(\|\Delta_{h_j e_j}(L(\mathbf{x}, D)(\phi u)) - [\Delta_{h_j e_j}, L(\mathbf{x}, D)](\phi u)\|_{s-k, \rho} + \|\Delta_{h_j e_j}(\phi u)\|_{s-1, \rho} \right) \\ &\leq C \left(\|\Delta_{h_j e_j}(L(\mathbf{x}, D)(\phi u))\|_{s-k, \rho} + \|[\Delta_{h_j e_j}, L(\mathbf{x}, D)](\phi u)\|_{s-k, \rho} + \|\Delta_{h_j e_j}(\phi u)\|_{s-1, \rho} \right) \\ &\leq C \left(\|\Delta_{h_j e_j}(L(\mathbf{x}, D)(\phi u))\|_{s-k, \rho} + C_1 \|\phi u\|_{s, \rho} + \|\Delta_{h_j e_j}(\phi u)\|_{s-1, \rho} \right) \\ &\leq C \left(\|L(\mathbf{x}, D)(\phi u)\|_{s-k+1, \rho} + C_1 \|\phi u\|_{s, \rho} + \|\phi u\|_{s-1, \rho} \right), \end{aligned}$$

where we have used Corollary 12.6.7 to obtain the second to last inequality with the constant C_1 , and Proposition 11.1.5 and Proposition 12.6.4 to obtain the last inequality.

We have therefore proved that $\limsup_{h \rightarrow 0} \|\Delta_{h_j e_j}(\phi u)\|_{s, \rho} < \infty$ for $1 \leq j \leq n$. It follows from Proposition 12.6.4 that $D_{x_j}(\phi u) \in H^{s,2}_{loc}(\mathbb{R}^n)$ for $1 \leq j \leq n$. Finally, Proposition 11.1.5 implies that $\phi u \in H^{s+1,2}(\mathbb{R}^n)$. Since $\phi \in \mathcal{D}(\Omega)$ was arbitrary, we have that $u \in H^{s+1,2}_{loc}(\Omega)$. ■

The mean result of this section is the following theorem.

Theorem 12.6.16 (Elliptic Regularity Theorem)

We consider the elliptic operator $L(\mathbf{x}, D)$ defined in (12.6.1) and assume that $s \in \mathbb{R}$. If $f \in H^{s,2}_{loc}(\Omega)$ and $u \in \mathcal{D}'(\Omega)$ satisfies $L(\mathbf{x}, D)u = f$ in the sense of distribution, then $u \in H^{s+k,2}_{loc}(\Omega)$.

Proof.

Given $\phi \in \mathcal{D}(\Omega)$, choose an open set U_0 such that $K = \text{supp } \phi \subset U_0 \subset \overline{U_0} \subset \Omega$. This can be done since K is compact and Ω is open. Choose $\psi_0 \in \mathcal{D}(\Omega)$ such that $\psi_0(\mathbf{x}) = 1$ for all $\mathbf{x} \in U_0$.

Since $\text{supp}(\psi_0 u) \subset \text{supp } \psi_0$ is compact, we have that $\psi_0 u \in \mathcal{E}'(\mathbb{R}^n)$; namely, $\psi_0 u$ is a distribution with a compact support⁵. We then have from Corollary 11.7.6 that $\psi_0 u \in H^{t,2}(\mathbb{R}^n)$ for some $t \in \mathbb{R}$. By decreasing t if necessary, we may assume that $m = s + k - t$ is a positive integer. We still have that $\psi_0 u \in H^{t,2}(\mathbb{R}^n)$ because we decrease t .

Choose an open set U_1 such that $K \subset U_1 \subset \overline{U_1} \subset U_0$ and $\psi_1 \in \mathcal{D}(\Omega)$ such that $\text{supp } \psi_1 \subset U_0$ and $\psi_1(\mathbf{x}) = 1$ for all $\mathbf{x} \in U_1$. Proceeding inductively, we can create a sequence $\{\psi_j, U_j\}_{j=0}^{m-1}$ such that $K \subset U_j \subset \overline{U_j} \subset U_{j-1}$, $\text{supp } \psi_j \subset U_{j-1}$ and $\psi_j(\mathbf{x}) = 1$ for all $\mathbf{x} \in U_j$. Finally, set $\psi_m = \phi$.

We prove by induction on j that $\psi_j u \in H^{t+j,2}(\mathbb{R}^n)$ for $0 \leq j \leq m$. The result is true for $j = 0$ by construction. suppose that $\psi_j u \in H^{t+j,2}(\mathbb{R}^n)$ for $j < m$. We have for all $\mu \in \mathcal{D}(\mathbb{R}^n)$ that

$$\begin{aligned} (\psi_{j+1} L(\mathbf{x}, D)(\psi_j u))(\mu) &= (L(\mathbf{x}, D)(\psi_j u))(\psi_{j+1} \mu) = \sum_{|\alpha| \leq k} a_\alpha(\mathbf{x}) D^\alpha(\psi_j u)(\psi_{j+1} \mu) \\ &= \sum_{|\alpha| \leq k} (-1)^{|\alpha|} a_\alpha(\mathbf{x}) (\psi_j u)(D^\alpha(\psi_{j+1} \mu)) = \sum_{|\alpha| \leq k} (-1)^{|\alpha|} a_\alpha(\mathbf{x}) u(D^\alpha(\psi_{j+1} \mu)) \\ &= \sum_{|\alpha| \leq k} a_\alpha(\mathbf{x}) D^\alpha u(\psi_{j+1} \mu) = L(\mathbf{x}, D)(\psi_{j+1} \mu) = \psi_{j+1} L(\mathbf{x}, D)(\mu) = (\psi_{j+1} f)(\mu), \end{aligned}$$

where the fourth equality comes from $(\psi_j u)(\eta) = u(\psi_j \eta) = u(\eta)$ for all test functions $\eta \in \mathcal{D}(\Omega)$ with $\text{supp } \eta \subset U_j$ because $\psi_j(\mathbf{x}) = 1$ for all $\mathbf{x} \in U_j$. Moreover, $\text{supp } D^\alpha(\psi_{j+1} \mu) \subset U_j$. Hence,

$$\begin{aligned} L(\mathbf{x}, D)(\psi_{j+1} u) &= L(\mathbf{x}, D)(\psi_{j+1} \psi_j u) = \psi_{j+1} L(\mathbf{x}, D)(\psi_j u) + [L(\mathbf{x}, D), \psi_{j+1}](\psi_j u) \\ &= \psi_{j+1} f + [L(\mathbf{x}, D), \psi_{j+1}](\psi_j u), \end{aligned}$$

where $\psi_{j+1} f \in H^{s,2}(\mathbb{R}^n)$ because $f \in H^{s,2}_{loc}(\mathbb{R}^n)$ and $\psi_{j+1} \in \mathcal{D}(\Omega)$, and $[L(\mathbf{x}, D), \psi_{j+1}](\psi_j u) \in H^{t+j-k+1,2}(\mathbb{R}^n)$ according to Lemma 12.6.14 because $\psi_j u \in H^{t+j,2}(\mathbb{R}^n)$. Therefore,

⁵Recall that $\psi_0 u$ is defined by $(\psi_0 u)(\mu) = u(\psi_0 \mu)$ for all $\mu \in \mathcal{D}(\mathbb{R}^n)$. Its support is a subset of $\text{supp } \psi_0$ because $(\psi_0 u)(\mu) = 0$ for all $\mu \in \mathcal{D}(\mathbb{R}^n)$ with $\text{supp } \mu \cap \text{supp } \psi_0 = \emptyset$.

$L(\mathbf{x}, D)(\psi_{j+1}u) \in H^{t+j-k+1,2}(\mathbb{R}^n)$. Moreover, $\psi_{j+1}u = \psi_{j+1}\psi_ju \in H^{t+j,2}(\mathbb{R}^n)$ according to Proposition 11.1.22 because $\psi_ju \in H^{t+j,2}(\mathbb{R}^n)$ by our hypothesis of induction. It then follows from Lemma 12.6.15 that $\psi_{j+1}u \in H_{loc}^{t+j+1,2}(\Omega)$.

Given $\eta \in \mathcal{D}(\Omega)$ with $\eta(\mathbf{x}) = 1$ for all $\mathbf{x} \in V_j$, we have that $\eta\psi_{j+1}u \in H^{t+j+1,2}(\mathbb{R}^n)$ according to Proposition 11.1.23. However, $\eta\psi_{j+1}u = \psi_{j+1}u$ on $\mathbb{R}^n$ because

$$(\eta\psi_{j+1}u)(\mu) = u(\eta\psi_{j+1}\mu) = u(\psi_{j+1}\mu) = (\psi_{j+1}u)(\mu)$$

for all $\mu \in \mathcal{D}(\mathbb{R}^n)$ because $\text{supp } \psi_{j+1} \subset V_j$ and so $\eta\psi_{j+1}\xi = \psi_{j+1}\xi$. Therefore, $\psi_{j+1}u \in H^{t+j+1,2}(\mathbb{R}^n)$. This complete the proof by induction.

For $j = m$, we get $\phi u = \psi_mu \in H^{t+m,2}(\mathbb{R}^n) = H^{s+k,2}(\mathbb{R}^n)$. Since ϕ is arbitrary, we get that $u \in H_{loc}^{s+k,2}(\mathbb{R}^n)$ according to Proposition 11.1.23. ■

It is true that the previous theorem is a very important theorem for the internal regularity of elliptic operators. However, its corollary given below is probably the most famous result.

Corollary 12.6.17

We consider the elliptic operator $L(\mathbf{x}, D)$ defined in (12.6.1). If $L(\mathbf{x}, D)u = f \in C^\infty(\Omega)$ in the sense of distributions, then $u \in C^\infty(\Omega)$. Namely, $L(\mathbf{x}, D)$ is hypoelliptic.

Proof.

Since $f \in C^\infty(\Omega)$, we have that $f \in H_{loc}^{s,2}(\Omega)$ for all s according to Proposition 11.1.23. Therefore, $u \in H_{loc}^{s+2,2}(\Omega)$ for all s according to Theorem 12.6.16.

By the Sobolev Lemma, Theorem 11.7.3, and the definition of the localized Sobolev spaces, this implies that $u \in C^\infty(\Omega)$. To be more precise, let V be an open bounded subset such that $\bar{V} \subset \Omega$ and let k be a positive integer. Choose $s > k + n/2$. By definition of the localized Sobolev spaces, there exists $v \in H^{s,2}(\mathbb{R}^n)$ such that $v = u$ on V in the sense of distributions. By the Sobolev Lemma, $v \in C^k(\mathbb{R}^n)$. Thus $u \in C^k(V)$. Since V and k are arbitrary, this implies that $u \in C^k(\Omega)$ for all non-negative integer k . ■

12.6.2 Regularity at the Boundary

The title of this section is a little bit misleading but it is traditionally used for the topic of this section. The results in this section address the regularity of weak solutions of elliptic partial differential equations on the entire bounded domain Ω of the equations. It is however true that the boundary plays a significant role.

Let Ω be a bounded open subset of $\mathbb{R}^n$. In this section, we consider the elliptic operator

$$L(\mathbf{x}, D) = \sum_{|\alpha| \leq 2} a_\alpha(\mathbf{x}) D^\alpha \quad (12.6.32)$$

on $\bar{\Omega}$ with $a_\alpha \in \mathcal{S}(\mathbb{R}^n)$ for all α . Our main result is the next theorem. For this theorem, we go back to the divergence form of the elliptic operator $L(\mathbf{x}, D)$ presented in Section 12.2.

Theorem 12.6.18

Let Ω be a bounded open subset of $\mathbb{R}^n$ (or $\mathbb{R}_+^n$) with a boundary $\partial\Omega$ is of class C^{q+2} . Let $X = H^{1,2}(\Omega)$ or $X = H_0^{1,2}(\Omega)$, and suppose that $u \in X$ is a solution of the variational problem

$$B(v, u) = \sum_{|\alpha|, |\beta| \leq 1} \int_{\Omega} \overline{b_{\alpha, \beta}(\mathbf{x})} D^{\beta} u D^{\alpha} v \, d\mathbf{x} = \int_{\Omega} \bar{f} v \, d\mathbf{x}$$

for $v \in X$, obtained from the elliptic operator (12.6.32) on $\bar{\Omega}$. If B is weakly coercive and $f \in H^{q,2}(\Omega)$, then $u \in H^{q+2,2}(\Omega)$. Moreover, there exists a constant C depending only on Ω such that

$$\|u\|_{q+2,2,\Omega} \leq C (\|f\|_{q,2,\Omega} + \|u\|_{2,\Omega}) . \quad (12.6.33)$$

The proof of Theorem 12.6.18 will be a consequence of the next proposition. Moreover, we will only prove the theorem for elliptic operators of order two; namely, $m = 2$ and $k = 1$.

We need a little lemma first.

Lemma 12.6.19

Let Ω be an open subset of $\mathbb{R}^n$, and k and q be two integers. There exists a constant C such that $\|\Delta_{\mathbf{h}}^{\gamma} f\|_{k,2,\Omega} \leq C \|D^{\gamma} f\|_{k,2,\Omega}$ for all $f \in H^{k,2}(\Omega)$ with compact support in Ω , multi-indices γ such that $|\gamma| \leq q$, and $\mathbf{h} \in \mathbb{R}^n$ such that $\text{supp } f + B_{q\|\mathbf{h}\|}(\mathbf{0}) \subset \Omega$.

Proof.

We have from Lemma 11.1.7 that $\underline{f} \in H^{k,2}(\mathbb{R}^n)$. It follows from Proposition 12.6.4 that

$$\|\Delta_{\mathbf{h}}^{\gamma} \underline{f}\|_{k,2,\rho} \leq \|D^{\gamma} \underline{f}\|_{k,2,\rho} .$$

Since the norms $\|\cdot\|_{k,2,\mathbb{R}^n}$ and $\|\cdot\|_{k,\rho}$ are equivalent, there exist positive constants C_1 and C_2 such

$$C_1 \|g\|_{k,2,\mathbb{R}^n} \leq \|g\|_{k,\rho} \leq C_2 \|g\|_{k,2,\mathbb{R}^n}$$

for all $g \in H^{k,2}(\mathbb{R}^n)$.

Since $\text{supp } D^{\alpha} f \subset D f \subset \Omega$ and $\text{supp } \Delta_{\mathbf{h}}^{\gamma} f \subset \text{supp } f + B_{q\|\mathbf{h}\|}(\mathbf{0}) \subset \Omega$, we have that $D^{\alpha} \underline{f} = \underline{D^{\alpha} f}$ and $\Delta_{\mathbf{h}}^{\gamma} \underline{f} = \underline{\Delta_{\mathbf{h}}^{\gamma} f}$. Thus,

$$\begin{aligned} \|\Delta_{\mathbf{h}}^{\gamma} f\|_{k,2,\Omega} &= \|\Delta_{\mathbf{h}}^{\gamma} \underline{f}\|_{k,2,\mathbb{R}^n} \leq C_1^{-1} \|\Delta_{\mathbf{h}}^{\gamma} \underline{f}\|_{k,\rho} \leq C_1^{-1} \|D^{\gamma} \underline{f}\|_{k,\rho} \\ &\leq C_1^{-1} C_2 \|D^{\gamma} \underline{f}\|_{k,2,\mathbb{R}^n} = C \|D^{\gamma} f\|_{k,2,\Omega} \end{aligned}$$

with $C = C_1^{-1} C_2$. Note that C does not depend of f or $\mathbf{h}$. ■

We begin by assuming that $\Omega = \mathbb{R}_+^n$.

Proposition 12.6.20

Let $\Omega = \mathbb{R}_+^n$ and $X = H^{1,2}(\Omega)$ or $X = H_0^{1,2}(\Omega)$. For $r > 0$, let $\Omega_r = \Omega \cap B_r(\mathbf{0})$ and $X_r = \{u \in X : \text{supp } u \subset \Omega_\rho \text{ for some } \rho < r\}$. Suppose that $f \in H^{q,2}(\Omega_r)$ for some $r > 0$. If $u \in X$ is a solution of the variational problem

$$B(v, u) = \sum_{|\alpha|, |\beta| \leq 1} \int_{\Omega_r} \overline{b_{\alpha, \beta}(\mathbf{x})} D^\beta u D^\alpha v \, d\mathbf{x} = \int_{\Omega_r} \bar{f} v \, d\mathbf{x} \quad (12.6.34)$$

for $v \in X_r$, where B is weakly coercive, then $u \in H^{q+2,2}(\Omega_\rho)$ for $\rho < r$ and there exists a constant $C = C(q, \rho)$ such that

$$\|u\|_{q+2,2,\Omega_\rho} \leq C (\|f\|_{q,2,\Omega_r} + \|u\|_{1,2,\Omega_r}) .$$

Proof.

Claim A: If $\rho < r$ and $j < q+2$, then $D^\gamma u \in H^{1,2}(\Omega_\rho)$ for all multi-indices γ such that $|\gamma| \leq j$ and $\gamma_n = 0$. Moreover, there exists a positive constant $C_1 = C_1(j, \rho)$ such that

$$\|D^\gamma u\|_{1,2,\Omega_\rho} \leq C_1 (\|f\|_{q,2,\Omega_r} + \|u\|_{1,2,\Omega_r}) \quad (12.6.35)$$

for all multi-indices γ such that $|\gamma| \leq j$ and $\gamma_n = 0$.

The proof is by induction on j . The claim is obviously true for $j = 0$. Suppose that the claim is true for $0 \leq j < q+1$. Let $r_1 = \frac{2\rho+r}{3}$ and $r_2 = \frac{\rho+2r}{3}$. We have $\rho < r_1 < r_2 < r$.

The induction hypothesis applied to the domain Ω_{r_2} instead of Ω_ρ says that $D^\gamma u \in H^{1,2}(\Omega_{r_2})$ for all multi-indices γ with $|\gamma| \leq j$ and $\gamma_n = 0$, and there exists a positive constant $C_2 = C_2(j, r_2)$ such that

$$\|D^\gamma u\|_{1,2,\Omega_{r_2}} \leq C_2 (\|f\|_{q,2,\Omega_r} + \|u\|_{1,2,\Omega_r}) \quad (12.6.36)$$

for all multi-indices γ such that $|\gamma| \leq j$ and $\gamma_n = 0$.

Choose $\psi \in \mathcal{D}(\mathbb{R}^n)$ such that $\text{supp } \psi \subset B_{r_1}(\mathbf{0})$ and $\psi = 1$ on $B_\rho(\mathbf{0})$. Let γ be a multi-index such that $|\gamma| = j+1$ with $\gamma_n = 0$. If $\|\mathbf{h}\| < (r_2 - r_1)/(j+1)$, then $\Delta_{\mathbf{h}}^\gamma(\psi u) \in X_{r_2}$ because $\text{supp } \psi \subset B_{r_1}(\mathbf{0})$, and $\Delta_{\mathbf{h}}^\gamma v \in X_{r_2}$ for $v \in X_{r_1}$. In the sequel, we always assume that $\|\mathbf{h}\|$ satisfies $\|\mathbf{h}\| < (r_2 - r_1)/(j+1)$.

i) We prove that there exists a constant $C_3 = C_3(j, r_2, \psi)$ such that

$$|B(v, \Delta_{\mathbf{h}}^\gamma(\psi u))| \leq C_3 \|v\|_{1,2,\Omega_{r_2}} (\|f\|_{q,2,\Omega_r} + \|u\|_{1,2,\Omega_r}) \quad (12.6.37)$$

for all $v \in X_{r_2}$.

Since the difference quotient operator (12.6.4) commutes with the derivative operator, we get

$$B(v, \Delta_{\mathbf{h}}^\gamma(\psi u)) = \sum_{|\alpha|, |\beta| \leq 1} \int_{\Omega_r} \overline{b_{\alpha, \beta}(\mathbf{x})} D^\beta \Delta_{\mathbf{h}}^\gamma(\psi u) D^\alpha v \, d\mathbf{x}$$

$$= \sum_{|\alpha|, |\beta| \leq 1} \int_{\Omega_r} \overline{b_{\alpha, \beta}(\mathbf{x}) \Delta_{\mathbf{h}}^{\gamma} D^{\beta}(\psi u)} D^{\alpha} v \, d\mathbf{x} = \sum_{|\alpha|, |\beta| \leq 1} \int_{\Omega_r} \overline{\Delta_{\mathbf{h}}^{\gamma} (b_{\alpha, \beta}(\mathbf{x}) D^{\beta}(\psi u))} D^{\alpha} v \, d\mathbf{x} + E_1$$

for all $v \in X_{r_2}$, where

$$\begin{aligned} E_1(v, u) &= \sum_{|\alpha|, |\beta| \leq 1} \int_{\Omega_r} \left(\overline{b_{\alpha, \beta}(\mathbf{x}) \Delta_{\mathbf{h}}^{\gamma} D^{\beta}(\psi u)} - \overline{\Delta_{\mathbf{h}}^{\gamma} (b_{\alpha, \beta}(\mathbf{x}) D^{\beta}(\psi u))} \right) D^{\alpha} v \, d\mathbf{x} \\ &= \sum_{|\alpha|, |\beta| \leq 1} \int_{\Omega_r} \overline{-[\Delta_{\mathbf{h}}^{\gamma}, b_{\alpha, \beta}(\mathbf{x})] D^{\beta}(\psi u)} D^{\alpha} v \, d\mathbf{x} \end{aligned}$$

for all $v \in X_{r_2}$. Furthermore, we may write

$$B(v, \Delta_{\mathbf{h}}^{\gamma}(\psi u)) = \sum_{|\alpha|, |\beta| \leq 1} \int_{\Omega_r} \overline{\Delta_{\mathbf{h}}^{\gamma} (b_{\alpha, \beta}(\mathbf{x}) (\psi D^{\beta} u))} D^{\alpha} v \, d\mathbf{x} + E_1 + E_2 \quad (12.6.38)$$

for all $v \in X_{r_2}$, where

$$E_2(v, u) = \sum_{|\alpha|, |\beta| \leq 1} \int_{\Omega_r} \overline{\Delta_{\mathbf{h}}^{\gamma} (b_{\alpha, \beta}(\mathbf{x}) (D^{\beta}(\psi u) - \psi D^{\beta} u))} D^{\alpha} v \, d\mathbf{x}$$

for all $v \in X_{r_2}$. It follows from Proposition 12.6.2 that

$$B(v, \Delta_{\mathbf{h}}^{\gamma}(\psi u)) = (-1)^{j+1} \sum_{|\alpha|, |\beta| \leq 1} \int_{\Omega_r} \overline{b_{\alpha, \beta}(\mathbf{x}) (\psi D^{\beta} u)} \Delta_{-\mathbf{h}}^{\gamma} D^{\alpha} v \, d\mathbf{x} + E_1 + E_2$$

for all $v \in X_{r_2}$. Again, since the difference quotient operator commutes with the derivative operator, we get

$$\begin{aligned} B(v, \Delta_{\mathbf{h}}^{\gamma}(\psi u)) &= (-1)^{j+1} \sum_{|\alpha|, |\beta| \leq 1} \int_{\Omega_r} \overline{b_{\alpha, \beta}(\mathbf{x}) (\psi D^{\beta} u)} D^{\alpha} \Delta_{-\mathbf{h}}^{\gamma} v \, d\mathbf{x} + E_1 + E_2 \\ &= (-1)^{j+1} \sum_{|\alpha|, |\beta| \leq 1} \int_{\Omega_r} \left(\overline{b_{\alpha, \beta}(\mathbf{x}) D^{\beta} u} \right) D^{\alpha} (\bar{\psi} \Delta_{-\mathbf{h}}^{\gamma} v) \, d\mathbf{x} + E_1 + E_2 + E_3 \end{aligned}$$

for all $v \in X_{r_2}$, where

$$E_3(v, u) = (-1)^{j+1} \sum_{|\alpha|, |\beta| \leq 1} \int_{\Omega_r} \left(\overline{b_{\alpha, \beta}(\mathbf{x}) D^{\beta} u} \right) (\bar{\psi} D^{\alpha} \Delta_{-\mathbf{h}}^{\gamma} v - D^{\alpha} (\bar{\psi} \Delta_{-\mathbf{h}}^{\gamma} v)) \, d\mathbf{x}$$

for all $v \in X_{r_2}$. Finally, since $\bar{\psi} \Delta_{-\mathbf{h}}^{\gamma} v \in X_{r_2} \subset X$, we have from (12.6.34) that

$$B(v, \Delta_{\mathbf{h}}^{\gamma}(\psi u)) = (-1)^{j+1} \int_{\Omega_r} (\bar{\psi} \Delta_{-\mathbf{h}}^{\gamma} v) \bar{f} \, d\mathbf{x} + E_1 + E_2 + E_3 = E_1 + E_2 + E_3 + E_4$$

for all $v \in X_{r_2}$, where

$$\begin{aligned} E_4(v, u) &= (-1)^{j+1} \int_{\Omega_r} (\Delta_{-\mathbf{h}}^{\gamma} v) \bar{\psi} f \, d\mathbf{x} = (-1)^{j+1} \int_{\Omega_r} (\Delta_{-\mathbf{h}}^{\tilde{\gamma}} (\Delta_{-h_i e_i} v)) \bar{\psi} f \, d\mathbf{x} \\ &= - \int_{\Omega_r} (\Delta_{h_i e_i} v) \overline{\Delta_{\mathbf{h}}^{\tilde{\gamma}} (\psi f)} \, d\mathbf{x} \end{aligned}$$

for all $v \in X_{r_2}$, where we have selected $1 \leq i < n$ such that $\gamma_i > 0$ and set $\tilde{\gamma} = \gamma - \mathbf{e}_i$. So $|\tilde{\gamma}| = j$ with $\tilde{\gamma}_n = 0$. We have used Proposition 12.6.2 for the last equality in the expression above.

We now find upperbounds of the form $C_{E_i} \|v\|_{1,2,\Omega_{r_2}} (\|f\|_{q,2,\Omega_r} + \|u\|_{1,2,\Omega_r})$ for each of the E_i , where the C_{E_i} are constants.

We first consider E_1 . Since $\text{supp}(\psi u) \subset \Omega_{r_1}$ and $\text{supp} \Delta_{\mathbf{h}}^{\gamma}(\psi u) \subset \Omega_{r_2}$, we may use Schwarz inequality in $L^2(\Omega_r)$ and Corollary 12.6.10 to get a constant Q_1 such that

$$\begin{aligned} |E_1(u, v)| &\leq \sum_{|\alpha|, |\beta| \leq 1} \|D^{\alpha} v\|_{2,\Omega_{r_2}} \left\| [\Delta_{\mathbf{h}}^{\gamma}, b_{\alpha,\beta}(\mathbf{x})] D^{\beta}(\psi u) \right\|_{2,\Omega_{r_2}} \\ &\leq Q_1 \|v\|_{1,2,\Omega_{r_2}} \sum_{\eta < \gamma, |\beta| \leq 1} \|D^{\eta} D^{\beta}(\psi u)\|_{2,\Omega_{r_2}} \leq Q_1 \|v\|_{1,2,\Omega_{r_2}} \sum_{\eta < \gamma} \|D^{\eta} u\|_{1,2,\Omega_{r_2}} \\ &\leq Q_1 \|v\|_{1,2,\Omega_{r_2}} \sum_{|\eta| \leq j, \eta_n = 0} \|D^{\eta} u\|_{1,2,\Omega_{r_2}} \end{aligned}$$

for all $v \in X_{r_2}$. It follows from (12.6.36), our induction hypothesis, that

$$|E_1(v, u)| \leq C_{E_1} \|v\|_{1,2,\Omega_{r_2}} (\|f\|_{q,2,\Omega_r} + \|u\|_{1,2,\Omega_r})$$

for all $v \in X_{r_2}$, where $C_{E_1} = Q_1 C_2 \sum_{|\eta| \leq j, \eta_n = 0} 1$.

For E_2 , we have

$$D^{\beta}(\psi u) - \psi D^{\beta} u = \begin{cases} (D^{\beta} \psi) u & \text{if } |\beta| = 1 \\ 0 & \text{if } |\beta| = 0 \end{cases}$$

Hence, we get from Schwarz inequality in $L^2(\Omega_r)$ that

$$\begin{aligned} |E_2(u, v)| &\leq \sum_{|\beta|=1, |\alpha| \leq 1} \|D^{\alpha} v\|_{2,\Omega_{r_2}} \left\| \Delta_{\mathbf{h}}^{\gamma} (b_{\alpha,\beta} (D^{\beta} \psi) u) \right\|_{2,\Omega_{r_2}} \\ &\leq \|v\|_{1,2,\Omega_{r_2}} \sum_{|\beta|=1, |\alpha| \leq 1} \left\| \Delta_{\mathbf{h}}^{\gamma} (b_{\alpha,\beta} (D^{\beta} \psi) u) \right\|_{2,\Omega_{r_2}} \leq P_1 \|v\|_{1,2,\Omega_{r_2}} \sum_{|\beta|=1, |\alpha| \leq 1} \|D^{\gamma} (b_{\alpha,\beta} (D^{\beta} \psi) u)\|_{2,\Omega_{r_2}} \\ &\leq P_1 \|v\|_{1,2,\Omega_{r_2}} \sum_{|\beta|=1, |\alpha| \leq 1} \left(\sum_{\eta \leq \gamma} Q_{\alpha,\beta,\eta} \|D^{\eta} u\|_{2,\Omega_{r_2}} \right) = P_1 \|v\|_{1,2,\Omega_{r_2}} \sum_{\eta \leq \gamma} \left(\sum_{|\beta|=1, |\alpha| \leq 1} Q_{\alpha,\beta,\eta} \right) \|D^{\eta} u\|_{2,\Omega_{r_2}} \\ &\leq P_1 Q_2 \|v\|_{1,2,\Omega_{r_2}} \sum_{\eta \leq \gamma} \|D^{\eta} u\|_{2,\Omega_{r_2}} \leq P_1 Q_2 \|v\|_{1,2,\Omega_{r_2}} \sum_{|\eta| \leq j, \eta_n = 0} \|D^{\eta} u\|_{1,2,\Omega_{r_2}} \end{aligned}$$

for all $v \in X_{r_2}$, where P_1 is a constant given by Lemma 12.6.19 with $\Omega = \Omega_{r_2}$ and $k = 0$. Moreover, the constants $Q_{\alpha,\beta,\eta}$ are summations of terms of the form $\sup_{\mathbf{x} \in \bar{\Omega}} |D^{\mu} b_{\alpha,\beta}(\mathbf{x}) D^{\beta+\nu} \psi(\mathbf{x})|$

for $\mu + \nu \leq \gamma - \eta$, and $Q_2 = \max_{\eta \leq \gamma} \sum_{|\beta|=1, |\alpha| \leq 1} Q_{\alpha,\beta,\eta}$.

As for E_1 , we can use the induction hypothesis to find a constant C_{E_2} such that

$$|E_2(v, u)| \leq C_{E_2} \|v\|_{1,2,\Omega_{r_2}} (\|f\|_{q,2,\Omega_r} + \|u\|_{1,2,\Omega_r})$$

for all $v \in X_{r_2}$.

For E_3 , we have

$$\bar{\psi} D^\alpha \Delta_{-\mathbf{h}}^\gamma v - D^\alpha (\bar{\psi} \Delta_{-\mathbf{h}}^\gamma v) = \begin{cases} - (D^\alpha \bar{\psi}) \Delta_{-\mathbf{h}}^\gamma v & \text{if } |\alpha| = 1 \\ 0 & \text{if } |\alpha| = 0 \end{cases}$$

Hence,

$$|E_3(v, u)| = \left| \sum_{|\alpha|=1, |\beta| \leq 1} \int_{\Omega_r} \overline{b_{\alpha, \beta}(\mathbf{x}) D^\beta u} ((D^\alpha \bar{\psi}) \Delta_{-\mathbf{h}}^\gamma v) d\mathbf{x} \right|$$

for all $v \in X_{r_2}$. Choose i such that $\gamma_i \neq 0$ and set $\tilde{\gamma} = \gamma - e_i$. Using proposition 12.6.2, we get

$$\begin{aligned} |E_3(v, u)| &= \left| \sum_{|\alpha|=1, |\beta| \leq 1} \int_{\Omega_r} \overline{b_{\alpha, \beta}(\mathbf{x}) D^\beta u D^\alpha \psi} (\Delta_{-\mathbf{h}}^{\tilde{\gamma}} (\Delta_{-h_i e_i} v)) d\mathbf{x} \right| \\ &= \left| (-1)^{|\tilde{\gamma}|} \sum_{|\alpha|=1, |\beta| \leq 1} \int_{\Omega_r} \overline{\Delta_{-\mathbf{h}}^{\tilde{\gamma}} (b_{\alpha, \beta}(\mathbf{x}) D^\beta u D^\alpha \psi)} \Delta_{-h_i e_i} v d\mathbf{x} \right| \end{aligned}$$

for all $v \in X_{r_2}$.

It follows from Schwarz inequality in $L^2(\Omega_r)$ that

$$\begin{aligned} |E_3(v, u)| &\leq \sum_{|\alpha|=1, |\beta| \leq 1} \|\Delta_{-h_i e_i} v\|_{2, \Omega_{r_2}} \|\Delta_{-\mathbf{h}}^{\tilde{\gamma}} (b_{\alpha, \beta}(\mathbf{x}) D^\beta u D^\alpha \psi)\|_{2, \Omega_{r_2}} \\ &\leq P_1^2 \sum_{|\alpha|=1, |\beta| \leq 1} \|D_{x_i} v\|_{2, \Omega_{r_2}} \|D^{\tilde{\gamma}} (b_{\alpha, \beta}(\mathbf{x}) D^\beta u D^\alpha \psi)\|_{2, \Omega_{r_2}} \\ &\leq P_1^2 \|v\|_{1, 2, \Omega_{r_2}} \sum_{|\alpha|=1, |\beta| \leq 1} \left(\sum_{\eta \leq \tilde{\gamma}} \tilde{Q}_{\alpha, \beta, \eta} \|D^{\beta + \eta} u\|_{2, \Omega_{r_2}} \right) \\ &= P_1^2 \|v\|_{1, 2, \Omega_{r_2}} \sum_{\eta \leq \tilde{\gamma}} \left(\sum_{|\alpha|=1, |\beta| \leq 1} \tilde{Q}_{\alpha, \beta, \eta} \right) \|D^\eta u\|_{1, 2, \Omega_{r_2}} \\ &\leq Q_3 \|v\|_{1, 2, \Omega_{r_2}} \sum_{\eta \leq \tilde{\gamma}} \|D^\eta u\|_{1, 2, \Omega_{r_2}} \leq Q_3 \|v\|_{1, 2, \Omega_r} \sum_{|\eta| \leq j, \eta_n = 0} \|u\|_{1, 2, \Omega_{r_2}} \end{aligned}$$

for all $v \in X_{r_2}$, where P_1 is defined above. Moreover, the constants $\tilde{Q}_{\alpha, \beta, \eta}$ are summations of terms of the form $\sup_{\mathbf{x} \in \overline{\Omega_r}} |D^\mu b_{\alpha, \beta}(\mathbf{x}) D^{\alpha + \nu} \psi(\mathbf{x})|$ for $\mu + \nu \leq \tilde{\gamma} - \eta$, and $Q_3 =$

$$P_1^2 \max_{\eta \leq \tilde{\gamma}} \sum_{|\beta|=1, |\alpha| \leq 1} \tilde{Q}_{\alpha, \beta, \eta}.$$

As for E_1 and E_2 , we can use the induction hypothesis to find a constant C_{E_3} such that

$$|E_3(v, u)| \leq C_{E_3} \|v\|_{1, 2, \Omega_{r_2}} (\|f\|_{q, 2, \Omega_r} + \|u\|_{1, 2, \Omega_r})$$

for all $v \in X_r$.

For E_4 , we first choose an index i such that $\gamma_i \neq 0$. Using Schwarz inequality in $L^2(\Omega_r)$, we get

$$|E_4(v, u)| \leq \|\Delta_{-h_i e_i}^{\mathbf{e}_i} v\|_{2, \Omega_r} \|\Delta_{-\mathbf{h}}^{\tilde{\gamma}} (\psi f)\|_{2, \Omega_r} \leq P_1^2 \|D_{x_i} v\|_{2, \Omega_r} \|D^{\tilde{\gamma}} (\psi f)\|_{2, \Omega_r}$$

$$\leq P_1^2 \|v\|_{1,2,\Omega_r} \|D^{\tilde{\gamma}}(\psi f)\|_{2,\Omega_r} \leq C_{E_4} \|v\|_{1,2,\Omega_{r_2}} \|f\|_{q,2,\Omega_r}$$

for all $v \in X_r$, where P_1 is defined above and C_{E_4} is defined as it follows. The last inequality comes from $|\tilde{\gamma}| \leq j \leq k$ since we assume that $j < k + 1$, and the relation

$$\|D^{\tilde{\gamma}}(\psi f)\|_{2,\Omega_r} \leq \sum_{\mu+\nu=\tilde{\gamma}} \binom{\tilde{\gamma}}{\nu} \|D^{\mu} f\|_{2,\Omega_r} \|D^{\nu} \psi\|_{2,\Omega_r} \leq \left(\sum_{\nu \leq \tilde{\gamma}} \binom{\tilde{\gamma}}{\nu} \|D^{\nu} \psi\|_{2,\Omega_r} \right) \|f\|_{q,2,\Omega_r}.$$

We therefore have that $C_{E_4} = P_1^2 \sum_{\nu \leq \tilde{\gamma}} \binom{\tilde{\gamma}}{\nu} \|D^{\nu} \psi\|_{2,\Omega_r}$. Hence,

$$|E_4(v, u)| \leq C_{E_4} \|v\|_{1,2,\Omega_{r_2}} (\|f\|_{q,2,\Omega_r} + \|u\|_{1,2,\Omega_r})$$

for all $v \in X_r$.

We have proved that (12.6.37) with $C_3 = C_{E_1} + C_{E_2} + C_{E_3} + C_{E_4}$. Note that C_3 depends indirectly of ρ because C_3 depends of r_2 and $r_2 = (\rho + 2r)/3$.

ii) Let C_0 and λ be the two constants in the definition of a weakly coercive form applied to B ; namely,

$$\operatorname{Re} B(v, v) \geq C_0 \|v\|_{1,2}^2 - \lambda \|v\|_2^2 \quad (12.6.39)$$

for all $v \in X_r$. If we substitute $v = \Delta_{\mathbf{h}}^{\alpha}(\psi u)$ in (12.6.39) and use (12.6.37), we get that

$$\begin{aligned} \|\Delta_{\mathbf{h}}^{\gamma}(\psi u)\|_{1,2,\Omega_{r_2}}^2 &\leq \frac{1}{C_0} \left(|B(\Delta_{\mathbf{h}}^{\gamma}(\psi u), \Delta_{\mathbf{h}}^{\gamma}(\psi u))| + \lambda \|\Delta_{\mathbf{h}}^{\gamma}(\psi u)\|_{2,\Omega_{r_2}}^2 \right) \\ &\leq \frac{1}{C_0} \left(C_3 \|\Delta_{\mathbf{h}}^{\gamma}(\psi u)\|_{1,2,\Omega_{r_2}} (\|f\|_{q,2,\Omega_r} + \|u\|_{1,2,\Omega_r}) + \lambda \|\Delta_{\mathbf{h}}^{\gamma}(\psi u)\|_{2,\Omega_{r_2}}^2 \right) \\ &\leq \frac{1}{C_0} \left(C_3 \|\Delta_{\mathbf{h}}^{\gamma}(\psi u)\|_{1,2,\Omega_{r_2}} (\|f\|_{q,2,\Omega_r} + \|u\|_{1,2,\Omega_r}) \right. \\ &\quad \left. + \lambda \|\Delta_{\mathbf{h}}^{\gamma}(\psi u)\|_{2,\Omega_{r_2}} \|\Delta_{\mathbf{h}}^{\gamma}(\psi u)\|_{1,2,\Omega_{r_2}} \right). \end{aligned}$$

Hence,

$$\|\Delta_{\mathbf{h}}^{\gamma}(\psi u)\|_{1,2,\Omega_{r_2}} \leq \frac{1}{C_0} \left(C_3 (\|f\|_{q,2,\Omega_r} + \|u\|_{1,2,\Omega_r}) + \lambda \|\Delta_{\mathbf{h}}^{\gamma}(\psi u)\|_{2,\Omega_{r_2}} \right).$$

Choose i such that $\gamma_i \neq 0$ and set $\tilde{\gamma} = \alpha - e_i$. From (12.6.36), that we can use because $|\tilde{\gamma}| \leq j$, we find that

$$\begin{aligned} \|\Delta_{\mathbf{h}}^{\gamma}(\psi u)\|_{2,\Omega_{r_2}} &= \|\Delta_{\mathbf{h}e_1}^{e_i} \Delta_{\mathbf{h}}^{\tilde{\gamma}}(\psi u)\|_{2,\Omega_{r_2}} \leq P_1 \|D^{e_i}(\Delta_{\mathbf{h}}^{\tilde{\gamma}}(\psi u))\|_{2,\Omega_{r_2}} \leq P_1 \|\Delta_{\mathbf{h}}^{\tilde{\gamma}}(\psi u)\|_{1,2,\Omega_{r_2}} \\ &\leq P_1 P_2 \|D^{\tilde{\gamma}}(\psi u)\|_{1,2,\Omega_{r_2}} \leq C_4 (\|f\|_{q,2,\Omega_r} + \|u\|_{1,2,\Omega_r}) \end{aligned}$$

for $\|\mathbf{h}\| < (r_2 - r_1)/(j+1)$, where P_1 is defined above, P_2 is a constant given by Lemma 12.6.19 with $\Omega = \Omega_{r_2}$ and $k = 1$, and $C_4 = P_1 P_2 C_2$. Hence,

$$\begin{aligned} \|\Delta_{\mathbf{h}}^{\gamma}(\psi u)\|_{1,2,\Omega_{r_2}} &\leq \frac{1}{C_0} \left(C_3 (\|f\|_{q,2,\Omega_r} + \|u\|_{1,2,\Omega_r}) + \lambda C_4 (\|f\|_{q,2,\Omega_r} + \|u\|_{1,2,\Omega_r}) \right) \\ &\leq \frac{1}{C_0} (C_3 + \lambda C_4) (\|f\|_{q,2,\Omega_r} + \|u\|_{1,2,\Omega_r}). \end{aligned}$$

With $C_5 = \max \left\{ C_2, \frac{1}{C_0} (C_3 + \lambda C_4) \right\}$, we get

$$\|\Delta_{\mathbf{h}}^{\gamma}(\psi u)\|_{1,2,\Omega_{r_2}} \leq C_5 (\|f\|_{q,2,\Omega_r} + \|u\|_{1,2,\Omega_r}) \quad (12.6.40)$$

for all multi-indices γ such that $|\gamma| \leq j+1$ and $\gamma_n = 0$, and all $\mathbf{h}$ such that $\|\mathbf{h}\| < (r_2 - r_1)/(j+1)$.

We have from Corollary 12.6.8 that $D^{\gamma}(\psi u) \in H^{1,2}(\Omega_{r_2})$ because $\limsup_{\mathbf{h} \rightarrow \mathbf{0}} \|\Delta_{\mathbf{h}}^{\gamma} f\|_{q,2,\Omega_r} \leq C_5 (\|f\|_{q,2,\Omega_r} + \|u\|_{1,2,\Omega_r}) < \infty$, and there exists a constant C_6 such that

$$\|D^{\gamma}(\psi u)\|_{1,2,\Omega_{r_2}} \leq C_6 \|\Delta_{\mathbf{h}}^{\gamma}(\psi u)\|_{1,2,\Omega_{r_2}} \quad (12.6.41)$$

for some (in fact infinitely many) values of $\mathbf{h}$ such that $\|\mathbf{h}\| < (r_2 - r_1)/(j+1)$. Since $\psi = 1$ on Ω_{ρ} , we have that

$$\|D^{\gamma} u\|_{1,2,\Omega_{\rho}} \leq \|D^{\gamma}(\psi u)\|_{1,2,\Omega_{r_2}} \quad (12.6.42)$$

It follows from (12.6.40), (12.6.41) and (12.6.42) that

$$\|D^{\gamma} u\|_{1,2,\Omega_{\rho}} \leq C_1 (\|f\|_{q,2,\Omega_r} + \|u\|_{1,2,\Omega_r})$$

for all multi-indices γ such that $|\gamma| \leq j+1$ and $\gamma_n = 0$, where $C_1 = C_5 C_6$. This complete the proof of claim A.

Claim B: Assume that $\rho < r$, $j \leq q+2$ and γ is a multi-index such that $|\gamma| \leq q+2$ and $\gamma_n \leq j$. Then, $D^{\gamma} u \in L^2(\Omega_{\rho})$ and there exists a constant $C_7 = C_7(j, \rho)$ such that

$$\|D^{\gamma} u\|_{2,\Omega_{\rho}} \leq C_7 (\|f\|_{q,2,\Omega_r} + \|u\|_{1,2,\Omega_r}) \quad (12.6.43)$$

The proof is by induction on j . Claim A shows that this new claim is true for $j = 0$ and $j = 1$. If $\gamma_n = 0$, let $\tilde{\gamma} = \gamma - \mathbf{e}_i$ for one index $0 \leq i < n$ such that $\alpha_i > 0$. If $\gamma_n = 1$, let $\tilde{\gamma} = \gamma - \mathbf{e}_n$. Then $|\tilde{\gamma}| < q+2$ with $\tilde{\gamma}_n = 0$. Hence, from claim A, we have that

$$\|D^{\gamma} u\|_{2,\Omega_{\rho}} \leq \|D^{\tilde{\gamma}} u\|_{1,2,\Omega_{\rho}} \leq C_1(q+1, \rho) (\|f\|_{q,2,\Omega_r} + \|u\|_{1,2,\Omega_r}) \quad .$$

So $C_7(j, \rho) = C_1(q+1, \rho)$ for $j = 0$ and $j = 1$ is acceptable.

Suppose that $j \geq 2$ and that the claim is true for $\gamma_n < j$. Consider γ such that $|\gamma| \leq q+2$ and $\gamma_n = j$. Let $\tilde{\gamma} = \gamma - 2\mathbf{e}_n$. Since the operator $L(\mathbf{x}, D)u = \sum_{|\alpha| \leq 2} a_{\alpha}(\mathbf{x}) D^{\alpha} u$, the initial formulation

of the variational problem (12.6.34), is elliptic on $\overline{\Omega}$, we must have for $\beta = (0, 0, \dots, 0, 2)$ that $a_{\beta}(\mathbf{x}) \neq 0$ for all $\mathbf{x} \in \overline{\Omega}$. We may therefore rewrite $L(\mathbf{x}, D)u = f$ as

$$D^{\beta} u = \frac{1}{a_{\beta}} \left(f - \sum_{|\alpha| \leq 2, \alpha_n \neq 2} a_{\alpha} D^{\alpha} u \right)$$

on Ω . Thus,

$$D^{\gamma} u = D^{\tilde{\gamma}} D^{\beta} u = D^{\tilde{\gamma}} \left(\frac{1}{a_{\beta}} f \right) - D^{\tilde{\gamma}} \left(\frac{1}{a_{\beta}} \sum_{|\alpha| \leq 2, \alpha_n \neq 2} a_{\alpha} D^{\alpha} u \right)$$

$$= \sum_{|\alpha| \leq q} g_\alpha D^\alpha f - \sum_{|\alpha| \leq q+2, \alpha_n \leq j-1} h_\alpha D^\alpha u ,$$

where the $g_\alpha, h_\alpha \in C^\infty(\overline{\Omega})$ for all α . The constraints on the indices for the second sum come from $\tilde{\gamma}_n \leq j-2$ and $\alpha_n < 2$. Hence,

$$\|D^\gamma u\|_{2, \Omega_\rho} \leq \sum_{|\alpha| \leq q} \left(\sup_{\mathbf{x} \in \overline{\Omega}_r} |g_\alpha(\mathbf{x})| \|D^\alpha f\|_{2, \Omega_\rho} \right) + \sum_{|\alpha| \leq q+2, \alpha_n \leq j-1} \left(\sup_{\mathbf{x} \in \overline{\Omega}_r} |h_\alpha(\mathbf{x})| \|D^\alpha u\|_{2, \Omega_\rho} \right) .$$

Since

$$\|D^\alpha f\|_{2, \Omega_\rho} \leq \|f\|_{q, 2, \Omega_\rho} \leq \|f\|_{q, 2, \Omega_r} ,$$

for $|\alpha| \leq q$ and

$$\|D^\alpha u\|_{2, \Omega_\rho} \leq C_7(j-1, \rho) (\|f\|_{q, 2, \Omega_r} + \|u\|_{1, 2, \Omega_r})$$

for $|\alpha| \leq q+2$ with $\alpha_n \leq j-1$ by induction, we get

$$\|D^\gamma u\|_{2, \Omega_\rho} \leq A \|f\|_{q, 2, \Omega_r} + B (\|f\|_{q, 2, \Omega_r} + \|u\|_{1, 2, \Omega_r}) ,$$

where $A = \sum_{|\alpha| \leq q} \sup_{\mathbf{x} \in \overline{\Omega}_r} |g_\alpha(\mathbf{x})|$ and $B = C_7(j-1, \rho) \sum_{|\alpha| \leq q+2, \alpha_n \leq j-1} \sup_{\mathbf{x} \in \overline{\Omega}_r} |h_\alpha(\mathbf{x})|$. Hence,

$$\|D^\gamma u\|_{2, \Omega_\rho} \leq C_7(j, \rho) (\|f\|_{q, 2, \Omega_r} + \|u\|_{1, 2, \Omega_r})$$

with $C_7(j, \rho) = A + B$. This complete the proof of Claim B.

C) To complete the proof, let $\tilde{C} = \max_{0 \leq j \leq q+2} C_7(j, \rho)$, where the $C_7(j, \rho)$ are given by claim B.

Thus $D^\alpha u \in L^2(\Omega_\rho)$ and

$$\|D^\alpha u\|_{2, \Omega_\rho} \leq \tilde{C} (\|f\|_{q, 2, \Omega_r} + \|u\|_{1, 2, \Omega_r})$$

for all multi-indices α such that $|\alpha| \leq q+2$. It follows that $u \in H^{k+2, 2}(\Omega_\rho)$ and

$$\|u\|_{q+2, 2, \Omega_\rho} \leq C (\|f\|_{q, 2, \Omega_r} + \|u\|_{1, 2, \Omega_r})$$

for $C = \tilde{C} \left(\sum_{|\alpha| \leq q+2} 1 \right)^{1/2}$. ■

Before giving the proof of Theorem 12.6.18, we need to provide a lemma like Lemma 12.6.14 but for Ω a bounded open subset of $\mathbb{R}^n$.

Lemma 12.6.21

We consider the elliptic operator $L(\mathbf{x}, D)$ defined in (12.6.1). Suppose that Ω is a bounded open subset of $\mathbb{R}^n$, $\phi \in \mathcal{D}(\Omega)$ and $q \in \mathbb{N}$. The operator $[L(\mathbf{x}, D), \phi]$ defined by

$$[L(\mathbf{x}, D), \phi](v) = L(\mathbf{x}, D)(\phi v) - \phi L(\mathbf{x}, D)v$$

for $v \in H^{q, 2}(\Omega)$ is a bounded mapping from $H^{q, 2}(\Omega)$ into $H^{q-k+1, 2}(\Omega)$.

Proof.

The proof is identical to the proof of Lemma 12.6.14 with $\mathbb{R}^n$ replaced by Ω except that (12.6.31) becomes

$$\|A_\eta(\mathbf{x}) D^\eta v\|_{q-k+1,2,\Omega} \leq C_\eta \|D^\eta v\|_{q-k+1,2,\Omega} \leq C_\eta \|v\|_{q-k+|\eta|+1,2,\Omega} \leq C_\eta \|v\|_{q,2,\Omega},$$

where Proposition 11.1.17 is used to get the constants C_η and the definition of the norm $\|\cdot\|_{m,2,\Omega}$ with $m \in \mathbb{N}$ is used to get the second inequality. ■

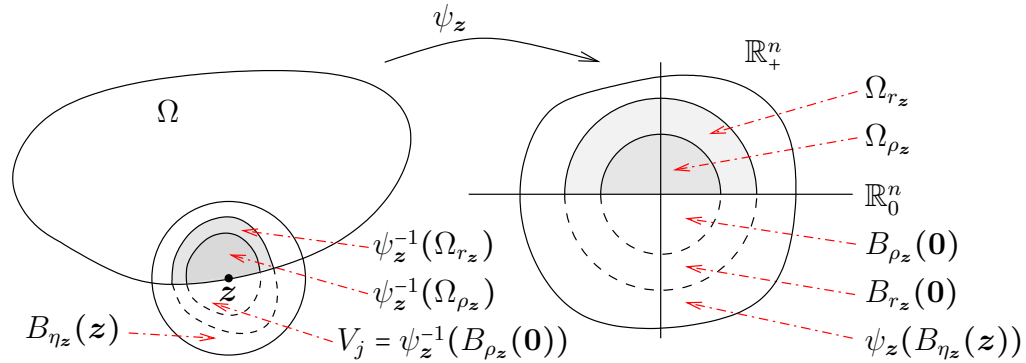


Figure 12.2: Representation local of the boundary of Ω as the image of the hyperplane $\mathbb{R}_0^n$ near the origin.

Proof (Theorem 12.6.18).

Because of the smoothness of the boundary of Ω , for each $\mathbf{z} \in \partial\Omega$, there exist an open ball $B_{\eta_z}(\mathbf{z})$ and a C^{q+2} -diffeomorphisms $\psi_z : B_{\eta_z}(\mathbf{z}) \rightarrow \mathbb{R}^n$ such that

1. $\psi_z(\mathbf{z}) = \mathbf{0}$,
2. $\psi_z(B_{\eta_z}(\mathbf{z}) \cap \Omega) \subset \mathbb{R}_+^n$, and
3. $\psi_z(B_{\eta_z}(\mathbf{z}) \cap \partial\Omega) \subset \mathbb{R}_0^n = \{\mathbf{x} \in \mathbb{R}^n : x_n = 0\}$.

Choose r_z small enough such that $\Omega_{r_z} = B_{r_z}(\mathbf{0}) \cap \mathbb{R}_+^n \subset \psi_z(B_{\eta_z}(\mathbf{z}) \cap \Omega)$, and $\rho_z < r_z$ (Figure 12.2).

Since $\overline{\partial\Omega}$ is compact because $\partial\Omega$ is bounded, we can cover $\partial\Omega$ with a finite number of open sets of the form $\psi_z^{-1}(B_{\rho_z}(\mathbf{0}))$ as described above. Let $V_1, V_2, \dots, V_s$ be those open sets.

Since $K = \overline{\Omega} \setminus \bigcup_{j=1}^s V_j$ is a compact subset of Ω , there exist an open set V_0 such that $K \subset V_0 \subset \overline{V_0} \subset \Omega$. Then, $\{V_j\}_{j=0}^s$ is an open cover of Ω .

i) We have that $u \in H^{q+2,2}(V_0)$ from Theorem 12.6.16. To get the inequality

$$\|u\|_{q+2,2,V_0} \leq C_0 (\|f\|_{q,2,\Omega} + \|u\|_{2,\Omega})$$

for some constant C_0 , we choose a sequence $\{W_j\}_{j=0}^{q+1}$ of open subsets of Ω such that

$$V_0 = W_0 \subset \overline{W_0} \subset W_1 \subset \overline{W_1} \subset W_2 \subset \dots \subset \overline{W_q} \subset W_{q+1} = \Omega .$$

For each $0 \leq j \leq q$, choose $\psi_j \in \mathcal{D}(\Omega)$ such that $\text{supp } \psi_j \subset W_{j+1}$ and $\psi_j = 1$ on W_j . From Corollary 12.6.13 with $t = 0$, there exists for each $0 \leq j \leq q$ a constant $C_{1,j}$ such that

$$\begin{aligned} \|u\|_{q+2-j,2,W_j} &\leq \|\psi_j u\|_{q+2-j,2,W_{j+1}} \leq C_{1,j} \left(\|L(\mathbf{x}, D)(\psi_j u)\|_{q-j,2,W_{j+1}} + \|\psi_j u\|_{2,W_{j+1}} \right) \\ &\leq C_{1,j} \left(\|\psi_j L(\mathbf{x}, D)u\|_{q-j,2,W_{j+1}} + \|[L(\mathbf{x}, D), \psi_j]u\|_{q-j,2,W_{j+1}} + \|\psi_j u\|_{2,W_{j+1}} \right) . \end{aligned}$$

The reason why we could use Corollary 12.6.13 was that $\text{supp}(\psi_j u) \subset W_{j+1}$ and so $\psi_j u \in H_0^{q+2-j,2}(W_{j+1})$ according to Lemma 11.2.3.

We have from Lemma 12.6.21 that

$$\|[L(\mathbf{x}, D), \psi_j]u\|_{q-j,2,W_{j+1}} \leq C_{2,j} \|u\|_{q+1-j,2,W_{j+1}}$$

for some constant $C_{2,j}$. We end up with

$$\begin{aligned} \|u\|_{q+2-j,2,W_j} &\leq C_{1,j} \left(\|L(\mathbf{x}, D)u\|_{q-j,2,W_{j+1}} + C_{2,j} \|u\|_{q+1-j,2,W_{j+1}} + \|u\|_{2,W_{j+1}} \right) \\ &\leq C_{1,j} \left(\|f\|_{q,2,\Omega} + C_{2,j} \|u\|_{q+1-j,2,W_{j+1}} + \|u\|_{2,\Omega} \right) \\ &\leq C_{1,j} \|f\|_{q,2,\Omega} + C_{1,j} C_{2,j} \|u\|_{q+1-j,2,W_{j+1}} + C_{1,j} \|u\|_{2,\Omega} \end{aligned} \quad (12.6.44)$$

for $0 \leq j \leq q$.

We prove by induction that, for each $0 \leq j \leq q$, there exists a constant K_j such

$$\|u\|_{q+2-j,2,W_j} \leq K_j (\|f\|_{q,2,\Omega} + \|u\|_{1,2,\Omega}) \quad (12.6.45)$$

The result is true for $j = q$ since (12.6.44) with $j = q$ yields

$$\begin{aligned} \|u\|_{2,2,W_q} &\leq C_{1,q} \|f\|_{q,2,\Omega} + C_{1,q} C_{2,q} \|u\|_{1,2,W_{q+1}} + C_{1,q} \|u\|_{2,\Omega} \\ &\leq C_{1,q} \|f\|_{q,2,\Omega} + (C_{1,q} C_{2,q} + C_{1,q}) \|u\|_{1,2,\Omega} \leq K_q (\|f\|_{q,2,\Omega} + \|u\|_{1,2,\Omega}) , \end{aligned}$$

where $K_q = C_{1,q} C_{2,q} + C_{1,q}$. Suppose that (12.6.45) is true for j between 1 and q . From (12.6.44) with j replaced by $j-1$, we get

$$\begin{aligned} \|u\|_{q+3-j,2,W_{j-1}} &\leq C_{1,j-1} \|f\|_{q,2,\Omega} + C_{1,j-1} C_{2,j-1} \|u\|_{q+2-j,2,W_j} + C_{1,j-1} \|u\|_{2,\Omega} \\ &\leq C_{1,j-1} \|f\|_{q,2,\Omega} + C_{1,j-1} C_{2,j-1} K_j (\|f\|_{q,2,\Omega} + \|u\|_{1,2,\Omega}) + C_{1,j-1} \|u\|_{2,\Omega} \\ &\leq (C_{1,j-1} + C_{1,j-1} C_{2,j-1} K_j) \|f\|_{q,2,\Omega} + (C_{1,j-1} C_{2,j-1} K_j + C_{1,j-1}) \|u\|_{1,2,\Omega} , \end{aligned}$$

where we have used the hypothesis of induction for the second inequality. If $K_{j-1} = C_{1,j-1} + C_{1,j-1} C_{2,j-1} K_j$, we get (12.6.45) for i replaced by $i-1$. This complete the proof by induction. Hence, with $j = 0$ in (12.6.45), we get

$$\|u\|_{q+2,2,V_0} = \|u\|_{q+2,2,W_0} \leq K_0 (\|f\|_{q,2,\Omega} + \|u\|_{1,2,\Omega}) . \quad (12.6.46)$$

ii) Suppose that $V_j = \psi_z^{-1}(B_{\rho_z}(\mathbf{0}))$ for some $\mathbf{z} \in \partial\Omega$ as described at the beginning of the proof and where $0 < j \leq s$. For this step, we sketch the reasoning and leave the many details to the reader. According to Proposition 11.4.2, we may use the isomorphism

$$\begin{aligned} T_{\mathbf{z}} : H^{q,2}(\Omega_{r_{\mathbf{z}}}) &\rightarrow H^{q,2}(\psi_z^{-1}(\Omega_{r_{\mathbf{z}}})) \\ g &\mapsto g \circ \psi_z \end{aligned}$$

to express (12.6.34) locally as the variational problem

$$\tilde{B}(\tilde{v}, \tilde{u}) = B(T_{\mathbf{z}}(\tilde{v}), T_{\mathbf{z}}(\tilde{u})) = \int_{\Omega_{r_{\mathbf{z}}}} \overline{f(\psi^{-1}(\mathbf{x}))} \tilde{v}(\mathbf{x}) |\det D \psi_z^{-1}(\mathbf{x})| d\mathbf{x} ,$$

where $\tilde{B}$ is still weakly coercive and $\tilde{v} \in X_r = \{\tilde{v} \in X : \text{supp } u \subset B_\rho(0) \cap \mathbb{R}_+^n \text{ for some } \rho < r_{\mathbf{z}}\}$ with $X = H^{q,2}(\mathbb{R}_+^n)$ or $X = H_0^{q,2}(\mathbb{R}_+^n)$. We can then conclude from Proposition 12.6.20 that $\tilde{u} \in H^{q+2,2}(\Omega_{\rho_{\mathbf{z}}})$ and there exists a constant $\tilde{K}_j = C(j, q, \rho_{\mathbf{z}})$ such that

$$\|\tilde{u}\|_{q+2,2,\Omega_{\rho_{\mathbf{z}}}} \leq \tilde{K}_j (\|f \circ \psi^{-1}\|_{q,2,\Omega_{r_{\mathbf{z}}}} + \|\tilde{u}\|_{1,2,\Omega_{r_{\mathbf{z}}}}) .$$

We then have that $u = T_{\mathbf{z}}(\tilde{u}) \in H^{q+2,2}(\psi_z^{-1}(\Omega_{\rho_{\mathbf{z}}})) = H^{q+2,2}(V_j \cap \Omega)$.

Since $T_{\mathbf{z}}$ and $T_{\mathbf{z}}^{-1}$ are bounded mappings, there exist constants C_1 and C_2 such that $\|T_{\mathbf{z}}(\tilde{v})\|_{q,2,\psi_z^{-1}(\Omega_{r_{\mathbf{z}}})} \leq C_1 \|\tilde{v}\|_{q,2,\Omega_{r_{\mathbf{z}}}}$ for $\tilde{v} \in H^{q,2}(\Omega_{r_{\mathbf{z}}})$ and $\|T_{\mathbf{z}}^{-1}(v)\|_{q,2,\Omega_{r_{\mathbf{z}}}} \leq C_2 \|v\|_{q,2,\psi_z^{-1}(\Omega_{r_{\mathbf{z}}})}$ and for $v \in H^{q,2}(\psi_z^{-1}(\Omega_{r_{\mathbf{z}}}))$. We then get that

$$\frac{1}{C_2} \|\tilde{v}\|_{q,2,\Omega_{r_{\mathbf{z}}}} \leq \|T_{\mathbf{z}}(\tilde{v})\|_{q,2,\psi_z^{-1}(\Omega_{r_{\mathbf{z}}})} \leq C_1 \|\tilde{v}\|_{q,2,\Omega_{r_{\mathbf{z}}}}$$

for $\tilde{v} \in H^{q,2}(\Omega_{r_{\mathbf{z}}})$. We can use this equivalence to find a constant K_j such that

$$\|u\|_{q+2,2,V_j \cap \Omega} = \|u\|_{q+2,2,\psi_z^{-1}(\Omega_{r_{\mathbf{z}}})} \leq K_j (\|f\|_{q,2,\psi_z^{-1}(\Omega_{r_{\mathbf{z}}})} + \|u\|_{1,2,\psi_z^{-1}(\Omega_{r_{\mathbf{z}}})}) \leq K_j (\|f\|_{q,2,\Omega} + \|u\|_{1,2,\Omega}) .$$

iii) Let $\{\phi_j\}_{j=1}^s$ be a partition of unity of Ω subordinated to $\{V_j\}_{j=0}^s$. We have that $u = \sum_{j=0}^s \phi_j u$, where $\text{supp}(\phi_j u) \subset V_j \cap \Omega$ for all j . Hence, from (i) and (ii), we have that $u \in H^{q+2,2}(\Omega)$ and

$$\begin{aligned} \|u\|_{q+2,2,\Omega} &\leq \sum_{j=0}^s \|\phi_j u\|_{q+2,2,\Omega} = \sum_{j=0}^s \|\phi_j u\|_{q+2,2,\Omega \cap V_j} \leq \sum_{j=0}^s C_j \|u\|_{q+2,2,\Omega \cap V_j} \\ &\leq \sum_{j=0}^s C_j K_j (\|f\|_{q,2,\Omega} + \|u\|_{1,2,\Omega}) = K (\|f\|_{q,2,\Omega} + \|u\|_{1,2,\Omega}) \end{aligned} \quad (12.6.47)$$

where $K = \sum_{j=0}^s C_j K_j$. We have used Proposition 11.1.17 to obtain the constants C_j .

iv) Finally, It follows from the definition of weakly coercive applied to the sesquilinear form B , that there exist two constants C_0 and λ such that

$$\|u\|_{1,2,\Omega}^2 \leq \frac{1}{C_0} (B(u, u) + \lambda \|u\|_{2,\Omega}^2) .$$

Recall that we prove the theorem only for second order elliptic operators, thus $k = 1$. Using $B(v, u) = \langle v, f \rangle_2$ for $v \in X$ and Schwarz inequality, we then get

$$\begin{aligned} \|u\|_{1,2,\Omega}^2 &\leq \frac{1}{C_0} \left(\int_{\Omega} f(\mathbf{x})u(\mathbf{x}) \, d\mathbf{x} + \lambda \|u\|_{2,\Omega}^2 \right) \leq \frac{1}{C_0} (\|f\|_{2,\Omega} \|u\|_{2,\Omega} + \lambda \|u\|_{2,\Omega}^2) \\ &= \frac{1}{C_0} \|u\|_2 (\|f\|_{2,\Omega} + \lambda \|u\|_{2,\Omega}) \leq \frac{1}{C_0} \|u\|_{1,2} (\|f\|_{2,\Omega} + \lambda \|u\|_{2,\Omega}) . \end{aligned}$$

Hence,

$$\|u\|_{1,2,\Omega} \leq \frac{1}{C_0} (\|f\|_{2,\Omega} + \lambda \|u\|_{2,\Omega}) \leq \frac{1}{C_0} (\|f\|_{q,2,\Omega} + \lambda \|u\|_{2,\Omega}) .$$

If we combine this inequality with (12.6.47), we get

$$\begin{aligned} \|u\|_{q+2,2,\Omega} &\leq K (\|f\|_{q,2,\Omega} + \|u\|_{1,2,\Omega}) \leq K \left(\|f\|_{q,2,\Omega} + \frac{1}{C_0} (\|f\|_{q,2,\Omega} + \lambda \|u\|_{2,\Omega}) \right) \\ &= K \left(1 + \frac{1}{C_0} \right) \|f\|_{q,2,\Omega} + K\lambda \|u\|_{2,\Omega} \leq C (\|f\|_{q,2,\Omega} + \|u\|_{2,\Omega}) , \end{aligned}$$

where $C = \max \left\{ K\lambda, K \left(1 + \frac{1}{C_0} \right) \right\}$. This complete the proof of Theorem 12.6.18. ■

Corollary 12.6.22

In Theorem 12.6.18, if $f \in C^\infty(\overline{\Omega})$ and $\partial\Omega$ is a C^∞ -manifold, then $u \in C^\infty(\overline{\Omega})$.

Proof.

Since $f \in H^{k,2}(\Omega)$ for all k because all the derivatives of f are bounded on $\overline{\Omega}$, we get from Theorem 12.6.18 that $u \in H^{k,2}(\Omega)$ for all k . It then follows from Corollary 11.7.9 that $u \in C^k(\overline{\Omega})$ for all k . ■

It follows from this corollary that eigenvectors of the strongly elliptic operator $L(\mathbf{x}, D)$ are of class C^∞ on $\overline{\Omega}$.

Corollary 12.6.23

Referring to Theorem 12.6.18, if $u \in X$ is an eigenvector of B associated to the eigenvalue λ (i.e. $B(v, u) = \langle v, \lambda u \rangle$ for all $v \in X$), then $u \in C^\infty(\overline{\Omega})$.

Proof.

The conclusion of this corollary is a consequence of the previous corollary applied to the variational problem $B_1(v, u) = B(v, u) - \bar{\lambda} \langle v, u \rangle = 0$ for $u, v \in X$. ■

12.7 Existence of Green's Functions

In this section, we finally prove Theorems 9.5.3 and 9.5.7.

Proposition 12.7.1

Let Ω be a bounded open subset of $\mathbb{R}^n$ with a boundary $\partial\Omega$ of class C^∞ . If $f \in C^\infty(\overline{\Omega})$, the solution of the Dirichlet problem $\Delta u = 0$ on Ω with $u = f$ on $\partial\Omega$ is in $C^\infty(\overline{\Omega})$.

Proof.

The existence and uniqueness of the solution has been proved in Theorem 9.9.40. Namely, there exists a unique $u \in C(\overline{\Omega})$ such that $\Delta u = 0$ on Ω in the sense of distributions and $u|_{\partial\Omega} = f$.

Let $w \in H_0^{1,2}(\Omega)$ be the weak solution of $-\Delta w = \Delta f$ on Ω with $w = 0$ on $\partial\Omega$. This solution exists and is unique according to Theorem 12.5.1 because the variational formulation of the Dirichlet problem $-\Delta w = \Delta f$ on Ω with $w = 0$ on $\partial\Omega$ is

$$B(w, v) = \int_{\Omega} \nabla w(\mathbf{x}) \cdot \nabla v(\mathbf{x}) \, d\mathbf{x} = \int_{\Omega} v(\mathbf{x}) \Delta f(\mathbf{x}) \, d\mathbf{x}$$

for all $v \in H_0^{1,2}(\Omega)$, where B is a bounded, symmetric and coercive bilinear form on $H_0^{1,2}(\Omega)$. From Corollary 12.6.22, we have that $w \in C^\infty(\overline{\Omega})$. Moreover, $w = 0$ on $\partial\Omega$ according to Proposition 11.8.14.

Since $\Delta(w + f) = 0$ on Ω and $w + f = f$ on $\partial\Omega$, we have by uniqueness of solutions that $u = w + f \in C^\infty(\overline{\Omega})$. ■

Proof of Theorem 9.5.3.

Let N be the fundamental solution of Δ given in Theorem 9.3.1.

i) For each $\mathbf{x} \in \Omega$, choose $\phi \in \mathcal{D}(\mathbb{R}^n)$ such that $\phi(\mathbf{x}) = 0$ and $\phi(\mathbf{y}) = 1$ for all $\mathbf{y} \in \partial\Omega$. Let

$$M(\mathbf{y}) = \begin{cases} \phi(\mathbf{y})N(\mathbf{x} - \mathbf{y}) & \text{if } \mathbf{y} \in \mathbb{R}^n \setminus \{\mathbf{x}\} \\ 0 & \text{if } \mathbf{y} = \mathbf{x} \end{cases}$$

Let $u_{\mathbf{x}}$ be the solution of the Dirichlet problem $\Delta_{\mathbf{y}} u = 0$ on Ω with $u = M$ on $\partial\Omega$. According to Proposition 12.7.1, $u_{\mathbf{x}} \in C^\infty(\overline{\Omega})$ since $M \in C^\infty(\overline{\Omega})$.

We have shown that, for each $\mathbf{x} \in \Omega$, the solution $u_{\mathbf{x}}$ of $\Delta_{\mathbf{y}} u = 0$ on Ω with $u(\mathbf{y}) = N(\mathbf{x} - \mathbf{y})$ for $\mathbf{y} \in \partial\Omega$ is $C^\infty(\overline{\Omega})$.

ii) The Green function for Ω is $G(\mathbf{x}, \mathbf{y}) = N(\mathbf{x} - \mathbf{y}) - u_{\mathbf{x}}(\mathbf{y})$ for $\mathbf{x} \in \Omega$ and $\mathbf{y} \in \overline{\Omega}$. It is clear that $\mathbf{y} \mapsto G(\mathbf{x}, \mathbf{y}) - N(\mathbf{x} - \mathbf{y}) = u_{\mathbf{x}}(\mathbf{y})$ is harmonic on Ω and continuous on $\overline{\Omega}$; it is in fact in $C^\infty(\overline{\Omega})$. Moreover, $G(\mathbf{x}, \mathbf{y}) = N(\mathbf{x} - \mathbf{y}) - u_{\mathbf{x}}(\mathbf{y}) = 0$ for $\mathbf{y} \in \partial\Omega$ by construction. ■

Proof of Theorem 9.5.7.

According to Theorem 9.9.40, there exists a unique solution $u \in C(\overline{\Omega})$ of the Dirichlet problem $\Delta u = 0$ on Ω in the sense of distributions with $u = g$ on $\partial\Omega$. Moreover, from Corollary 12.6.17, $u \in C^\infty(\Omega)$.

We only have to show that

$$u(\mathbf{x}) = w(\mathbf{x}) \equiv \int_{\partial\Omega} g(\mathbf{y}) \frac{\partial G_{\mathbf{x}}}{\partial \nu}(\mathbf{y}) \, dS_{\mathbf{y}}$$

for all $\mathbf{x} \in \Omega$.

From Weierstrass approximation theorem, we can construct a sequence $\{g_j\}_{j=1}^\infty$ of polynomials such that $\sup_{\mathbf{x} \in \partial\Omega} |g_j(\mathbf{x}) - g(\mathbf{x})| \leq 1/j$ for all j . Let u_j be the solution of the Dirichlet problem $\Delta u = 0$ in Ω with $u = g_j$ on $\partial\Omega$.

From Proposition 12.7.1, we have that $u_j \in C^\infty(\bar{\Omega})$. Hence, as it was proved after Theorem 9.5.7, we have that

$$u_j(\mathbf{x}) = \int_{\partial\Omega} u_j(\mathbf{y}) \frac{\partial G_{\mathbf{x}}}{\partial \nu}(\mathbf{y}) dS_{\mathbf{y}} = \int_{\partial\Omega} g_j(\mathbf{y}) \frac{\partial G_{\mathbf{x}}}{\partial \nu}(\mathbf{y}) dS_{\mathbf{y}} \quad (12.7.1)$$

for all $\mathbf{x} \in \Omega$. If $j \rightarrow \infty$, the right side of (12.7.1) converges to

$$\int_{\partial\Omega} g(\mathbf{y}) \frac{\partial G_{\mathbf{x}}}{\partial \nu}(\mathbf{y}) dS_{\mathbf{y}}$$

because $g_j(\mathbf{y}) \frac{\partial G_{\mathbf{x}}}{\partial \nu}(\mathbf{y})$ converges uniformly to $g(\mathbf{y}) \frac{\partial G_{\mathbf{x}}}{\partial \nu}(\mathbf{y})$ on $\partial\Omega$. For the left side of (12.7.1), we apply the maximum principle, Corollary 9.2.5, to $u_j - u$ to conclude that

$$\sup_{\mathbf{x} \in \bar{\Omega}} |u_j(\mathbf{x}) - u(\mathbf{x})| \leq \sup_{\mathbf{x} \in \partial\Omega} |g_j(\mathbf{x}) - g(\mathbf{x})| \rightarrow 0 \quad \text{as } j \rightarrow \infty.$$

Thus, the left hand side of (12.7.1) converges to $u(\mathbf{x})$. This prove the Theorem. ■

12.8 Maximum Principle

We have already seen in Theorem 9.2.3 a maximum principle for the Laplace equation. Our goal in this section is to generalize if possible this theorem to a general family of elliptic partial differential equations. As we will see, the maximum principles that we obtain will not always be as strong as the maximum principle for the Laplace equation.

We begin by presenting the strong maximum principle for a class of elliptic partial differential equations on a domain Ω that admit solutions in $C^2(\Omega) \cap C(\bar{\Omega})$. After that, we present a slightly weaker maximum principle that applies to a more general class of strongly elliptic partial differential equations. Though weaker, this principle is still very useful.

A good reference about maximum principles for partial differential equations is [25]. We follow the presentation in [35] for the strong maximum principle.

12.8.1 Strong Maximum Principle

All the theorems in this subsection form what is usually called the strong maximum principle. Several equivalent formulations are presented in the literature. Be aware that most of the formulations of the strong maximum principle are formulated in terms of the standard form of the linear elliptic differential operator $L(\mathbf{x}, D)$, not the divergence form. Therefore, some

of the inequalities in the hypothesis of the next theorem may be in the other direction than usually found. It is sometime amazing what an extra minus sign can do.

Before proving any theorem, we need a little result from linear algebra.

Proposition 12.8.1

Suppose that A and B are two $n \times n$ matrices, and A is symmetric and **positive semi-definite**; namely, A is symmetric and $\mathbf{x}^\top A \mathbf{x} \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n$. If B is positive semi-definite, then $\text{tr}(AB) \geq 0$. If B is negative semi-definite, then $\text{tr}(AB) \leq 0$.

Proof.

We prove the result for B negative semi-definite and leave the proof of the other result to the reader.

Since A is symmetric, there exists an orthogonal matrix P such that $Q = P^\top A P$ is a diagonal matrix. Since A is positive semi-definite, the elements $q_{i,i}$ on the diagonal of Q (which are the eigenvalues of A) are greater or equal to zero.

Let $C = P^\top B P$. We have that C is negative semi-definite because $\mathbf{x}^\top C \mathbf{x} = (P \mathbf{x})^\top B (P \mathbf{x}) \leq 0$ for all $\mathbf{x} \in \mathbb{R}^n$. Therefore $c_{i,i} = \mathbf{e}_i^\top C \mathbf{e}_i \leq 0$ for $1 \leq i \leq n$.

Hence,

$$\text{tr}(AB) = \text{tr}(P^\top A B P) = \text{tr}((P^\top A P)(P^\top B P)) = \text{tr}(QC) = \sum_{i=1}^n q_{i,i} c_{i,i} \leq 0 .$$

The fact that $\text{tr}(AB) = \text{tr}(P^\top A B P)$ comes from the fact that the trace of a $n \times n$ matrix is the sum of its eigenvalues, and conjugated matrices have the same eigenvalues. ■

The first theorem is not the most general result but it is the crux of the proof of all the other results in this subsection.

Theorem 12.8.2

Let Ω be a connected open subset of $\mathbb{R}^n$ with a boundary $\partial\Omega$ of class C^2 . Suppose that

$$L(\mathbf{x}, D)u = - \sum_{i,j=1}^n D_{x_i} (b_{i,j}(\mathbf{x}) D_{x_j} u)$$

is an elliptic operator, where $b_{i,j} \in C^1(\overline{\Omega})$ for $1 \leq i, j \leq n$. Let $u \in C^2(\Omega) \cap C(\overline{\Omega})$ be the solution of $L(\mathbf{x}, D)u = f$, where $f \in C(\overline{\Omega})$.

1. If $f(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in \overline{\Omega}$ and $\min_{\mathbf{x} \in \overline{\Omega}} u(\mathbf{x}) = u(\mathbf{p})$ for some $\mathbf{p} \in \Omega$, or
2. if $f(\mathbf{x}) \leq 0$ for all $\mathbf{x} \in \overline{\Omega}$ and $\max_{\mathbf{x} \in \overline{\Omega}} u(\mathbf{x}) = u(\mathbf{p})$ for some $\mathbf{p} \in \Omega$,

then $u(\mathbf{x}) = u(\mathbf{p})$ for all $\mathbf{x} \in \Omega$. In particular, a non-constant function u may only reach its minimum in (1) or its maximum in (2) on $\partial\Omega$.

Proof.

We have that

$$L(\mathbf{x}, D)u = - \sum_{i,j=1}^n D_{x_i} (b_{i,j}(\mathbf{x}) D_{x_j} u) = - \sum_{i,j=1}^n b_{i,j}(\mathbf{x}) D_{x_i x_j} u - \sum_{j=1}^n b_j(\mathbf{x}) D_{x_j} u, \quad (12.8.1)$$

where $b_j = \sum_{i=1}^n D_{x_i} b_{i,j}$ for $1 \leq j \leq n$.

We prove that (2) implies that $u(\mathbf{x}) = u(\mathbf{p})$ for all $\mathbf{x} \in \Omega$ and leave the proof of the other implication to the reader. The proof of the other implication is similar.

Let $M = \{\mathbf{x} \in \Omega : u(\mathbf{x}) = u(\mathbf{p})\}$. If the conclusion is false, then M is a proper subset of Ω . Moreover, M is non-empty because $\mathbf{p} \in M$.

The discussion of the next two paragraphs is summarized in Figure 12.3.

Choose $\mathbf{q} \in \Omega \setminus M$. Since Ω is connected, we can select a continuous path $\gamma : [0, 1] \rightarrow \Omega$ from $\mathbf{q}$ to $\mathbf{p}$; in particular, $\gamma(0) = \mathbf{q}$ and $\gamma(1) = \mathbf{p}$. Since $\gamma([0, 1])$ is a compact subset of Ω , we have that

$$\delta = \text{dist}(\partial\Omega, \gamma([0, 1])) = \sup \{\|\mathbf{x} - \gamma(t)\| : \mathbf{x} \in \partial\Omega \text{ and } 0 \leq t \leq 1\} > 0.$$

Since $u(\mathbf{q}) < u(\mathbf{p})$, there exists $0 < \delta_1 < \delta/2$ such that $u(\mathbf{x}) < u(\mathbf{p})$ for all $\mathbf{x} \in B_{\delta_1}(\mathbf{q})$. Let t_1 be the smallest value of $t \in [0, 1]$ such that $M \cap \partial B_{\delta_1}(\gamma(t_1)) \neq \emptyset$. Let $\mathbf{q}_1 = \gamma(t_1)$ and $B_1 = B_{\delta_1}(\mathbf{q}_1)$. Then, $\overline{B_1} \subset \Omega$ with $B_1 \cap M = \emptyset$ and $M \cap \partial B_1 = \{\mathbf{q}_2\}$ for some $\mathbf{q}_2 \in \gamma([0, 1])$. Choose $\mathbf{r} \in B_1$ and δ_2 such that $B_{\delta_2}(\mathbf{r}) \subset B_1$ and $\partial B_{\delta_2}(\mathbf{r}) \cap M = \{\mathbf{q}_2\}$. Let $B_2 = B_{\delta_2}(\mathbf{r})$. Finally, choose $\delta_3 < \delta_2$ and set $B_3 = B_{\delta_3}(\mathbf{q}_2)$.

We have that $\partial B_3 = S_1 \cup S_2$, where $S_1 = \overline{B_2} \cap \partial B_3$ and $S_1 \cap S_2 = \emptyset$. Since $u \in C(\Omega)$, $u(\mathbf{x}) < u(\mathbf{q}_2) = u(\mathbf{p})$ for all $\mathbf{x} \in S_1$ and S_1 is a compact set, there exists $\epsilon > 0$ such that $u(\mathbf{x}) < u(\mathbf{p}) - \epsilon$ for all $\mathbf{x} \in S_1$.

Let

$$h(\mathbf{x}) = e^{-\alpha\|\mathbf{x}-\mathbf{r}\|^2} - e^{-\alpha\delta_2^2}, \quad (12.8.2)$$

where $\alpha > 0$ will be selected later. We have that

$$\begin{aligned} & e^{-\alpha\|\mathbf{x}-\mathbf{r}\|^2} L(\mathbf{x}, D)h(\mathbf{x}) \\ &= -4\alpha^2 \sum_{i,j=1}^n b_{i,j}(\mathbf{x})(x_i - r_i)(x_j - r_j) - 2\alpha \sum_{j=1}^n (b_{j,j}(\mathbf{x}) + b_j(\mathbf{x})(x_j - r_j)). \end{aligned} \quad (12.8.3)$$

Since $\mathbf{r} \notin \overline{B_3}$ because $\delta_3 < \delta_2$, we have that $\sum_{i,j=1}^n b_{i,j}(\mathbf{x})(x_i - r_i)(x_j - r_j) > 0$ for all $\mathbf{x} \in \overline{B_3}$ because $L(\mathbf{x}, D)$ is elliptic on Ω ⁶. It follows from (12.8.3) and the compactness of $\overline{B_3}$ that

$$L(\mathbf{x}, D)h(\mathbf{x}) < 0 \quad (12.8.4)$$

for all $\mathbf{x} \in \overline{B_3}$ if we choose α large enough. Suppose that this has been done.

⁶The reader should review the definition of elliptic partial differential operator that we have adopted for this chapter at the beginning of Section 12.2.

Let $m = \max_{\mathbf{x} \in S_1} h(\mathbf{x})$. Note that $m > 0$ because $S_1 \cap B_2 \neq \emptyset$. Choose $\eta < \epsilon/m$ and let $v = u + \eta h$. We get that

$$v(\mathbf{x}) = u(\mathbf{x}) + \eta h(\mathbf{x}) < u(\mathbf{p}) - \epsilon + \eta m < u(\mathbf{p})$$

for all $\mathbf{x} \in S_1$. Since $S_2 \cap \overline{B_2} = \emptyset$, we have that $\|\mathbf{x} - \mathbf{r}\| > \delta_2$ for all $\mathbf{x} \in S_2$. Thus, $h(\mathbf{x}) < 0$ for all $\mathbf{x} \in S_2$ and we get that

$$v(\mathbf{x}) = u(\mathbf{x}) + \eta h(\mathbf{x}) < u(\mathbf{x}) \leq u(\mathbf{p})$$

for all $\mathbf{x} \in S_2$. Since $v(\mathbf{x}) < u(\mathbf{p})$ for $\mathbf{x} \in \partial B_3 = S_1 \cap S_2$ and $v(\mathbf{q}_2) = u(\mathbf{q}_2) + \eta h(\mathbf{q}_2) = u(\mathbf{p})$, and since this is true for any positive δ_3 less than δ_2 , we find that v has a local maximum at $\mathbf{q}_2$. Therefore, the gradient of v is null and its Hessian matrix is negative semi-definite at $\mathbf{q}_2$; namely, $\nabla v(\mathbf{q}_2) = \mathbf{0}$ and $\sum_{i,j=1}^n D_{x_i x_j} v(\mathbf{q}_2) y_i y_j \leq 0$ for all $\mathbf{y} \in \mathbb{R}^n$. Moreover, since $L(\mathbf{x}, D)$ is elliptic on Ω , we also get $\sum_{i,j=1}^n b_{i,j}(\mathbf{q}_2) y_i y_j \geq 0$ for all $\mathbf{y} \in \mathbb{R}^n$.

Let A and $B^\top$ be the matrices with entries $b_{i,j}(\mathbf{q}_2)$ and $-D_{x_i x_j} v(\mathbf{q}_2)$ respectively. It follows from the previous proposition that

$$\begin{aligned} 0 \leq \text{tr}(AB) &= - \sum_{i,j=1}^n b_{i,j}(\mathbf{q}_2) D_{x_i x_j} v(\mathbf{q}_2) = L(\mathbf{x}, D)v(\mathbf{x})|_{\mathbf{x}=\mathbf{q}_2} \\ &= L(\mathbf{x}, D)u(\mathbf{x})|_{\mathbf{x}=\mathbf{q}_2} + \eta L(\mathbf{x}, D)h(\mathbf{x})|_{\mathbf{x}=\mathbf{q}_2} < L(\mathbf{x}, D)u(\mathbf{x})|_{\mathbf{x}=\mathbf{q}_2}, \end{aligned}$$

where the second equality comes from $\nabla v(\mathbf{q}_2) = \mathbf{0}$ and the last inequality comes from (12.8.4). This is a contradiction of our original assumption that $L(\mathbf{x}, D)u(\mathbf{x}) = f(\mathbf{x}) \leq 0$ for all $\mathbf{x} \in \Omega$. $\blacksquare$

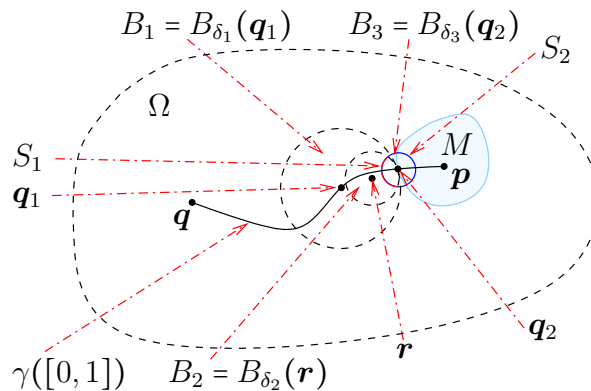


Figure 12.3: Construction for the proof of Theorem 12.8.2

We now generalize the previous theorem. However, the conclusion will not be as strong as for the previous theorem.

Theorem 12.8.3

Let Ω be a connected open subset of $\mathbb{R}^n$ with a boundary $\partial\Omega$ of class C^2 . Suppose that

$$L(\mathbf{x}, D)u = - \sum_{i,j=1}^n D_{x_i} (b_{i,j}(\mathbf{x}) D_{x_j} u) + b_0(\mathbf{x})u$$

be an elliptic operator, where $b_{i,j} \in C^1(\overline{\Omega})$ for $1 \leq i, j \leq n$ and $b_0 \in C(\overline{\Omega})$. Moreover, we assume that $b_0(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in \overline{\Omega}$. Let $u \in C^2(\Omega) \cap C(\overline{\Omega})$ be the solution of $L(\mathbf{x}, D)u = f$, where $f \in C(\overline{\Omega})$.

1. If $f(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in \overline{\Omega}$ and $\min_{\mathbf{x} \in \overline{\Omega}} u(\mathbf{x}) = u(\mathbf{p}) < 0$ for some $\mathbf{p} \in \Omega$, or
2. if $f(\mathbf{x}) \leq 0$ for all $\mathbf{x} \in \overline{\Omega}$ and $\max_{\mathbf{x} \in \overline{\Omega}} u(\mathbf{x}) = u(\mathbf{p}) > 0$ for some $\mathbf{p} \in \Omega$,

then $u(\mathbf{x}) = u(\mathbf{p})$ for all $\mathbf{x} \in \Omega$. In particular, u may only reach its minimum in (1) or its maximum in (2) on $\partial\Omega$.

Proof.

Let

$$H(\mathbf{x}, D)u = - \sum_{i,j=1}^n D_{x_i} (b_{i,j}(\mathbf{x}) D_{x_j} u) = - \sum_{i,j=1}^n b_{i,j}(\mathbf{x}) D_{x_i x_j} u - \sum_{j=1}^n b_j(\mathbf{x}) D_{x_j} u, \quad (12.8.5)$$

where $b_j = \sum_{i=1}^n D_{x_i} b_{i,j}$ for $1 \leq j \leq n$.

We prove that (1) implies that $u(\mathbf{x}) = u(\mathbf{p})$ for all $\mathbf{x} \in \Omega$ and leave the proof of the other implication to the reader.

Let $M = \{\mathbf{x} \in \Omega : u(\mathbf{x}) = u(\mathbf{p})\}$. If the conclusion is false, then M is a proper subset of Ω . Moreover, M is non-empty because $\mathbf{p} \in M$ and it is closed in Ω because $u \in C(\overline{\Omega})$.

Given $\mathbf{y} \in M$, since $u(\mathbf{y}) = u(\mathbf{p}) < 0$ and u is continuous on Ω , there exists $\rho > 0$ such $B_\rho(\mathbf{y}) \subset \Omega$ and $0 > u(\mathbf{x}) \geq u(\mathbf{p})$ for all $\mathbf{x} \in B_\rho(\mathbf{y})$. Hence, $H(\mathbf{x}, D)u(\mathbf{x}) = -b_0(\mathbf{x})u(\mathbf{x}) + f(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in B_\rho(\mathbf{y})$. We get from (1) of Theorem 12.8.2 (with $H(\mathbf{x}, D)$ instead of $L(\mathbf{x}, D)$ and $B_\rho(\mathbf{y})$ instead of Ω) that $u(\mathbf{x}) = u(\mathbf{p})$ for all $\mathbf{x} \in B_\rho(\mathbf{y})$; namely, $B_\rho(\mathbf{y}) \subset M$. Thus $M \neq \emptyset$ is open and close in Ω . Since Ω is connected, we get that $M = \Omega$. ■

As expected, Theorem 12.8.3 can be used to prove the uniqueness of the solution for an elliptic partial differential equation.

Corollary 12.8.4

Consider Ω , $L(\mathbf{x}, D)$ and f defined in Theorem 12.8.3. If $u_1, u_2 \in C^2(\Omega) \cap C(\overline{\Omega})$ satisfy $L(\mathbf{x}, D)u = f$ on Ω and $u_1 = u_2$ on $\partial\Omega$, then $u_1 = u_2$ on Ω .

Proof.

Since $u = u_1 - u_2 \in C^2(\Omega) \cap C(\overline{\Omega})$ is a solution of $L(\mathbf{x}, D)u = 0$ on Ω with $u = 0$ on $\partial\Omega$, a simple case elimination using Theorem 12.8.3 yields $u = 0$. ■

The next theorems provide useful information for the Neumann problem.

Theorem 12.8.5

Consider Ω , $L(\mathbf{x}, D)$ and f defined in Theorem 12.8.2. Suppose that $u \in C^2(\Omega) \cap C(\overline{\Omega})$ is a non-constant solution of $L(\mathbf{x}, D)u = f$ on Ω . Let $\mu(\mathbf{x})$ be any outward-pointing unit vector at $\mathbf{x} \in \partial\Omega$.

1. If $f(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in \overline{\Omega}$ and u reaches its minimum at $\mathbf{p} \in \partial\Omega$, then $\frac{\partial u}{\partial \mu}(\mathbf{p}) < 0$.
2. If $f(\mathbf{x}) \leq 0$ for all $\mathbf{x} \in \overline{\Omega}$ and u reaches its maximum at $\mathbf{p} \in \partial\Omega$, then $\frac{\partial u}{\partial \mu}(\mathbf{p}) > 0$.

Proof.

There is a lot of similarity between the proof of this theorem and the proof of the Theorem 12.8.2.

We proof (2) and leave the proof (1) to the reader. Since u is not constant, we get from (2) of Theorem 12.8.2 that $u(\mathbf{x}) < u(\mathbf{p})$ for all $\mathbf{x} \in \Omega$.

Since $\partial\Omega$ is of class C^2 , we can find $\mathbf{r} \in \Omega$ and $\delta_1 > 0$ such that $B_{\delta_1}(\mathbf{r}) \subset \Omega$ and $\partial B_{\delta_1}(\mathbf{r}) \cap \partial\Omega = \{\mathbf{p}\}$ (Figure 12.4). Let $B_1 = B_{\delta_1}(\mathbf{r})$. Choose $\delta_2 < \delta_1$ and set $B_2 = B_{\delta_2}(\mathbf{p})$.

As in the proof of Theorem 12.8.2, we consider h defined in (12.8.2) and select α large enough to have

$$L(\mathbf{x}, D)h(\mathbf{x}) < 0 \quad (12.8.6)$$

for all $\mathbf{x} \in \overline{B_2}$.

Let $S_1 = \overline{B_1} \cap \partial B_2$, and set $m_1 = \max_{\mathbf{x} \in S_1} h(\mathbf{x})$ and $m_2 = \max_{\mathbf{x} \in S_1} u(\mathbf{x})$. Since h and u are continuous on the compact set S_1 , these maximums are reached at some points of S_1 . Since $u(\mathbf{x}) < u(\mathbf{p})$ for all $\mathbf{x} \in S_1 \subset B_1 \setminus \{\mathbf{p}\}$, we have that $m_2 < u(\mathbf{p})$. Moreover, $m_1 > 0$ because $S_1 \cap B_1 \neq \emptyset$. Choose $0 < \eta < (u(\mathbf{p}) - m_2)/m_1$ and let $v = u + \eta h$. Then, $v(\mathbf{x}) = u(\mathbf{x}) + \eta h(\mathbf{x}) < u(\mathbf{p}) = v(\mathbf{p})$ for all $\mathbf{x} \in S_1$.

Let $S_2 = B_2 \cap \partial B_1$. Since $h(\mathbf{x}) = 0$ for all $\mathbf{x} \in S_2$, we have that $v(\mathbf{x}) = u(\mathbf{x}) < u(\mathbf{p}) = v(\mathbf{p})$ for all $\mathbf{x} \in S_2 \setminus \{\mathbf{p}\}$ because $u(\mathbf{x}) < u(\mathbf{p})$ for all $\mathbf{x} \in \Omega$.

We have shown that $v(\mathbf{x}) \leq v(\mathbf{p})$ for all $\mathbf{x} \in \partial(B_1 \cap B_2) = S_1 \cup S_2$. Moreover, we get from (12.8.6) that

$$L(\mathbf{x}, D)v(\mathbf{x}) = L(\mathbf{x}, D)u(\mathbf{x}) + \eta L(\mathbf{x}, D)h(\mathbf{x}) < 0$$

for all $\mathbf{x} \in B_1 \cap B_2$. Hence, we get from (2) of Theorem 12.8.2 (with $B_1 \cap B_2$ instead of Ω) that v has an isolate maximum at $\mathbf{p}$ because v is not constant on $B_1 \cap B_2$. Therefore

$$0 \leq \frac{\partial v}{\partial \mu}(\mathbf{p}) = \frac{\partial u}{\partial \mu}(\mathbf{p}) + \eta \frac{\partial h}{\partial \mu}(\mathbf{p}). \quad (12.8.7)$$

However, if $\mu(\mathbf{p}) = \nu(\mathbf{p}) = \|\mathbf{p} - \mathbf{r}\|^{-1}(\mathbf{p} - \mathbf{r})$, the outward unit normal to $\partial(B_1 \cap B_2)$ at $\mathbf{p}$, then

$$\begin{aligned} \frac{\partial h}{\partial \mu}(\mathbf{p}) &= \nabla h(\mathbf{p}) \cdot \nu(\mathbf{p}) = \sum_{i=1}^n (-2\alpha(p_i - r_i) e^{-\alpha\|\mathbf{p}-\mathbf{r}\|^2}) \left(\frac{p_i - r_i}{\|\mathbf{p} - \mathbf{r}\|} \right) \\ &= \sum_{i=1}^n \left(-2\alpha \frac{(p_i - r_i)^2}{\|\mathbf{p} - \mathbf{r}\|} e^{-\alpha\|\mathbf{p}-\mathbf{r}\|^2} \right) = -2\alpha \|\mathbf{p} - \mathbf{r}\| e^{-\alpha\|\mathbf{p}-\mathbf{r}\|^2} < 0 . \end{aligned}$$

Therefore, we get $0 < \frac{\partial u}{\partial \mu}(\mathbf{p})$ from (12.8.7).

If $\mu(\mathbf{p})$ is any outward-pointing unit vector to Ω at $\mathbf{p} \in \partial\Omega$, then $\mu(\mathbf{p}) \cdot \nu(\mathbf{p}) > 0$. Hence, $\text{sgn } \nabla h(\mathbf{p}) \cdot \mu(\mathbf{p}) = \text{sgn } \nabla h(\mathbf{p}) \cdot \nu(\mathbf{p}) > 0$. ■

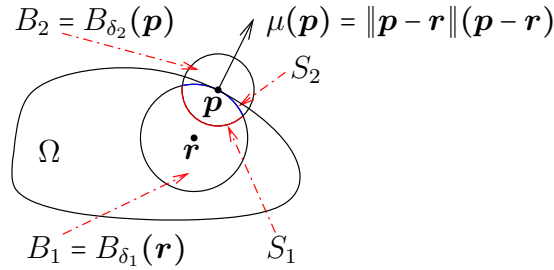


Figure 12.4: Construction for the proof of Theorem 12.8.5

Theorem 12.8.6

Consider Ω , $L(x, D)$ and f defined in Theorem 12.8.3. Suppose that $u \in C^2(\Omega) \cap C(\overline{\Omega})$ is a non-constant solution of $L(x, D)u = f$ on Ω . Let $\mu(x)$ be any outward-pointing unit vector at $x \in \partial\Omega$.

1. If $f(x) \geq 0$ for all $x \in \overline{\Omega}$ and u reaches its negative minimum at $\mathbf{p} \in \partial\Omega$, then $\frac{\partial u}{\partial \mu}(\mathbf{p}) < 0$.
2. If $f(x) \leq 0$ for all $x \in \overline{\Omega}$ and u reaches its positive maximum at $\mathbf{p} \in \partial\Omega$, then $\frac{\partial u}{\partial \mu}(\mathbf{p}) > 0$.

Proof.

We prove the case (2) and leave the proof of the other case to the reader.

Let

$$H(x, D)u = - \sum_{i,j=1}^n D_{x_i} (b_{i,j}(x) D_{x_j} u) = - \sum_{i,j=1}^n b_{i,j}(x) D_{x_i x_j} u - \sum_{j=1}^n b_j(x) D_{x_j} u ,$$

where $b_j = \sum_{i=1}^n D_{x_i} b_{i,j}$ for $1 \leq i \leq n$.

Since u is not constant, we get from (2) of Theorem 12.8.3 that $u(\mathbf{x}) < u(\mathbf{p})$ for all $\mathbf{x} \in \Omega$.

Since $\partial\Omega$ is of class C^2 , we can select a ball $B \subset \Omega$ such that $\overline{B} \cap \partial\Omega = \{\mathbf{p}\}$. In particular, ∂B is tangent to $\partial\Omega$ at $\mathbf{p}$.

Since $u(\mathbf{p}) > 0$ and $u \in C(\overline{\Omega})$, we may assume that B is small enough to have $u(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in B$. Hence,

$$H(\mathbf{x}, D)u(\mathbf{x}) = -b_0(\mathbf{x})u(\mathbf{x}) + f(\mathbf{x}) \leq 0$$

for all $\mathbf{x} \in B$. It follows from Theorem 12.8.5 that $\frac{\partial u}{\partial \mu}(\mathbf{p}) > 0$. ■

The fact that $\frac{\partial u}{\partial \nu}u(\mathbf{x}) \leq 0$ at a point $\mathbf{p} \in \partial\Omega$ where u reach a minimum or that $\frac{\partial u}{\partial \nu}u(\mathbf{x}) \geq 0$ at a point $\mathbf{p} \in \partial\Omega$ where u reach a maximum is not surprising and can be deduced from basic calculus. The strict inequality is more surprising.

12.8.2 General Maximum Principle

We have a maximum principle for the second order coercive elliptic partial differential equations if the coefficients are not smooth functions as we had in the previous subsection. We begin with maximum principle for the Dirichlet problem. As in the previous section, we consider in this section only real valued functions.

Our maximum principle for the Dirichlet problem will be deduced from the following proposition.

Proposition 12.8.7

Let Ω be an open subset of $\mathbb{R}^n$, Suppose that $f \in L^2(\Omega)$ and that $u \in H^{1,2}(\Omega) \cap C(\overline{\Omega})$ is a solution of

$$B(v, u) = \int_{\Omega} (\nabla u \cdot \nabla v + uv) \, d\mathbf{x} = \int_{\Omega} f v \, d\mathbf{x} \quad (12.8.8)$$

for all $v \in H_0^{1,2}(\Omega)$. Then

$$\min \left\{ \inf_{\mathbf{x} \in \partial\Omega} u(\mathbf{x}), \inf_{\mathbf{x} \in \Omega} f(\mathbf{x}) \right\} \leq u(\mathbf{x}) \leq \max \left\{ \sup_{\mathbf{x} \in \partial\Omega} u(\mathbf{x}), \sup_{\mathbf{x} \in \Omega} f(\mathbf{x}) \right\}$$

for all $\mathbf{x} \in \Omega$.

Proof.

i) We first prove the second inequality. Let

$$K = \max \left\{ \sup_{\mathbf{x} \in \partial\Omega} u(\mathbf{x}), \sup_{\mathbf{x} \in \Omega} f(\mathbf{x}) \right\} .$$

If $K = \infty$, we are done. Suppose that $K < \infty$. We choose a function $g \in C^1(\mathbb{R})$ such that $0 < g'(x) < 1$ for all $x \in]0, \infty[$ and $g(x) = 0$ for all $x \in]-\infty, 0]$. Let

$$\begin{aligned} v : \overline{\Omega} &\rightarrow \mathbb{R} \\ \mathbf{x} &\mapsto g(u(\mathbf{x}) - K) \end{aligned}$$

i.a) Suppose that Ω is bounded. Then $v \in L^2(\Omega) \cap L^1(\Omega)$ because v is continuous on the compact set $\overline{\Omega}$. Moreover, for any multi-index α with $|\alpha| = 1$, we have $D^\alpha v = g'(u - K) D^\alpha u$ in the sense of distributions with $|g'| \leq 1$ and $D^\alpha u \in L^2(\Omega)$ because $u \in H^{1,2}(\Omega)$. Thus, $D^\alpha v \in L^2(\Omega)$. Therefore, $v \in H^{1,2}(\Omega)$.

Since $v \in H^{1,2}(\Omega) \cap C(\overline{\Omega})$ and $v(\mathbf{x}) = g(u(\mathbf{x}) - K) = 0$ for $\mathbf{x} \in \partial\Omega$ by construction, we get from Proposition 11.8.14 that $v \in H_0^{1,2}(\Omega)$.

With this choice of v in (12.8.8), we get

$$\int_{\Omega} ((\nabla u \cdot \nabla u) g'(u - K) + u g(u - K)) \, d\mathbf{x} = \int_{\Omega} f g(u - K) \, d\mathbf{x} .$$

Since $v \in L^1(\Omega)$, we may subtract $K \int_{\Omega} g(u - K) \, d\mathbf{x}$ from each side of the previous equation to get

$$\int_{\Omega} (\|\nabla u\|_2^2 g'(u - K) + (u - K) g(u - K)) \, d\mathbf{x} = \int_{\Omega} (f - K) g(u - K) \, d\mathbf{x} .$$

Since $\|\nabla u\|_2^2 g'(u - K) \geq 0$ and $(f - K) g(u - K) \leq 0$ almost everywhere on Ω , we get

$$\int_{\Omega} (u - K) g(u - K) \, d\mathbf{x} \leq 0 .$$

However $(u(\mathbf{x}) - K) g(u(\mathbf{x}) - K) > 0$ for $\mathbf{x} \in \Omega$ such that $u(\mathbf{x}) > K$ and $(u(\mathbf{x}) - K) g(u(\mathbf{x}) - K) = 0$ for $\mathbf{x} \in \Omega$ such that $u(\mathbf{x}) \leq K$. Therefore, $(u(\mathbf{x}) - K) g(u(\mathbf{x}) - K) = 0$ for all $\mathbf{x} \in \Omega$ because the function defined by $\mathbf{x} \mapsto (u(\mathbf{x}) - K) g(u(\mathbf{x}) - K)$ is continuous on Ω . Thus $u(\mathbf{x}) \leq K$ for all $\mathbf{x} \in \Omega$.

i.b) If Ω is not bounded, we must have $K \geq 0$ because $f(\mathbf{x}) \leq K < 0$ for almost all $\mathbf{x}$ and Ω not bounded implies that $f \notin L^2(\Omega)$. Choose $K_0 > K$ and let

$$\begin{aligned} v : \overline{\Omega} &\rightarrow \mathbb{R} \\ \mathbf{x} &\mapsto g(u(\mathbf{x}) - K_0) \end{aligned}$$

for g defined above. To prove that $v \in H_0^{1,2}(\Omega)$, we note that for every $x \in \mathbb{R}$, there exists y between $x - K_0$ and $-K_0$ such that

$$|g(x - K_0)| = |g(x - K_0) - g(-K_0)| = |g'(y)| |x| \leq |x| .$$

Hence $|g(u - K_0)| \leq |u|$. Since $u \in L^2(\Omega)$, we have that $v = g(u - K_0) \in L^2(\Omega)$. The proof that $D^\alpha v \in L^2(\Omega)$ for the multi-index α such that $|\alpha| = 1$ is identical to the proof given in (i.a). As in (i.a), $v \in H^{1,2}(\Omega) \cap C(\overline{\Omega})$ and $v(\mathbf{x}) = g(u(\mathbf{x}) - K_0) = 0$ for $\mathbf{x} \in \partial\Omega$ implies that $v \in H_0^{1,2}(\Omega)$.

With this choice of v in (12.8.8), we get

$$\int_{\Omega} ((\nabla u \cdot \nabla u)g'(u - K_0) + u g(u - K_0)) \, d\mathbf{x} = \int_{\Omega} f g(u - K_0) \, d\mathbf{x} . \quad (12.8.9)$$

Moreover, it follows from $|g(u - K_0)| \leq |u|$ that $v \in L^1(\Omega)$ as we now prove. Let $S = \{\mathbf{x} : u(\mathbf{x}) > K_0\}$. We have that

$$\int_{\Omega} |g(u - K_0)| \, d\mathbf{x} = \int_S |g(u - K_0)| \, d\mathbf{x} \leq \int_S |u| \, d\mathbf{x} \leq \int_S \frac{1}{K_0} |u|^2 \, d\mathbf{x} \leq \frac{1}{K_0} \int_{\Omega} |u|^2 \, d\mathbf{x} < \infty$$

because $u \in H^{1,2}(\Omega)$ and $u^2(\mathbf{x}) \geq K_0 |u(\mathbf{x})|$ for all $\mathbf{x} \in S$. It is here that we need $K_0 > K \geq 0$ to get $K_0 \neq 0$.

Since $K_0 \int_{\Omega} g(u - K_0) \, d\mathbf{x} \in \mathbb{R}$, we may subtract it from each side of the equality in (12.8.9) to get

$$\int_{\Omega} (\|\nabla u\|_2^2 g'(u - K_0) + (u - K_0) g(u - K_0)) \, d\mathbf{x} = \int_{\Omega} (f - K_0) g(u - K_0) \, d\mathbf{x} .$$

As we have done in (i-a), we may use $(u(\mathbf{x}) - K_0) g(u(\mathbf{x}) - K_0) > 0$ for $\mathbf{x} \in \Omega$ such that $u(\mathbf{x}) > K_0$ and $(u(\mathbf{x}) - K_0) g(u(\mathbf{x}) - K_0) = 0$ for $\mathbf{x} \in \Omega$ such that $u(\mathbf{x}) \leq K_0$ to conclude that $u(\mathbf{x}) \leq K_0$ for all $\mathbf{x} \in \Omega$. Since $K_0 > K$ is arbitrary, we get $u(\mathbf{x}) \leq K$ for all $\mathbf{x} \in \Omega$.

ii) To prove the first inequality in the conclusion of the theorem, it suffices to consider $-u$. ■

The next theorem is similar to Theorem 12.8.3 but it is not as strong. In particular, it does not rule out the possibility that the maximum be also reached inside the set Ω .

Theorem 12.8.8

Let Ω be an open subset of $\mathbb{R}^n$ and

$$L(\mathbf{x}, D)u = - \sum_{i,j=1}^n D_{x_i} (b_{i,j}(\mathbf{x}) D_{x_j} u) + b_0(\mathbf{x})u$$

be a strongly elliptic operator. We assume that $b_{i,j} \in H^{1,\infty}(\Omega)$ for all i and j , that $b_j \equiv \sum_{i=1}^n D_{x_i} b_{i,j} = 0$ almost everywhere on Ω for all j , and that $b_0 \in L^\infty(\Omega)$ satisfies $b_0 \geq 0$ almost everywhere on Ω . Suppose that $f \in L^2(\Omega)$ and that $u \in H^{1,2}(\Omega) \cap C(\overline{\Omega})$ is a weak solution of $L(\mathbf{x}, D)u = f$; namely, a solution of

$$B(v, u) = \sum_{i,j=1}^n \int_{\Omega} b_{i,j} D_{x_j} u D_{x_i} v \, d\mathbf{x} + \int_{\Omega} b_0 u v \, d\mathbf{x} = \int_{\Omega} f v \, d\mathbf{x} \quad (12.8.10)$$

for all $v \in H_0^{1,2}(\Omega)$. Then

1. $u \geq 0$ on $\partial\Omega$ and $f \geq 0$ almost everywhere on Ω implies that $u \geq 0$ on Ω .

Moreover, if $b_0 = 0$ almost everywhere on Ω and Ω is bounded, then

2. $f \geq 0$ almost everywhere on Ω implies that $u(\mathbf{x}) \geq \inf_{\mathbf{y} \in \partial\Omega} u(\mathbf{y})$ for all $\mathbf{x} \in \Omega$, and
3. $f \leq 0$ almost everywhere on Ω implies that $u(\mathbf{x}) \leq \sup_{\mathbf{y} \in \partial\Omega} u(\mathbf{y})$ for all $\mathbf{x} \in \Omega$, and
4. $f = 0$ almost everywhere on Ω implies that $\inf_{\mathbf{y} \in \partial\Omega} u(\mathbf{y}) \leq u(\mathbf{x}) \leq \sup_{\mathbf{y} \in \partial\Omega} u(\mathbf{y})$ for all $\mathbf{x} \in \Omega$.

Proof.

The variational formulation of our elliptic partial differential equation is given by (12.8.10) because we assume that $b_j \equiv \sum_{i=1}^n D_{x_i} b_{i,j} = 0$ almost everywhere on Ω for all j . According to [3], there is version of the present theorem with b_j not almost everywhere 0 but the prove is much harder than the present proof.

1) Let $g \in C^1(\mathbb{R})$ be the function defined in the proof of Proposition 12.8.7. We define $h \in C^1(\mathbb{R})$ by $h(x) = -g(-x)$ for $x \in \mathbb{R}$. As in the proof of Proposition 12.8.7, we have that $v = h(u) \in H_0^{1,2}(\Omega)$. If we substitute $v = h(u)$ in (12.8.10), we get

$$\sum_{i,j=1}^n \int_{\Omega} b_{i,j} D_{x_j} u D_{x_i} u h'(u) d\mathbf{x} + \int_{\Omega} b_0 u h(u) d\mathbf{x} = \int_{\Omega} f h(u) d\mathbf{x} .$$

Note that $f h(u) \leq 0$ and $b_0 u h(u) \geq 0$ almost everywhere in Ω because $h(u(\mathbf{x})) = 0$ for $\mathbf{x}$ such that $u(\mathbf{x}) > 0$ and $h(u(\mathbf{x})) < 0$ for $\mathbf{x}$ such that $u(\mathbf{x}) < 0$ by construction. Thus

$$\sum_{i,j=1}^n \int_{\Omega} b_{i,j} D_{x_j} u D_{x_i} u h'(u) d\mathbf{x} \leq 0 .$$

Since the partial differential equation is strongly elliptic, and $h'(x) = g'(-x) > 0$ for $x < 0$ and $h'(x) = 0$ for $x \geq 0$, there exists $\theta > 0$ such that

$$\sum_{i,j=1}^n \int_{\Omega} b_{i,j} D_{x_j} u D_{x_i} u h'(u) d\mathbf{x} \geq \theta \int_{\Omega} \|\nabla u\|_2^2 h'(u) d\mathbf{x} .$$

Hence,

$$\int_{\Omega} \|\nabla u\|_2^2 h'(u) d\mathbf{x} \leq 0 .$$

We get $\|\nabla u\|_2^2 h'(u) = 0$ almost everywhere on Ω .

Let

$$E(t) = \int_0^t \sqrt{h'(s)} ds . \quad (12.8.11)$$

$E \circ u \in C(\overline{\Omega})$ because $u \in C(\overline{\Omega})$. Moreover, $E(u(\mathbf{x})) = 0$ for $\mathbf{x} \in \partial\Omega$ because $u(\mathbf{x}) \geq 0$ for $\mathbf{x} \in \partial\Omega$ and $h'(s) = 0$ for $s > 0$. Since $\|\nabla E(u(\mathbf{x}))\|_2^2 = \|\nabla u(\mathbf{x})\|_2^2 h'(u(\mathbf{x})) = 0$ almost everywhere on Ω , we have that $E \circ u$ is constant on $\overline{\Omega}$. Thus, $E \circ u$ constant on $\overline{\Omega}$ and null on $\partial\Omega$ imply that $E(u(\mathbf{x})) = 0$ for all $\mathbf{x} \in \overline{\Omega}$. This implies that $u(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in \Omega$ because $u(\mathbf{x}) < 0$ for some $\mathbf{x}$ implies that $E(u(\mathbf{x})) > 0$ since $h'(s) > 0$ for $s < 0$.

- 2) Let $K = \inf_{\mathbf{x} \in \partial\Omega} u(\mathbf{x})$ and consider the function $u - K$. Since Ω is bounded, we have that $u - K \in H^{1,2}(\Omega)$. Moreover, since $b_0 = 0$ almost everywhere on Ω , we have that $u - K$ satisfies (12.8.10). Since $u - K \geq 0$ on $\partial\Omega$, we may use the first part of the theorem to conclude that $u - K \geq 0$ on Ω .
- 3) The proof is very similar to the proof (2). Let $K = \max_{\mathbf{x} \in \partial\Omega} u(\mathbf{x})$ and consider the function $K - u$. Since Ω is bounded, we have that $K - u \in H^{1,2}(\Omega)$. Moreover, since $b_0 = 0$ almost everywhere on Ω , we have that $K - u$ satisfies (12.8.10) with f replaced by $-f$. Since $K - u \geq 0$ on $\partial\Omega$, we may use the first part of the theorem (with f replaced by $-f$ and u by $K - u$) to conclude that $K - u \geq 0$ on Ω .
- 4) This follows from (2) and (3). ■

We have the following maximum principle for the Neumann problem.

Proposition 12.8.9

Let Ω be an open subset of $\mathbb{R}^n$. Suppose that $f \in L^2(\Omega)$ and that $u \in H^{1,2}(\Omega)$ is a solution of

$$B(v, u) = \int_{\Omega} (\nabla u \cdot \nabla v + uv) \, d\mathbf{x} = \int_{\Omega} f v \, d\mathbf{x}$$

for all $v \in H^{1,2}(\Omega)$. Then

$$\inf_{\mathbf{x} \in \Omega} f(\mathbf{x}) \leq u \leq \sup_{\mathbf{x} \in \Omega} f(\mathbf{x})$$

almost everywhere in Ω .

Proof.

The proof is similar to the proof of Proposition 12.8.7. ■

12.9 Addendum

We have presented some examples of the variational method to solve elliptic partial differential equations in the previous sections. More examples can be found in [3, 32]. The presentation of the variational method in [32] is slightly different from our presentation. In particular, they do not use Stampacchia's theorem, Theorem 4.5.4. Nevertheless, they have many interesting examples with their physical motivation and interpretation. This is an excellent reference for a reader interested in the applications of partial differential equations.

12.10 Exercises

Suggested exercises:

- In [3]: number 8.23 to 8.42 in Chapter 8; problems 45 to 51 in the chapter Problems.

- In [22]: number 4 in Section 8.2; number 2 in Section 8.3.

Chapter 13

Shock Waves

In this chapter, we will consider equations of the form

$$\frac{\partial u}{\partial y} + \frac{\partial}{\partial x} G(u) = 0, \quad (13.0.1)$$

where $u : \mathbb{R} \times [0, \infty[\rightarrow \mathbb{R}^n$ ¹ and $G : \mathbb{R}^n \rightarrow \mathbb{R}^n$. We also assume that an initial condition is given; namely, that $u(x, 0)$ is given for all $x \in \mathbb{R}$. Note that y should be interpreted as the time².

Nonlinear partial differential equations of this type are called **conservation laws**. To formally justify this name, we consider a long thin pipe filled with water to which a substance has been added. Let u be the concentration of this substance at time t and position x along the pipe, and let $G(u)$ be the flux of the substance in the selected positive direction of the pipe when the concentration is u . The quantity of the substance in a tiny section $[x_1, x_2]$ of the pipe at time y is $\int_{x_1}^{x_2} u(x, y) dx$. Assuming no external contribution, a conservation law states that the rate of change of this quantity is determined by the difference between the flux of the substance at each end of the tiny section of the pipe. Mathematically, the law states that

$$\frac{\partial}{\partial y} \int_{x_1}^{x_2} u(x, y) dx = -G(u(x_2, t)) + G(u(x_1, t)).$$

If we assume that u and G are sufficiently differentiable functions, we may rewrite this equation as

$$\int_{x_1}^{x_2} \frac{\partial u}{\partial y}(x, y) dx = - \int_{x_1}^{x_2} \frac{\partial}{\partial x} G(u(x, y)) dx.$$

Hence,

$$\int_{x_1}^{x_2} \left(\frac{\partial u}{\partial y}(x, y) + \frac{\partial}{\partial x} G(u(x, y)) \right) dx = 0.$$

¹Systems of conservation laws with $u : \mathbb{R}^k \times [0, \infty[\rightarrow \mathbb{R}^n$ and $k > 1$ could also be studied. However, except for an increase in the complexity of the analysis, there is no specific benefit or new insight gained in doing that at the present level. See [17].

²We should have used t instead of y as it is generally done in textbooks on this subject but we choose for consistency to use the same notation that we have used when introducing first order partial differential equations.

Since this is true for all tiny section $[x_1, x_2]$, we should have that

$$\frac{\partial u}{\partial y}(x, y) + \frac{\partial}{\partial x} G(u(x, y)) = 0$$

for all x and y . This is (13.0.1) for $n = 1$.

One of the first reference to conservation laws was in physics to describe gas dynamics. This type of equations are usually obtain as models from physical systems where the total amount of some quantity in a given region is constant; any internal variation is compensated by the flux across the boundary of the region. We only give a very brief introduction to the subject. A good reference on this subject is [35].

13.1 Discontinuous Solutions

We consider the nonlinear partial differential equation

$$\frac{\partial u}{\partial y} + g(u) \frac{\partial u}{\partial x} = 0, \quad (13.1.1)$$

where $g : \mathbb{R} \rightarrow \mathbb{R}$ is a continuous function. (13.1.1) is of the form (13.0.1) where G is a primitive of g . When $g(u) = u$, this partial differential equation is called the **Burgers' equation**. It has been used as a simple model to explain turbulence, where u represents a concentration and y is the time.

We use the method of characteristics to solve (13.1.1) with the initial condition $u(x, 0) = h(x)$. The characteristic equations are

$$x'(t) = g(u(t)) \quad , \quad y'(t) = 1 \quad \text{and} \quad u'(t) = 0.$$

The initial condition is given by

$$x(0) = s \quad , \quad y(0) = 0 \quad \text{and} \quad u(0) = h(s)$$

for $s \in \mathbb{R}$. From $u'(t) = 0$ and $u(0) = h(s)$, we get $u = u(t, s) = h(s)$. We have that u does not depend on t . From $y'(t) = 1$ and $y(0) = 0$, we get $y = y(t, s) = t$. Finally, from $x'(t) = g(u) = g(h(s))$ and $x(0) = s$, we get $x = x(t, s) = t g(h(s)) + s$.

If we substitute $h(s) = u$ and $t = y$ in $x = t g(h(s)) + s$, we get $s = x - g(u) y$. Hence,

$$u = h(x - g(u) y) \quad (13.1.2)$$

is an implicit solution of the partial differential equation. This gives some information on the solution but not enough when the implicit equation for u cannot be solved explicitly.

It follows from the characteristic curves

$$\{(x, y, u) = (t g(h(s)) + s, t, h(s)) : t \in \mathbb{R}\}$$

for $s \in \mathbb{R}$ that u is constant to $h(s)$ along (sections of) the lines

$$\Sigma_s = \{(x, y) = (t g(h(s)) + s, t) : t \in \mathbb{R}\} ;$$

namely, along the lines $x = g(h(s))y + s$ in the x, y plane. Each line has a slope of $g(h(s))$ if we consider x as a function of y .

Since $\frac{dx}{dy} = g(u(x, y))$ and y represents the time, g is called the **(signal) speed** and u is called the **signal**. In most textbooks, the curves Σ_s are called characteristics. However, we have already used this term when studying first order partial differential equations. Since Σ_s is the projection of a characteristic curve for each s , we will call Σ_s an **x, y characteristics**.

If $g(h(s_1)) > g(h(s_2))$ for $s_1 < s_2$, the lines $\Sigma_{s_i} = \{(g(h(s_i))y + s_i, y) : y \geq 0\}$ for $i = 1$ and 2 intersect at the point

$$\mathbf{p} = \left(\frac{g(h(s_1))s_2 - g(h(s_2))s_1}{g(h(s_1)) - g(h(s_2))}, \frac{s_2 - s_1}{g(h(s_1)) - g(h(s_2))} \right)$$

(Figure 13.1). Thus, u will be discontinuous at this point since $u = h(s_1)$ when we approach $\mathbf{p}$ along Σ_{s_1} and $u = h(s_2)$ when we approach $\mathbf{p}$ along Σ_{s_2} with $h(s_1) \neq h(s_2)$.

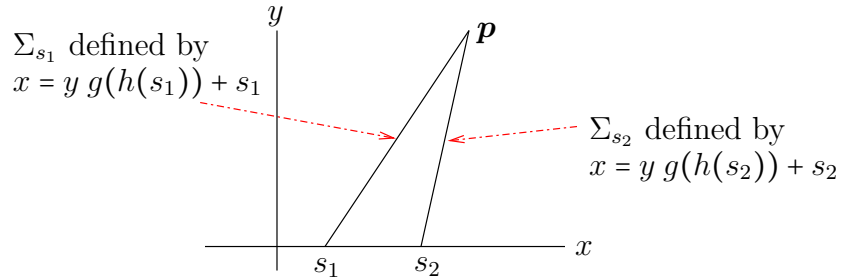


Figure 13.1: The two lines intersect if $g(h(s_1)) > g(h(s_2))$.

When $g \circ h$ is a decreasing function, the solution u is called a **shock wave**. The solution u is discontinuous along the curve $\Gamma = \{(\gamma(y), y) : y \geq 0\}$ of class C^1 (Figure 13.2) given by the points $\mathbf{p}$ described above, and thus (13.1.1) does not have a classical solution. It is important to note that we obtain a discontinuous solution independently of the smoothness of the initial condition. We still get a discontinuous solution even if h is C^∞ or analytic. The nonlinearity of the partial differential equation is to be blamed for this phenomena.

13.2 Strong Solutions

We consider

$$\frac{\partial u}{\partial y} + \frac{\partial}{\partial x} G(u) = 0 \tag{13.2.1}$$

with the initial condition $u(x, 0) = h(x)$ for $x \in \mathbb{R}$.

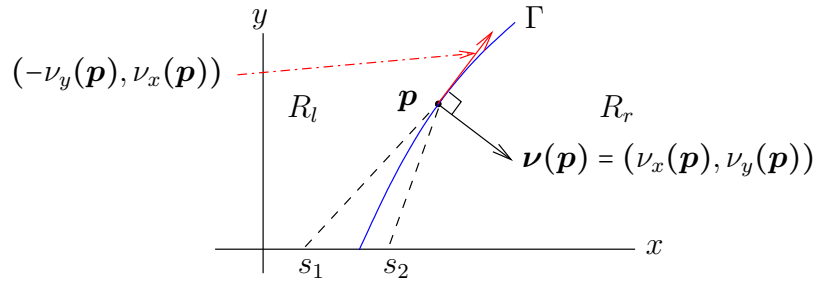


Figure 13.2: Γ is the curve along which the solution u of (13.1.1) is discontinuous.

Definition 13.2.1

A **strong solution** of (13.2.1) is a measurable function $u : \mathbb{R} \times [0, \infty[\rightarrow \mathbb{R}$ such that

$$\int_{[0, \infty[} \int_{\mathbb{R}} \left(u \frac{\partial \psi}{\partial y} + G(u) \frac{\partial \psi}{\partial x} \right) dx dy + \int_{\mathbb{R}} h(x) \psi(x, 0) dx = 0 \quad (13.2.2)$$

for all functions $\psi \in \mathcal{D}(\mathbb{R} \times [0, \infty[)$.

To justify this definition, we assume that $u \in C^1(\mathbb{R} \times]0, \infty[) \cap C(\mathbb{R} \times [0, \infty[)$. Since

$$\frac{\partial}{\partial y}(u\psi) + \frac{\partial}{\partial x}(G(u)\psi) = \left(u \frac{\partial \psi}{\partial y} + G(u) \frac{\partial \psi}{\partial x} \right) + \left(\frac{\partial u}{\partial y} + \frac{\partial}{\partial x} G(u) \right) \psi.$$

we get from the divergence theorem that

$$\begin{aligned} \int_{\partial B} (G(u)\psi, u\psi) \cdot \boldsymbol{\nu} ds &= \int_B \left(\frac{\partial}{\partial y}(u\psi) + \frac{\partial}{\partial x}(G(u)\psi) \right) dx dy \\ &= \int_B \left(u \frac{\partial \psi}{\partial y} + G(u) \frac{\partial \psi}{\partial x} \right) dx dy + \int_B \left(\frac{\partial u}{\partial y} + \frac{\partial}{\partial x} G(u) \right) \psi dx dy \end{aligned} \quad (13.2.3)$$

for all bounded open set $B \subset \mathbb{R} \times [0, \infty[$ and $\psi \in \mathcal{D}(\mathbb{R} \times [0, \infty[)$, where $\boldsymbol{\nu}(x, y)$ is the outward unit normal to ∂B at $(x, y) \in \partial B$. We assume that ∂B is of class C^1 almost everywhere.

Suppose that $u \in C^1(\mathbb{R} \times]0, \infty[) \cap C(\mathbb{R} \times [0, \infty[)$ satisfies (13.2.1) in the classical sense in $\mathbb{R} \times]0, \infty[$ and $u(x, 0) = h(x)$ for $x \in \mathbb{R}$. Given $\psi \in \mathcal{D}(\mathbb{R} \times [0, \infty[)$, choose a box $B =]-a, a[\times [0, b[$ such that $\text{supp } \psi \subset B$. Since $\frac{\partial u}{\partial y} + \frac{\partial}{\partial x} G(u) = 0$ in $\mathbb{R} \times [0, \infty[$ and $u(x, 0) = h(x)$ for $x \in \mathbb{R}$, it follows from (13.2.3) that

$$\begin{aligned} - \int_{-a}^a h(x) \psi(x, 0) dx &= - \int_{-a}^a u(x, 0) \psi(x, 0) dx \\ &= \int_{-a}^a \int_0^b \left(u \frac{\partial \psi}{\partial y} + G(u) \frac{\partial \psi}{\partial x} \right) dy dx + \int_{-a}^a \int_0^b \left(\frac{\partial u}{\partial y} + \frac{\partial}{\partial x} G(u) \right) \psi dy dx \\ &= \int_{-a}^a \int_0^b \left(u \frac{\partial \psi}{\partial y} + G(u) \frac{\partial \psi}{\partial x} \right) dy dx. \end{aligned}$$

We get the negative sign because $\nu(x, y) = (0, -1)$ on the segment $S = \{(x, 0) : -a \leq x \leq a\}$. The positive orientation along ∂B is counterclockwise. Note that $\psi(x, y) = 0$ for $(x, y) \in (\partial B) \setminus S$ because $\text{supp } \psi \subset B$. Therefore, (13.2.2) is satisfied because $\psi(x, y) = 0$ for all $(x, y) \notin B$ since $\text{supp } \psi \subset B$.

Suppose now that (13.2.2) is satisfied and that $u \in C^1(\mathbb{R} \times]0, \infty[) \cap C(\mathbb{R} \times [0, \infty[)$. Given $B =]-a, a[\times [0, b[\subset \mathbb{R} \times [0, \infty[$ and any $\psi \in \mathcal{D}(B)$, we get from (13.2.2) and (13.2.3) that

$$\begin{aligned} - \int_{-a}^a u(x, 0) \psi(x, 0) dx &= \int_0^b \int_{-a}^a \left(u \frac{\partial \psi}{\partial y} + G(u) \frac{\partial \psi}{\partial x} \right) dx dy + \int_0^b \int_{-a}^a \left(\frac{\partial u}{\partial y} + \frac{\partial}{\partial x} G(u) \right) \psi dx dy \\ &= - \int_{-a}^a h(x) \psi(x, 0) dx + \int_0^b \int_{-a}^a \left(\frac{\partial u}{\partial y} + \frac{\partial}{\partial x} G(u) \right) \psi dx dy. \end{aligned}$$

Therefore,

$$\int_0^b \int_{-a}^a \left(\frac{\partial u}{\partial y} + \frac{\partial}{\partial x} G(u) \right) \psi dx dy + \int_{-a}^a (u(x, 0) - h(x)) \psi(x, 0) dx = 0$$

for all $\psi \in \mathcal{D}(B)$. We get that

$$\int_0^b \int_{-a}^a \left(\frac{\partial u}{\partial y} + \frac{\partial}{\partial x} G(u) \right) \mu dx dy = 0$$

for all $\mu \in \mathcal{D}(]-a, a[\times]0, b[) \subset \mathcal{D}(B)$ because $\mu(x, 0) = 0$ for $x \in \mathbb{R}$. Since $\frac{\partial u}{\partial y} + \frac{\partial}{\partial x} G(u)$ is continuous in B , we get that $\frac{\partial u}{\partial y} + \frac{\partial}{\partial x} G(u) = 0$ in $] -a, a[\times]0, b[$ in the classical sense. It then follows that

$$\int_{-a}^a (h(x) - u(x, 0)) \psi(x, 0) dx = 0$$

for all $\psi \in \mathcal{D}(B)$ because $\frac{\partial u}{\partial y} + \frac{\partial}{\partial x} G(u) = 0$ in $] -a, a[\times]0, b[$. Since $\mathcal{D}B|_{\{(x, 0) : -a < x < a\}} \equiv \mathcal{D}(]-a, a[)$ and $x \mapsto h(x) - u(x, 0)$ is continuous, we get $u(x, 0) = h(x)$ for $-a < x < a$. Since this is true for all boxes $B \subset \mathbb{R} \times]0, \infty[$, we get that u satisfies (13.2.1) in the classical sense in $\mathbb{R} \times]0, \infty[$ and $u(x, 0) = h(x)$ for $x \in \mathbb{R}$.

Considering strong solutions enlarges the set of possible solutions and increase the likelihood that a conservation law with an initial condition has a solution. However, it also increase the possibility that there may be more than one solution. This will be addressed by imposing additional conditions on the solutions to ensure that they have a physical meaning.

13.3 Rankine-Hugoniot Formula

Not all shock waves are admissible. They must satisfy a condition that we now deduce.

As before, we assume that u is solution of (13.2.1) which is discontinuous along the curve $\Gamma = \{(\gamma(y), y) : y \geq y_0\}$ of class C^1 for some $y_0 \geq 0$.

Let V be a bounded open subset of $\mathbb{R} \times [0, \infty[$ such that V is split in two by Γ . Given $\psi \in \mathcal{D}(V)$, the first integral in (13.2.2) can be split into two integrals and the second integral is null. We get

$$\iint_{V_l} \left(u \frac{\partial \psi}{\partial y} + G(u) \frac{\partial \psi}{\partial x} \right) dx dy + \iint_{V_r} \left(u \frac{\partial \psi}{\partial y} + G(u) \frac{\partial \psi}{\partial x} \right) dx dy = 0, \quad (13.3.1)$$

where the first integral is over the region $V_l = \{(x, y) \in V : x < \gamma(y)\}$ to the “left” of Γ . Similarly, the second integral is over the region $V_r = \{(x, y) \in V : x > \gamma(y)\}$ to the “right” of Γ . In both regions, we assume that the function u is of class C^1 , can be extended continuously to the closure of the region, and satisfies (13.2.1) in the classical sense.

We may use the Divergence Theorem to evaluate the first integral in (13.3.1).

$$\begin{aligned} & \iint_{V_l} \left(u \frac{\partial \psi}{\partial y} + G(u) \frac{\partial \psi}{\partial x} \right) dx dy \\ &= \iint_{V_l} \left(- \left(\frac{\partial u}{\partial y} + \frac{\partial}{\partial x} G(u) \right) \psi + \left(\frac{\partial}{\partial y} (u \psi) + \frac{\partial}{\partial x} (G(u) \psi) \right) \right) dx dy \\ &= - \iint_{V_l} \left(\frac{\partial u}{\partial y} + \frac{\partial}{\partial x} G(u) \right) \psi dx dy + \int_{\Gamma \cap \partial V_l} (G(u^-) \psi, u^- \psi) \cdot \boldsymbol{\nu} ds, \end{aligned} \quad (13.3.2)$$

where $\boldsymbol{\nu}(\mathbf{p})$ is the unit normal to Γ at $\mathbf{p} \in \Gamma$ pointing to the right of Γ , so outside V_l , and

$$u^-(\mathbf{p}) = \lim_{t \rightarrow 0^-} u(\mathbf{p} + t \boldsymbol{\nu}(\mathbf{p}))$$

for $\mathbf{p} \in \Gamma$. We have replaced the domain of integration of the second integral from ∂V_l to $\Gamma \cap \partial V_l$ because $\psi(\mathbf{p}) = 0$ for all $\mathbf{p} \in (\partial V_l) \setminus \Gamma$. The positive orientation of the contour integral along ∂V_l is counterclockwise.

It follows from (13.2.1) that, on the open domain of integration V_l where $u \in C^1(V_l)$, the first integral in (13.3.2) is null. Thus,

$$\iint_{V_l} \left(u \frac{\partial \psi}{\partial y} + G(u) \frac{\partial \psi}{\partial x} \right) dx dy = \int_{\Gamma \cap \partial V_l} (G(u^-) \psi, u^- \psi) \cdot \boldsymbol{\nu} ds. \quad (13.3.3)$$

Similarly, we have from the second integral in (13.3.1) that

$$\begin{aligned} & \iint_{V_r} \left(u \frac{\partial \psi}{\partial y} + G(u) \frac{\partial \psi}{\partial x} \right) dx dy \\ &= - \iint_{V_r} \left(\frac{\partial u}{\partial y} + \frac{\partial}{\partial x} G(u) \right) \psi dx dy + \int_{\Gamma \cap \partial V_r} (G(u^+) \psi, u^+ \psi) \cdot (-\boldsymbol{\nu}) ds \\ &= - \int_{\Gamma \cap \partial V_r} (G(u^+) \psi, u^+ \psi) \cdot \boldsymbol{\nu} ds, \end{aligned} \quad (13.3.4)$$

where

$$u^+(\mathbf{p}) = \lim_{t \rightarrow 0^+} u(\mathbf{p} + t \boldsymbol{\nu}(\mathbf{p}))$$

for $\mathbf{p} \in \Gamma$. The minus sign in front of the previous line integral comes from the counterclockwise direction of the contour integral along ∂V_r while the unit normal $\boldsymbol{\nu}$ is pointing inside V_r .

Hence, we get from (13.3.1), (13.3.3) and (13.3.4) that

$$\begin{aligned} & \int_{\Gamma \cap \partial V_l} (G(u^-) \psi, u^- \psi) \cdot \boldsymbol{\nu} \, ds - \int_{\Gamma \cap \partial V_r} (G(u^+) \psi, u^+ \psi) \cdot \boldsymbol{\nu} \, ds \\ &= \int_{\Gamma \cap \bar{V}} \psi (G(u^-) - G(u^+), u^- - u^+) \cdot \boldsymbol{\nu} \, ds = 0 \end{aligned} \quad (13.3.5)$$

because

$$\{\mathbf{p} \in \Gamma \cap \partial V_l : \psi(\mathbf{p}) \neq 0\} = \{\mathbf{p} \in \Gamma \cap \partial V_r : \psi(\mathbf{p}) \neq 0\} = \{\mathbf{p} \in \Gamma \cap \bar{V} : \psi(\mathbf{p}) \neq 0\} .$$

Since (13.3.5) is true for all $\psi \in \mathcal{D}(V)$ and open sets V which are split by Γ , we get

$$(G(u^-) - G(u^+), u^- - u^+) \cdot \boldsymbol{\nu} = 0$$

on Γ . If $\boldsymbol{\nu} = (\nu_x, \nu_y)$, this gives

$$\frac{G(u^+(\mathbf{p})) - G(u^-(\mathbf{p}))}{u^+(\mathbf{p}) - u^-(\mathbf{p})} = -\frac{\nu_y(\mathbf{p})}{\nu_x(\mathbf{p})} = \frac{d\gamma}{dy}(y) \quad (13.3.6)$$

for all $\mathbf{p} = (\gamma(y), y) \in \Gamma$.

The relation (13.3.6) is known as the **Rankine-Hugoniot formula**. All shock waves must satisfy this condition. $\frac{d\gamma}{dy}(y)$ is the speed of the shock wave.

Example 13.3.1

Consider the Burgers' equation

$$\frac{\partial u}{\partial y} + u \frac{\partial u}{\partial x} = 0 \quad (13.3.7)$$

with the initial condition

$$u(x, 0) = h(x) = \begin{cases} 1 & \text{if } x \leq 0 \\ 1 - x & \text{if } 0 < x \leq 1 \\ 0 & \text{if } 1 < x \end{cases}$$

This equation is of the form (13.1.1) with $g(u) = u$. We may expect a shock wave since $g \circ h$ is a decreasing function on the interval $[0, 1]$.

We first use the method of characteristic to solve this partial differential equation. The characteristic equations are

$$x'(t) = u(t) \quad , \quad y'(t) = 1 \quad \text{and} \quad u'(t) = 0 .$$

The initial condition is given by

$$x(0) = s \quad , \quad y(0) = 0 \quad \text{and} \quad u(0) = h(s)$$

for all s . From $u'(t) = 0$ and $u(0) = h(s)$, we get $u = u(t, s) = h(s)$. From $y'(t) = 1$ and $y(0) = 0$, we get $y = y(t, s) = t$. Finally, from $x'(t) = u = h(s)$ and $x(0) = s$, we get $x = x(t, s) = t h(s) + s$.

There are three cases to consider: $s < 0$, $0 < s < 1$ and $s > 1$.

1. For $s < 0$, we have that $x = t + s$, $y = t$ and $u = 1$. Thus $u = 1$ along the lines $x = y + s$.
2. For $0 < s < 1$, we have that $x = (1 - s)t + s$, $y = t$ and $u = 1 - s$. Thus $u = \frac{x - 1}{y - 1}$ along the lines $x = (1 - s)y + s$.
3. Finally, for $s > 1$, we have that $x = s$, $y = t$ and $u = 0$. Thus $u = 0$ along the lines $x = s$.

These results are summarized in Figure 13.3.

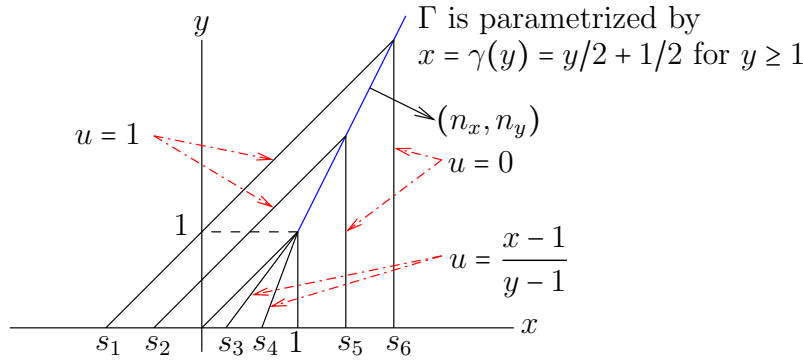


Figure 13.3: The solution predicted by the method of characteristics.

To determine the curve $\Gamma = \{(\gamma(y), y) : y \geq y_0\}$, we use the Rankine-Hugoniot formula

$$\begin{aligned} \frac{d\gamma}{dy}(y) &= \frac{G(u^+(\mathbf{p})) - G(u^-(\mathbf{p}))}{u^+(\mathbf{p}) - u^-(\mathbf{p})} = \frac{1}{2} \frac{(u^+(\mathbf{p}))^2 - (u^-(\mathbf{p}))^2}{u^+(\mathbf{p}) - u^-(\mathbf{p})} \\ &= \frac{1}{2} (u^+(\mathbf{p}) + u^-(\mathbf{p})) \end{aligned} \quad (13.3.8)$$

for all $\mathbf{p} = (\gamma(y), y) \in \Gamma$. We have chosen G defined by $G(u) = \frac{u^2}{2}$ as the primitive of g .

The curve Γ is defined only for $y \geq y_0 = 1$. Whatever the choice of Γ , we have that $u^+(\mathbf{p}) = 1$ and $u^-(\mathbf{p}) = 0$ for all $\mathbf{p} \in \Gamma$. Hence, $\frac{d\gamma}{dy}(y) = \frac{1}{2}$ for all $y > 1$. Thus $\gamma(y) = \frac{1}{2}y + \frac{1}{2}$ for $y \geq 1$ because Γ must contain the point $(1, 1)$ since u is discontinuous at $(1, 1)$.

The graph of $u(x, y)$ for some fixed values of y can be found in Figure 13.4. The shock wave travels at a speed of $1/2$. For the Burgers' equation, we have that

$$\frac{\partial x}{\partial y}(x_l, y) = u(x_l, y) = \frac{dG}{du}(u(x_l, y)) > \gamma'(y) > \frac{dG}{du}(u(x_r, y)) = u(x_r, y) = \frac{\partial x}{\partial y}(x_r, y)$$

for all $(x_l, y) \in R_l$ and $(x_r, y) \in R_u$. Namely, the speed of the wave behind the shock wave is larger than the speed of the wave ahead of the shock wave. This is the fundamental reason behind the existence of a shock wave. This relation will also play a fundamental role in determining a physically valid solution as we will show in the next section. ♣

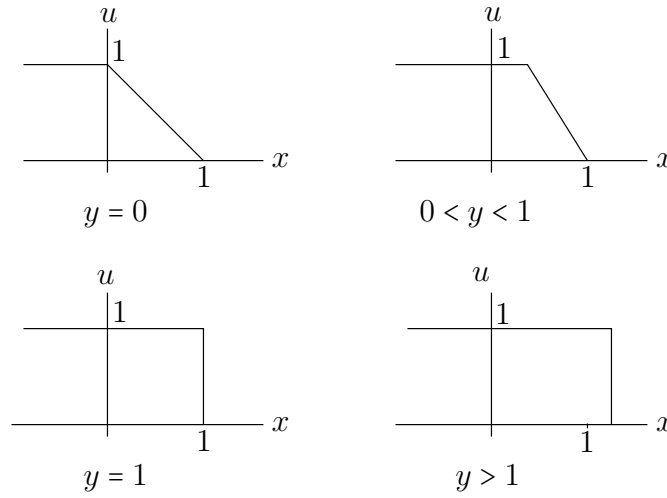


Figure 13.4: A shock wave for the Burgers' equation travels at a speed of $1/2$.

The next example illustrates some unexpected consequence of the Rankine-Hugoniot formula.

Example 13.3.2

We have seen that the solution of the Burgers' equation $\frac{\partial u}{\partial y} + \frac{\partial}{\partial x} \left(\frac{u^2}{2} \right) = 0$ must satisfy the Rankine-Hugoniot formula $\frac{d\gamma}{dy}(y) = \frac{1}{2} (u^+ + u^-)$, where $u^+ = u^+(\mathbf{p})$, $u^- = u^-(\mathbf{p})$ and $\mathbf{p} = (\gamma(y), y)$.

If we consider instead the conservation law

$$u \frac{\partial u}{\partial y} + u^2 \frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} + \frac{\partial}{\partial x} \left(\frac{2\sqrt{2}}{3} v^{3/2} \right) = 0 ,$$

where $v = u^2/2$. This is simply the Burgers' equation multiplied by u . Then, the solution v must now satisfy the Rankine-Hugoniot formula

$$\frac{d\gamma}{dy}(y) = \frac{2\sqrt{2}}{3} \left(\frac{(v^+)^{3/2} - (v^-)^{3/2}}{v^+ - v^-} \right) = \frac{2}{3} \left(\frac{(u^+)^2 + u^+ u^- + (u^-)^2}{u^+ + u^-} \right)$$

where $v^+ = (u^+)^2/2$ and $v^- = (u^-)^2/2$. But this is clearly not possible for the same curve Γ . The curve where the solution is discontinuous and where the shock wave is generated is not the same for both system. The two shock waves travel at different speeds. We therefore get different solutions despite the fact that the second conservation law is just the first one multiplied by u . This is certainly not what we have with ordinary differential equations. ♣

13.4 Existence, Uniqueness and Asymptotic Behaviour of Solutions

When $g \circ h$ is a continuously increasing function, a classical solution u of (13.1.1) exists (Figure 13.5).

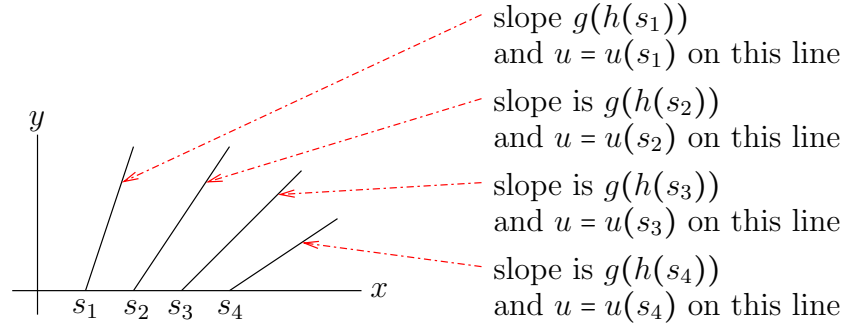


Figure 13.5: When $g \circ h$ is continuously increasing, we get a nice travelling wave.

However, we may, and often have, many solutions when $g \circ h$ fails to be continuously increasing.

Example 13.4.1

A simple example of non-uniqueness of the solution is given by the Burgers' equation (13.3.7) with the initial condition

$$u(x, 0) = h(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ 1 & \text{if } x > 0 \end{cases}$$

It is easy to verify that

$$u(x, y) = u_0(x, y) \equiv \begin{cases} 0 & \text{if } x < t/2 \\ 1 & \text{if } x \geq t/2 \end{cases}$$

and

$$u(x, y) = u_1(x, y) \equiv \begin{cases} 0 & \text{if } x \leq 0 \\ x/y & \text{if } 0 < x < y \\ 1 & \text{if } x \geq y \end{cases}$$

are both strong solutions of the Burgers' equation with the initial condition. We have sketched these solutions in Figure 13.6.

The solution u_0 represents a “shock wave” given by the Rankine-Hugoniot formula. More specifically, if the solution u is discontinuous along a curve $\Gamma = \{(\gamma(y), y) : y \geq y_0\}$ for some $y_0 \geq 0$, then (13.3.8) yields $\gamma'(y) = (u^+(\mathbf{p}) + u^-(\mathbf{p}))/2$ for all $\mathbf{p} = (\gamma(y), y) \in \Gamma$. If we set $u(x, y) = 0$ for (x, y) on the left side of Γ and 1 on the right side of Γ , we get $\gamma'(y) = 1/2$

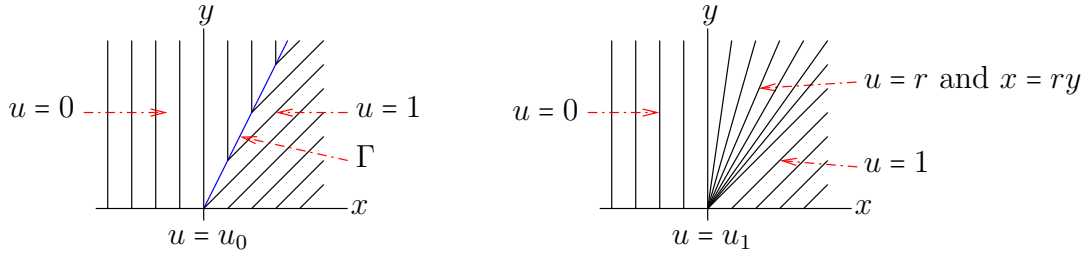


Figure 13.6: Several solutions of the Burgers' equation. The solution on the right is called a rarefaction wave.

and the solution u_0 . As we will see shortly, this type of “shock waves” is not physically acceptable.

The solution u_1 represents an **expansive wave** or a **rarefaction wave**.

There are in fact a continuum of solutions given by

$$u_r(x, y) = \begin{cases} 0 & \text{if } x \leq 0 \\ x/y & \text{if } 0 < x \leq ry \\ r & \text{if } ry < x \leq (r+1)y/2 \\ 1 & \text{if } x > (r+1)y/2 \end{cases}$$

for $0 \leq r \leq 1$. This family of solutions is sketched in Figure 13.7. Each member of the family is a combination of a shock wave and a rarefaction wave. ♣

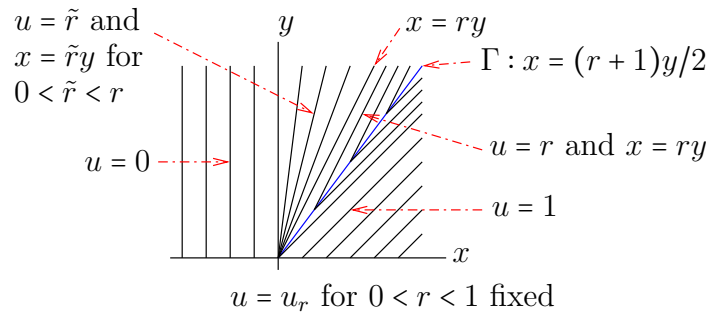


Figure 13.7: Several solutions of the Burgers' equation.

13.4.1 Rarefaction Waves

We first address the issues related to rarefaction waves. We consider (13.2.1) with the initial condition $u(x, 0) = h(x) = u^-$ for $x < 0$ and $u(x, 0) = h(x) = u^+$ for $x > 0$. Conservation laws

with such initial conditions are called **Riemann problems** in honour of Bernhard Riemann who was one of the first person to rigorously study (using mathematics) the gas dynamics experiment called “shock tube”. This experiment consists in a long tube separated in the middle by a membrane. The tube is filled with a gas that has a different pressure and density on both sides of the membrane. Then the membrane is removed.

We assume that $\frac{d^2G}{du^2}(u) > 0$ for all $u \in \mathbb{R}$ to simplify the discussion. This is the situation that we have for the Burgers’ equation for instance.

We note that $u_\mu(x, y) \equiv u(\mu x, \mu y)$ is a solution of (13.2.1) for all $\mu > 0$ ³. Therefore, we seek a solution of the form $u(x, y) = w(x/y)$ where $w \in C^1(\mathbb{R})$. If we set $z = x/y$ and substitute $u(x, y) = w(z)$ in (13.2.1), we get

$$\begin{aligned} 0 &= \frac{\partial u}{\partial y}(x, y) + \frac{\partial}{\partial x}G(u(x, y)) = -\frac{x}{y^2} \frac{dw}{dz}(z) + \frac{1}{y} \frac{dG}{dw}(w(z)) \frac{dw}{dz}(z) \\ &= \frac{1}{y} \left(-z + \frac{dG}{dw}(w(z)) \right) \frac{dw}{dz}(z). \end{aligned}$$

Thus

$$\left(\frac{dG}{dw}(w(z)) - z \right) \frac{dw}{dz}(z) = 0.$$

If $\frac{dw}{dz}(z) \neq 0$ for all z in an open subset V of $\mathbb{R}$, then $\frac{dG}{dw}(w(z)) = z$ for all $z \in V$. In fact, the converse is also true. If we derive $\frac{dG}{dw}(w(z)) = z$ with respect to z , we get $\frac{d^2G}{dw^2}(w(z)) \frac{dw}{dz}(z) = 1$ for $z \in V$. Since $\frac{d^2G}{dw^2}(w) > 0$, we get that $\frac{dw}{dz}(z) \neq 0$ for $z \in V$.

Since $\frac{d^2G}{dw^2}(w) > 0$, we have that $\frac{dG}{dw}(w) > 0$ is an increasing function. In particular, $\frac{dG}{dw}$ is invertible. The unique solution for w in V is therefore given by solving $\frac{dG}{dw}(w(z)) = z$ for $z \in V$. Therefore, $u(x, y) = w(z) = \left(\frac{dG}{dw} \right)^{-1}(z)$ for $z \in V$ and $z = x/y$.

Definition 13.4.2

We say that u^+ is **connected** to u^- by a rarefaction wave if $\frac{dG}{du}(u^+) > \frac{dG}{du}(u^-)$, and $\frac{dG}{dw}(w(z)) = z$ for $\frac{dG}{du}(u^+) > z > \frac{dG}{du}(u^-)$.

The last condition in the definition is used to define u in the wedge region where u is not defined by the initial conditions.

A rarefaction wave for the Burgers’ equation satisfying definition 13.4.2 is sketched in Figure 13.8. The solution u_1 in Example 13.4.1 is an example of a rarefaction wave satisfying Definition 13.4.2 and connecting 0 to 1.

³ u_μ is in fact a solution of (13.2.1) for all $\mu \in \mathbb{R}$ but the initial condition will not be satisfied for $\mu \leq 0$.

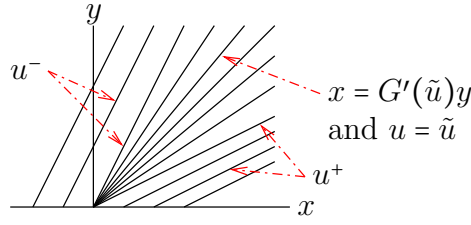


Figure 13.8: Example of a rarefaction wave given by Definition 13.4.2. We assume that $u(x, 0) = h(x) = u^-$ for $x < 0$ and $u(x, 0) = h(x) = u^+$ for $x > 0$, and $G'(u^-) > 0$. The lines where u is constant are given by $x = G'(u^-)y + s$ for $s \neq 0$ where $u = u^-$, $x = G'(u^+)y + s$ for $s \neq 0$ where $u = u^+$, and $x = G'(\tilde{u})y$ where $u = \tilde{u}$ is between u^+ and u^- .

13.4.2 Shock Waves

For mathematicians, there is no reason to reject any solution. However, for physicists, the conservation laws represent real natural phenomena with only one physically relevant solution. It is therefore natural to add an extra condition to conservation laws. Unfortunately, there is not a single condition that fits all the physical applications. The condition that we present below, due to Peter Lax, best addresses the problems in gas dynamics.

Definition 13.4.3 (Lax Shock Condition)

We consider the partial differential equation (13.2.1). Suppose that the strong solution u is discontinuous along a curve $\Gamma = \{(\gamma(y), y) : y \geq y_0\}$ of class C^1 for some $y_0 \geq 0$. The solution u must satisfy the Rankine-Hugoniot formula (13.3.6) and

$$\frac{dG}{du}(u^-(\mathbf{p})) > \frac{\partial \gamma}{\partial y}(y) > \frac{dG}{du}(u^+(\mathbf{p})) \quad (13.4.1)$$

for all $\mathbf{p} = (\gamma(y), y) \in \Gamma$.

In Example 13.4.1, the only solution satisfying Lax shock condition is u_1 since there is no discontinuity and so no shock wave. The other solutions including shock waves do not satisfy

$$G'(u^-(\mathbf{p})) > \gamma'(y) = \frac{u^+(\mathbf{p}) + u^-(\mathbf{p})}{2} = \frac{1+r}{2} > G'(u^+(\mathbf{p}))$$

for all $\mathbf{p} = (\gamma(y), y) \in \Gamma$, where $G(u) = u^2/2$.

Suppose that $G''(u) > 0$ for all $u \in \mathbb{R}$. Then, $G'(u(x_l, y)) > G'(u(x_r, y))$ for $x_l < x_r$ only if $u(x_l, y) > u(x_r, y)$ for $x_l < x_r$. This is the situation for u_1 in Example 13.4.1. It then follows that $G'(u^-(\mathbf{p})) > G'(u^+(\mathbf{p}))$ for all $\mathbf{p} \in \Gamma$ because $u^-(\mathbf{p}) > u^+(\mathbf{p})$. This is typical when $G''(u) > 0$ for all $u \in \mathbb{R}$. We leave it to the reader to deduce the typical case when $G''(u) < 0$ for all $u \in \mathbb{R}$. Conservation laws that satisfy the condition $G''(u) \neq 0$ for all $u \in \mathbb{R}$

are often called **genuinely nonlinear**. This is the one-dimensional version of what is called a **genuinely nonlinear system of conservation laws** that we will see later.

The relation (13.4.1) is called the **entropy inequality**. There is a physical motivation for (13.4.1). It is based on the notion from gas dynamics that entropy increases across a shock wave. In other words, there is a loss of information across a shock wave. Mathematically, this means that the solution representing the shock wave is not time reversible, you cannot determine the state leading to the shock wave from the form of the shock wave only. This is illustrated in the following example from [35].

Example 13.4.4

The functions defined by

$$u_r(x, y) = \begin{cases} 1 & \text{if } x < y - r/2 \\ (x - r/2)/(y - r) & \text{if } y - r/2 \leq x < r/2 \\ 0 & \text{if } x \geq r/2 \end{cases}$$

for $x \in \mathbb{R}$ and $0 \leq y < r$, and by

$$u_r(x, y) = \begin{cases} 1 & \text{if } x < y/2 \\ 0 & \text{if } x \geq y/2 \end{cases}$$

for $x \in \mathbb{R}$ and $y \geq r$, are all solutions of Burgers' equation (13.3.7) if $0 \leq r \leq 1$ (Figure 13.9). They are equal at $y = 1$. In fact, they are equal for $y \geq 1$. However, it is not possible to determine the form of the wave for $y < 1$ from its form at $y = 1$; namely, from which of the functions u_r the shock wave comes from. ♣

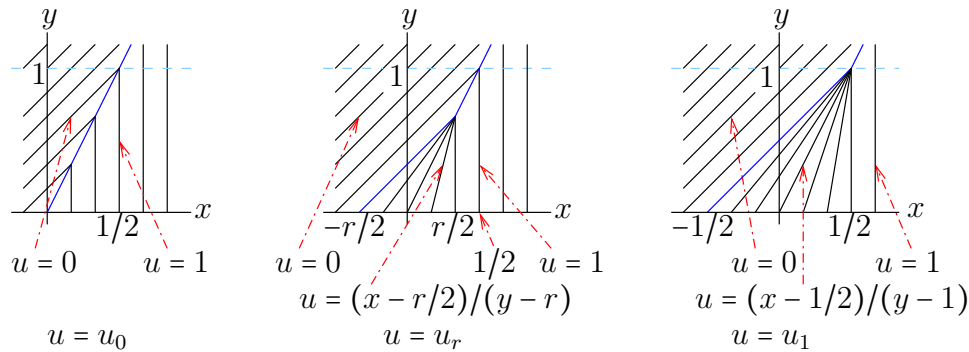


Figure 13.9: Several solutions of the Burgers' equation satisfying the same condition at $y = 1$.

In the systems of conservation laws with shock waves that we have studied, the physically relevant solution u is defined along lines that start on the x axis and end at the curve Γ associated to the shock wave as in Figure 13.3; it is not the opposite as we have in Figure 13.7. Such a situation is sometime expressed informally by saying that “characteristic curves end

on the shock wave but cannot emanate from the shock wave.” This is called an **entropy condition** for systems of conservation laws. This means that some “information” is lost at the front of a shock wave and none is created.

13.4.3 Existence and Uniqueness

In the rest of this section, we adopt the following convention. Given $g : \mathbb{R} \times [0, \infty[\rightarrow \mathbb{R}$, then $g_x : [0, \infty[\rightarrow \mathbb{R}$ and $g_y : \mathbb{R} \rightarrow \mathbb{R}$ denote the functions defined by $g_x(y) = g_y(x) \equiv g(x, y)$ for all $(x, y) \in \mathbb{R} \times [0, \infty[$.

Theorem 13.4.5

Suppose that $h \in L^\infty(\mathbb{R})$ and $G \in C^2(\mathbb{R})$. Let $H = \|h\|_\infty$, $D = \{u \in \mathbb{R} : |u| \leq H\}$, and $M = \max_{u \in D} |G'(u)|$. Assume that $m = \min_{u \in D} G''(u) > 0$. Then, there exists a strong solution of (13.2.1) with the initial condition $u(x, 0) = h(x)$ for $x \in \mathbb{R}$ that satisfies.

1. $|u(x, y)| \leq H$ for all $(x, y) \in \mathbb{R} \times [0, \infty[$.
2. There exists a constant C that depends on H , m and M , such that

$$\frac{u(x+s, y) - u(x, y)}{s} \leq \frac{C}{y} \quad (13.4.2)$$

for $(x, y) \in \mathbb{R} \times]0, \infty[$ and $s > 0$.

Moreover, suppose that u_1 and u_2 are two strong solutions satisfying (1) and (2), and $u_{i, y_1} \rightarrow u_{i, y_2}$ in $L^1(K)$ as $y_1 \rightarrow y_2$ for all compact sets $K \subset \mathbb{R}$ and $y_1, y_2 \in [0, \infty[$ ⁴, then $u_1 = u_2$ almost everywhere in $\mathbb{R} \times [0, \infty[$.

Proof.

A proof of this theorem can be found in [35]. A finite-difference method is used to prove the existence of the solution. It is a very long but elementary proof. In fact, finite-difference methods are often used to prove the existence of solutions of partial differential equations.

We will however prove the uniqueness of the solution. It is an interesting proof.

Suppose that u_1 and u_2 are the two strong solutions of (13.2.1) mentioned in the statement of the theorem. We then have that

$$\int_{[0, \infty[} \int_{\mathbb{R}} \left(u_i \frac{\partial \psi}{\partial y} + G(u_i) \frac{\partial \psi}{\partial x} \right) dx dy + \int_{\mathbb{R}} h(x) \psi(x, 0) dx = 0 \quad (13.4.3)$$

for $i \in \{1, 2\}$, and all $\psi \in \mathcal{D}(\mathbb{R} \times [0, \infty[)$. If we subtract (13.4.3) with $i = 1$ from (13.4.3) with

⁴This hypothesis is not in the original version of the theorem in [35] but we do not see how the theorem can be proved without such type of statement. This is not an unreasonable assumption. Similar discrete versions of this assumption ensure the convergence and stability of finite difference methods.

$i = 2$, we get

$$\begin{aligned} 0 &= \int_{[0, \infty[} \int_{\mathbb{R}} \left((u_2 - u_1) \frac{\partial \psi}{\partial y} + (G(u_2) - G(u_1)) \frac{\partial \psi}{\partial x} \right) dx dy \\ &= \int_{[0, \infty[} \int_{\mathbb{R}} (u_2 - u_1) \left(\frac{\partial \psi}{\partial y} + \mathcal{G}(u_1, u_2) \frac{\partial \psi}{\partial x} \right) dx dy \end{aligned} \quad (13.4.4)$$

for all $\psi \in \mathcal{D}(\mathbb{R} \times [0, \infty[)$, where

$$\mathcal{G}(p, q) = \begin{cases} \frac{G(q) - G(p)}{q - p} & \text{if } p \neq q \\ G'(p) & \text{if } p = q \end{cases}$$

We have that $\mathcal{G} \in C^1(\mathbb{R}^2)$ with $D\mathcal{G}(p, p) = (G''(p) \ G''(p))$ because $G \in C^2(\mathbb{R})$. Moreover, it follows from the hypothesis that $|\mathcal{G}(p, q)| \leq M$ for all $(p, q) \in D \times D$.

To prove that $u_1 = u_2$ almost everywhere in $\mathbb{R} \times]0, \infty[$, It would be nice if we could solve

$$\frac{\partial \psi}{\partial y} + \mathcal{G}(u_1, u_2) \frac{\partial \psi}{\partial x} = \phi$$

for all $\phi \in \mathcal{D}(\mathbb{R} \times]0, \infty[)$ because it would then follow from (13.4.4) that $u_1 = u_2$ almost everywhere. Unfortunately, this is not always possible because $\mathcal{G}(x, y) \equiv \mathcal{G}(u_1(x, y), u_2(x, y))$ for $(x, y) \in \mathbb{R} \times [0, \infty[$ may be discontinuous when u_1 and u_2 are discontinuous functions. To get around this difficulty, we define some functions $u_i^{[j]} \in C(\mathbb{R} \times [0, \infty[)$ for $j \in \mathbb{N}$ which are continuously differentiable with respect to their first variable.

We then show that

$$\frac{\partial \psi^{[j]}}{\partial y} + \mathcal{G}^{[j]} \frac{\partial \psi^{[j]}}{\partial x} = \phi \quad (13.4.5)$$

for $\phi \in \mathcal{D}(\mathbb{R} \times]0, \infty[)$ can always be solved, where $\mathcal{G}^{[j]}(x, y) \equiv \mathcal{G}(u_1^{[j]}(x, y), u_2^{[j]}(x, y))$ for $(x, y) \in \mathbb{R} \times [0, \infty[$.

Given $\phi \in \mathcal{D}(\mathbb{R} \times]0, \infty[)$ and $\psi^{[j]}$ satisfying (13.4.5), we have that

$$\begin{aligned} &\int_{[0, \infty[} \int_{\mathbb{R}} (u_2 - u_1) \phi dy dx \\ &= \int_{[0, \infty[} \int_{\mathbb{R}} (u_2 - u_1) \frac{\partial \psi^{[j]}}{\partial y} dy dx + \int_{[0, \infty[} \int_{\mathbb{R}} (u_2 - u_1) \mathcal{G}^{[j]} \frac{\partial \psi^{[j]}}{\partial x} dy dx \\ &= - \int_{[0, \infty[} \int_{\mathbb{R}} (G(u_2) - G(u_1)) \frac{\partial \psi^{[j]}}{\partial x} dy dx + \int_{[0, \infty[} \int_{\mathbb{R}} (u_2 - u_1) \mathcal{G}^{[j]} \frac{\partial \psi^{[j]}}{\partial x} dy dx \\ &= \int_{[0, \infty[} \int_{\mathbb{R}} (u_2 - u_1) (\mathcal{G}^{[j]} - \mathcal{G}) \frac{\partial \psi^{[j]}}{\partial x} dy dx, \end{aligned} \quad (13.4.6)$$

where (13.4.4) has been used to obtain the second equality.

Finally, we show that the limit as $j \rightarrow \infty$ in (13.4.6) yields

$$\int_{[0, \infty[} \int_{\mathbb{R}} (u_2 - u_1) \phi dx dy = 0.$$

Since $\phi \in \mathcal{D}(\mathbb{R} \times]0, \infty[$ is arbitrary, this will imply that $u_1 = u_2$ almost everywhere.

To glue all these arguments together will require several steps.

i) Choose $\eta \in \mathcal{D}(\mathbb{R})$ such that $\eta(w) = 0$ for $|w| > 1$, $\eta(w) \geq 0$ for all $w \in \mathbb{R}$, and $\int_{\mathbb{R}} \eta(w) dw = 1$. Let $\eta_j(w) = j\eta(jw)$ for $w \in \mathbb{R}$ and $j \in \mathbb{N}$. The functions $u_i^{[j]}$ are defined by $u_i^{[j]}(x, y) \equiv (u_{i,y} * \eta_j)(x)$ for all $(x, y) \in \mathbb{R} \times [0, \infty[$, $j \in \mathbb{N}$ and $i \in \{1, 2\}$.

We have from Proposition 6.5.5 that $u_{i,y}^{[j]} \in C^\infty(\mathbb{R})$ for all $y \geq 0$, $j \in \mathbb{N}$ and $i \in \{1, 2\}$. In particular, this implies that $u_i^{[j]}$ is continuous with respect to its first variable for all $j \in \mathbb{N}$ and $i \in \{1, 2\}$.

We now show that, for any $y_1, y_2 \in [0, \infty[$, we have that $u_{i,y_1}^{[j]} \rightarrow u_{i,y_2}^{[j]}$ uniformly on any compact set $K \subset \mathbb{R}$ as $y_1 \rightarrow y_2$ for all $j \in \mathbb{N}$ and $i \in \{1, 2\}$. We have by hypothesis that

$$\int_{K+[-1,1]} |u_{i,y_1}(x) - u_{i,y_2}(x)| dx \rightarrow 0 \quad \text{as } y_1 \rightarrow y_2$$

because $K + [-1, 1] = \{x + y : x \in K \text{ and } y \in [-1, 1]\}$ is a compact set. Therefore, given $\epsilon > 0$, there exists $\delta > 0$ such that

$$\int_{K+[-1,1]} |u_{i,y_1}(x) - u_{i,y_2}(x)| dx < \frac{\epsilon}{\max_{x \in \mathbb{R}} \eta(x)}$$

for $|y_1 - y_2| < \delta$. It follows that

$$\begin{aligned} |u_{i,y_1}^{[j]}(x) - u_{i,y_2}^{[j]}(x)| &= \left| \int_{\mathbb{R}} (u_i(x-s, y_1) - u_i(x-s, y_2)) \eta_j(s) ds \right| \\ &= \left| \int_{\mathbb{R}} (u_i(x-t/j, y_1) - u_i(x-t/j, y_2)) \eta(t) dt \right| \\ &= \left| \int_{|t| \leq 1} (u_i(x-t/j, y_1) - u_i(x-t/j, y_2)) \eta(t) dt \right| \\ &\leq \max_{x \in \mathbb{R}} \eta(x) \int_{|t| \leq 1} |u_i(x-t/j, y_1) - u_i(x-t/j, y_2)| dt \\ &\leq \max_{x \in \mathbb{R}} \eta(x) \int_{K+[-1,1]} |u_i(s, y_1) - u_i(s, y_2)| ds < \epsilon \end{aligned}$$

for all $x \in K$ if $|y_1 - y_2| < \delta$. This obviously implies that $u_i^{[j]}$ is continuous with respect to its second variable for all $j \in \mathbb{N}$ and $i \in \{1, 2\}$. But this extra uniform convergence on compact sets also ensures that $u_i^{[j]} \in C(\mathbb{R} \times [0, \infty[)$.

ii) We prove that $\{u_i^{[j]}\}_{j=0}^\infty$ converges to u_i in $L^1(K)$ for all compact sets $K \subset \mathbb{R} \times [0, \infty[$ and $i \in \{1, 2\}$.

We have from Proposition 6.5.3 that $\{u_{i,y}^{[j]}\}_{j=0}^\infty$ converges uniformly on compact sets to $u_{i,y}$ for all $y \geq 0$ and $i \in \{1, 2\}$ ⁵. Moreover, it follows from Theorem 6.5.2 and hypothesis (1) that $\|u_{i,y}^{[j]}\|_\infty \leq \|u_{i,y}\|_\infty \|\eta\|_1 = \|u_{i,y}\|_\infty \leq H$ for all $y \geq 0$, $j \in \mathbb{N}$ and $i \in \{1, 2\}$.

⁵To be precise, we have to slightly adapt the case for $L^\infty(\mathbb{R}^n)$ in Proposition 6.5.3. We do not need $f \in L^\infty(\mathbb{R}^n)$ as long as f is uniformly continuous on compact sets and ϕ has a compact support. In the proof, it suffices to take W large enough to have $\text{supp } \phi \subset W$.

Hence, we get from the Lebesgue Dominated Convergence Theorem that

$$\iint_K |u_i^{[j]}(x, y) - u_i(x, y)| \, dx \, dy \rightarrow 0 \quad \text{as } j \rightarrow \infty$$

because $|u_{i,y}^{[j]}(x) - u_{i,y}(x)| \rightarrow 0$ for all $(x, y) \in K \subset \mathbb{R} \times [0, \infty[$ and $|u_{i,y}^{[j]}(x) - u_{i,y}(x)| \leq g(x, y) \equiv 2H$ for all $(x, y) \in K$ with $g \in L^1(K)$ since K is bounded.

iii) Given $\phi \in \mathcal{D}(\mathbb{R} \times]0, \infty[)$, we choose a particular solution of (13.4.5). Choose $Y > 0$ large enough to have $\text{supp } \phi \subset \mathbb{R} \times]0, Y[$. Let

$$\psi^{[j]}(x, y) = \int_Y^y \phi(x^{[j]}(s, x, y), s) \, ds, \quad (13.4.7)$$

where $x^{[j]}$ is the solution of the ordinary differential equation

$$\frac{dx^{[j]}}{ds}(s, x, y) = \mathcal{G}^{[j]}(x^{[j]}(s, x, y), s) \quad (13.4.8)$$

with $x^{[j]}(y, x, y) = x$. Such a unique solution exists because $u_i^{[j]} \in C(\mathbb{R} \times [0, \infty[)$ and $u_{i,y}^{[j]} \in C^1(\mathbb{R})$ for all $y \geq 0$ imply that $\mathcal{G}^{[j]} \in C(\mathbb{R} \times [0, \infty[)$ and $\mathcal{G}_y^{[j]} \in C^1(\mathbb{R})$ for all $y \geq 0$ and $j \in \mathbb{N}$. Moreover, this implies that $(s, x, y) \mapsto x^{[j]}(s, x, y)$ is differentiable on $]0, \infty[\times \mathbb{R} \times]0, \infty[$ (Theorems 3.1 and 3.3 in Chapter 1 of [11]).

To prove that $\psi^{[j]}$ is a solution of (13.4.5), we note that

$$\begin{aligned} & \frac{\partial \psi^{[j]}}{\partial y}(x, y) + \mathcal{G}^{[j]}(x, y) \frac{\partial \psi^{[j]}}{\partial x}(x, y) \\ &= \phi(x^{[j]}(y, x, y), y) + \int_Y^y \frac{\partial \phi}{\partial x}(x^{[j]}(s, x, y), s) \frac{\partial x^{[j]}}{\partial y}(s, x, y) \, ds \\ & \quad + \mathcal{G}^{[j]}(x, y) \int_Y^y \frac{\partial \phi}{\partial x}(x^{[j]}(s, x, y), s) \frac{\partial x^{[j]}}{\partial x}(s, x, y) \, ds \\ &= \phi(x, y) + \underbrace{\int_Y^y \frac{\partial \phi}{\partial x}(x^{[j]}(s, x, y), s) \left(\frac{\partial x^{[j]}}{\partial y}(s, x, y) + \mathcal{G}^{[j]}(x, y) \frac{\partial x^{[j]}}{\partial x}(s, x, y) \right) \, ds}_{=0} \\ &= \phi(x, y) \end{aligned}$$

for $(x, y) \in \mathbb{R} \times]0, \infty[$, where the fact that the integrant in the last integral is null comes from equation (3.6) in the statement of Theorem 3.3 in Chapter 1 of [11].

iv) We prove that there exists a bounded set $R \subset \mathbb{R} \times [0, \infty[$ such that $\text{supp } \psi^{[j]} \subset R$ for all $j \in \mathbb{N}$.

Since $\|u_{i,y}^{[j]}\|_\infty \leq H$ for all $j \in \mathbb{N}$ and $i \in \{1, 2\}$, we get that

$$|\mathcal{G}^{[j]}(x, y)| = |\mathcal{G}(u_1^{[j]}(x, y), u_2^{[j]}(x, y))| \leq M \quad (13.4.9)$$

for all $(x, y) \in \mathbb{R} \times [0, \infty[$ and all $j \in \mathbb{N}$. Let $S = \text{supp } \phi$ for ϕ chosen initially. Since S is a compact, and so bounded, subset of $\mathbb{R} \times]0, \infty[$, we can find $a < b$ such that

$$S \subset R = \{(x, y) : 0 < y < Y \text{ and } y/M + a < x < -y/M + b\}.$$

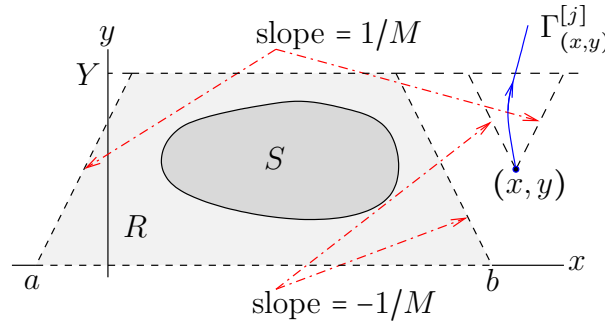


Figure 13.10: Figure associated to part (iv) of the proof of Theorem 13.4.5.

To prove that $\text{supp } \psi^{[j]} \subset R$, we prove that $\psi^{[j]}(x, y) = 0$ for all $(x, y) \notin R$. If $y \geq Y$, we have from (13.4.7) that $\psi^{[j]}(x, y) = 0$ for all $x \in \mathbb{R}$ because $\phi(x, y) = 0$ for all $x \in \mathbb{R}$ and $y \geq Y$. We only have to consider $y < Y$. It follows from (13.4.9) that any orbit $\Gamma_{(x,y)}^{[j]} \equiv \{(x^{[j]}(s, x, y), s) : s > 0\}$ of (13.4.8) that starts at $(x, y) \notin R$ does not intersect R and eventually crosses the line $y = Y$ (Figure 13.10). We therefore have that $\phi = 0$ along the orbits $\Gamma_{(x,y)}^{[j]}$ with $(x, y) \notin R$. It then follows from (13.4.5) that

$$\begin{aligned} \frac{\partial}{\partial s} \psi^{[j]}(x^{[j]}(s, x, y), s) &= \frac{\partial \psi^{[j]}}{\partial y}(x^{[j]}(s, x, y), s) + \frac{\partial \psi^{[j]}}{\partial x}(x^{[j]}(s, x, y), s) \frac{\partial x^{[j]}(s, x, y)}{\partial s} \\ &= \frac{\partial \psi^{[j]}}{\partial y}(x^{[j]}(s, x, y), s) + \mathcal{G}^{[j]}(x^{[j]}(s, x, y), s) \frac{\partial \psi^{[j]}}{\partial x}(x^{[j]}(s, x, y), s) \\ &= \phi(x^{[j]}(s, x, y), s) = 0 \end{aligned} \quad (13.4.10)$$

for $s \geq 0$. Thus $\psi^{[j]}$ is constant along the orbits $\Gamma_{(x,y)}^{[j]}$ with $(x, y) \notin R$. Since $\psi^{[j]}(x^{[j]}(Y, x, y), Y) = 0$, we get that $\psi^{[j]} = 0$ along the orbits $\Gamma_{(x,y)}^{[j]}$ with $(x, y) \notin R$. Thus, $\psi^{[j]}(x^{[j]}(y, x, y), y) = \psi^{[j]}(x, y) = 0$ for $(x, y) \notin R$.

v) We prove that $\{\mathcal{G}^{[j]}\}_{j=0}^{\infty}$ converges to $\mathcal{G}$ in $L^1(K)$ for all compact sets $K \subset \mathbb{R} \times [0, \infty[$.

We have that

$$\mathcal{G}(p, q) = \int_0^1 G'(tq + (1-t)p) dt \quad (13.4.11)$$

because

$$\int_0^1 G'(tq + (1-t)p) dt = \frac{1}{q-p} \int_p^q G'(r) dr = \frac{G(q) - G(p)}{q-p}$$

for $p \neq q$ and

$$\int_0^1 G'(tq + (1-t)p) dt = \int_0^1 G'(p) dt = G'(p)$$

for $p = q$. Hence,

$$\mathcal{G}^{[j]}(x, y) - \mathcal{G}(x, y) = \mathcal{G}(u_1^{[j]}(x, y), u_2^{[j]}(x, y)) - \mathcal{G}(u_1(x, y), u_2(x, y))$$

$$\begin{aligned}
&= \int_0^1 \left(G'(tu_2^{[j]}(x, y) + (1-t)u_1^{[j]}(x, y)) - G'(tu_2(x, y) + (1-t)u_1(x, y)) \right) dt \\
&= \int_0^1 \left(G''(\xi(x, y)) \left(t(u_2^{[j]}(x, y) - u_2(x, y)) + (1-t)(u_1^{[j]}(x, y) - u_1(x, y)) \right) \right) dt
\end{aligned}$$

for $\xi(x, y)$ between $tu_2^{[j]}(x, y) + (1-t)u_1^{[j]}(x, y)$ and $tu_2(x, y) + (1-t)u_1(x, y)$. Since $|\xi(x, y)| \leq H$ for all (x, y) because $\|u_i\|_\infty \leq H$ and $\|u_i^{[j]}\|_\infty \leq H$ for all $j \in \mathbb{N}$ and $i \in \{1, 2\}$, we get that $|G''(\xi(x, y))| \leq \tilde{M} \equiv \max_{p \in D} G''(p) < \infty$ for all (x, y) . The previous maximum is finite because G'' is continuous on the compact set D . Moreover, recall that $m = \min_{p \in D} G''(p) > 0$ by hypothesis. Thus,

$$\begin{aligned}
|\mathcal{G}^{[j]}(x, y) - \mathcal{G}(x, y)| &\leq \tilde{M} \int_0^1 \left(t |u_2^{[j]}(x, y) - u_2(x, y)| + (1-t) |u_1^{[j]}(x, y) - u_1(x, y)| \right) dt \\
&= \frac{\tilde{M}}{2} \left(|u_2^{[j]}(x, y) - u_2(x, y)| + |u_1^{[j]}(x, y) - u_1(x, y)| \right).
\end{aligned}$$

Hence,

$$\begin{aligned}
&\iint_K |\mathcal{G}^{[j]}(x, y) - \mathcal{G}(x, y)| \, dx \, dy \\
&\leq \frac{\tilde{M}}{2} \left(\iint_K |u_2^{[j]}(x, y) - u_2(x, y)| \, dx \, dy + \iint_K |u_1^{[j]}(x, y) - u_1(x, y)| \, dx \, dy \right) \rightarrow 0
\end{aligned}$$

as $j \rightarrow \infty$ because of (ii) above.

vi) We now address the convergence of the last integral in (13.4.6). The major issue is the behaviour of $\frac{\partial \psi^{[j]}}{\partial x}$.

vi.a) Given $\alpha > 0$, let $h_i(x, y) \equiv u_i(x, y) - (C/\alpha)g(x, y)$ where $g(x, y) = x$ for all $(x, y) \in \mathbb{R}^2$ and $i \in \{1, 2\}$. It follows from (13.4.2) that

$$\begin{aligned}
h_i(x+s, y) - h_i(x, y) &= u_i(x+s, y) - \frac{C}{\alpha}g(x+s) - u_i(x, y) + \frac{C}{\alpha}g(x) \\
&= s \left(\frac{u_i(x+s, y) - u_i(x, y)}{s} \right) - \frac{Cs}{\alpha} \leq \frac{Cs}{y} - \frac{Cs}{\alpha} = Cs \left(\frac{1}{y} - \frac{1}{\alpha} \right) \leq 0
\end{aligned}$$

for $y \geq \alpha$ and $i \in \{1, 2\}$. Thus $h_i(x, y)$ is non-increasing with respect to x for $y \geq \alpha$ and $i \in \{1, 2\}$. Hence,

$$u_i^{[j]}(x, y) - \frac{C}{\alpha}(\eta_j * g)(x, y) = \left(\eta_i * \left(u_i - \frac{C}{\alpha}g \right) \right)(x, y) = (\eta_i * h_i)(x, y)$$

is non-increasing with respect to x for $y \geq \alpha$ and $i \in \{1, 2\}$ ⁶. Therefore,

$$0 \geq \frac{\partial}{\partial x} \left(u_{i,y}^{[j]}(x) - \frac{C}{\alpha}(\eta_j * g)(x) \right) = \frac{\partial u_{i,y}^{[j]}(x)}{\partial x} - \frac{C}{\alpha}$$

⁶Suppose that $p \in L^1(\mathbb{R})$ and $q \in L^m(\mathbb{R})$ for $m \geq 1$. If q is non-increasing then

$$(p * q)(x_1) = \int_{-\infty}^{\infty} p(s)q(x_1 - s) \, ds \leq \int_{-\infty}^{\infty} p(s)q(x_2 - s) \, ds = (p * q)(x_2)$$

for $x_1 \geq x_2$.

for $x \in \mathbb{R}$, $y \geq \alpha$ and $i \in \{1, 2\}$. Thus

$$\frac{du_{i,y}^{[j]}(x)}{dx} \leq \frac{C}{\alpha} \quad (13.4.12)$$

for $x \in \mathbb{R}$, $y \geq \alpha$ and $i \in \{1, 2\}$.

It follows from (13.4.11) and (13.4.12) that

$$\begin{aligned} \frac{\partial \mathcal{G}^{[j]}}{\partial x}(x, y) &= \frac{\partial}{\partial x} \mathcal{G} \left(u_1^{[j]}(x, y), u_2^{[j]}(x, y) \right) \\ &= \int_0^1 G'' \left(tu_2^{[j]}(x, y) + (1-t)u_1^{[j]}(x, y) \right) \left(t \frac{\partial u_2^{[j]}}{\partial x}(x, y) + (1-t) \frac{\partial u_1^{[j]}}{\partial x}(x, y) \right) dt \\ &\leq \frac{C\tilde{M}}{\alpha} \end{aligned} \quad (13.4.13)$$

for $x \in \mathbb{R}$, $y \geq \alpha$ and $j \in \mathbb{N}$ because $tu_2^{[j]}(x, y) + (1-t)u_1^{[j]}(x, y) \in D$ for all $(x, y) \in \mathbb{R} \times [0, \infty[$ and $t \in [0, 1]$, and $m = \min_{u \in D} G''(u) > 0$ by hypothesis.

Let $w_j(s, x, y) = \frac{\partial x^{[j]}}{\partial s}(s, x, y)$ for $s \in]0, \infty[$, $(x, y) \in \mathbb{R} \times]0, \infty[$ and $j \in \mathbb{N}$.

Since $x^{[j]}(y, x, y) = x$ for all $(x, y) \in \mathbb{R} \times]0, \infty[$ and $j \in \mathbb{N}$, we get that $w_j(y, x, y) = 1$ for all $(x, y) \in \mathbb{R} \times]0, \infty[$ and $j \in \mathbb{N}$. Moreover,

$$\begin{aligned} \frac{\partial w_j}{\partial s}(s, x, y) &= \frac{\partial}{\partial s} \left(\frac{\partial x^{[j]}}{\partial s}(s, x, y) \right) = \frac{\partial}{\partial x} \left(\frac{\partial x^{[j]}}{\partial s}(s, x, y) \right) = \frac{\partial}{\partial x} \left(\mathcal{G}^{[j]}(x^{[j]}(s, x, y), s) \right) \\ &= \frac{\partial \mathcal{G}^{[j]}}{\partial x}(x^{[j]}(s, x, y), s) \frac{\partial x^{[j]}}{\partial s}(s, x, y) = \frac{\partial \mathcal{G}^{[j]}}{\partial x}(x^{[j]}(s, x, y), s) w_j(s, x, y) \end{aligned}$$

for all $(x, y) \in \mathbb{R} \times]0, \infty[$ and $j \in \mathbb{N}$. The solution of

$\frac{\partial w_j}{\partial s}(s, x, y) = \frac{\partial \mathcal{G}^{[j]}}{\partial x}(x^{[j]}(s, x, y), s) w_j(s, x, y)$ with the initial condition $w_j(y, x, y) = 1$ is

$$w_j(s, x, y) = \exp \left(\int_y^s \frac{\partial \mathcal{G}^{[j]}}{\partial x}(x^{[j]}(t, x, y), t) dt \right)$$

for all $j \in \mathbb{N}$.

Since we are only interested in the values of $\psi^{[j]}(x, y)$ for $(x, y) \in R$. We only need to consider $s \leq Y$. Moreover, we assume that $\alpha < Y$ and consider only $\alpha \leq s \leq Y$. We then get from (13.4.13) that $|w_j(s, x, y)| \leq e^{C\tilde{M}|s-y|/\alpha} \leq e^{C\tilde{M}(Y-\alpha)/\alpha}$ for $\alpha \leq s, y \leq Y$.

We get from (13.4.7) that

$$\frac{\partial \psi^{[j]}}{\partial x}(x, y) = \int_Y^y \frac{\partial \phi}{\partial x}(x^{[j]}(s, x, y), s) \frac{\partial x^{[j]}}{\partial s}(s, x, y) ds = \int_Y^y \frac{\partial \phi}{\partial x}(x^{[j]}(s, x, y), s) w_j(s, x, y) ds .$$

Thus, for $\alpha \leq y \leq Y$, we have that

$$\left| \frac{\partial \psi^{[j]}}{\partial x}(x, y) \right| \leq (Y - \alpha) P e^{C\tilde{M}(Y-\alpha)/\alpha} \quad (13.4.14)$$

for all $x \in \mathbb{R}$ and $j \in \mathbb{N}$, where $P = \max_{(x,y) \in \mathbb{R} \times [0, \infty[} \left| \frac{\partial \phi}{\partial x}(x, y) \right|$. Note that (13.4.14) is also true for $y > Y$ because $\psi^{[j]}(x, y) = 0$ for $(x, y) \in \mathbb{R} \times [Y, \infty[$ and $j \in \mathbb{N}$.

vi.b) We now consider $\frac{\partial \psi^{[j]}}{\partial x}(x, y)$ for $0 \leq y \leq \alpha$, $x \in \mathbb{R}$ and $j \in \mathbb{N}$. Let $V(f) \equiv \int_{\mathbb{R}} |f'(x)| dx$ for all functions $f \in C^1(\mathbb{R})$. Since $\text{supp } \psi^{[j]} \subset R$ for all $j \in \mathbb{N}$, we have that $V(\psi_y^{[j]}) < \infty$ for all $y \geq 0$ and $j \in \mathbb{N}$.

We show that there exist a positive integer N and a constant Q such that $V(\psi_y^{[j]}) \leq Q$ for all $0 < y \leq 1/N$ and $j \in \mathbb{N}$. Since $\text{supp } \phi \in \mathcal{D}(\mathbb{R} \times]0, \infty[)$, there exists a positive integer N such that $\phi(x, y) = 0$ for $(x, y) \in \mathbb{R} \times]0, 1/N]$. It then follows from (13.4.5) that

$$\frac{\partial \psi^{[j]}}{\partial y}(x, y) + \mathcal{G}^{[j]}(x, y) \frac{\partial \psi^{[j]}}{\partial x}(x, y) = 0$$

for $(x, y) \in \mathbb{R} \times]0, 1/N]$. Proceeding as we did in (13.4.10), we find that $\psi^{[j]}$ is constant along the section $\{(x^{[j]}(s, x, y), s) : 0 < s \leq 1/N\}$ of the orbit $\Gamma_{(x,y)}^{[j]}$ of (13.4.8) for all $(x, y) \in \mathbb{R} \times]0, 1/N]$ (Figure 13.11).

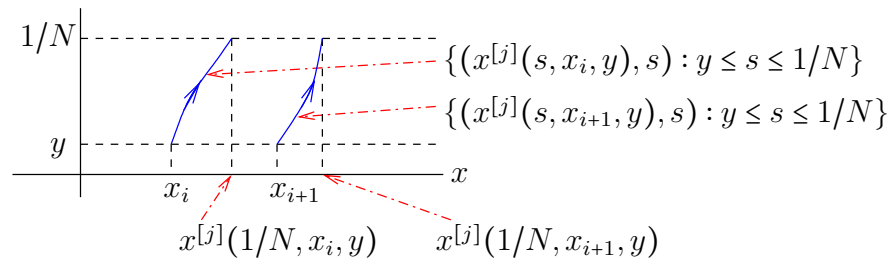


Figure 13.11: Figure associated to part (vi.b) of the proof of Theorem 13.4.5.

We now show that $V(\psi_y^{[j]}) = V(\psi_{1/N}^{[j]})$ for all $0 < y < 1/N$. Given any sequence $x_1 < x_2 < \dots < x_n$ in $\mathbb{R}$, we have that

$$\sum_{i=1}^{n-1} \left| \psi_y^{[j]}(x_{i+1}) - \psi_y^{[j]}(x_i) \right| = \sum_{i=1}^{n-1} \left| \frac{\partial \psi^{[j]}}{\partial x}(z_i, y) \right| (x_{i+1} - x_i)$$

for some $z_i \in]x_i, x_{i+1}[$ for $1 \leq i < N$. We also have that

$$\begin{aligned} \sum_{i=1}^{n-1} \left| \psi_y^{[j]}(x_{i+1}) - \psi_y^{[j]}(x_i) \right| &= \sum_{i=1}^{n-1} \left| \psi_y^{[j]}(x^{[j]}(1/N, x_{i+1}, y)) - \psi_y^{[j]}(x^{[j]}(1/N, x_i, y)) \right| \\ &= \sum_{i=1}^{n-1} \left| \frac{\partial \psi^{[j]}}{\partial x}(w_i, y) \right| (x^{[j]}(1/N, x_{i+1}, y) - x^{[j]}(1/N, x_i, y)) \end{aligned}$$

for some $w_i \in]x^{[j]}(s_{1/n, x_i, y}, x_i, y), x^{[j]}(s_{1/n, x_{i+1}, y}, x_{i+1}, y)[$ for $1 \leq i < N$.

Taking the supremum over all possible sequences $x_1 < x_2 < \dots < x_n$ in $\mathbb{R}$ in the two previous series, we find that $V(\psi_y^{[j]}) = V(\psi_{1/N}^{[j]})$ for $0 < y < 1/N$ by definition of the integral.

We take $\alpha < 1/N$. Let $R_{\alpha,+} = \{(x, y) \in R : y \geq \alpha\}$ and $R_{\alpha,-} = \{(x, y) \in R : y < \alpha\}$. We have from (13.4.6) that

$$\begin{aligned} \left| \int_0^\infty \int_{-\infty}^\infty (u_2 - u_1) \phi \, dy \, dx \right| &= \left| \int_0^\infty \int_{-\infty}^\infty (u_2 - u_1) (\mathcal{G}^{[j]} - \mathcal{G}) \frac{\partial \psi^{[j]}}{\partial x} \, dy \, dx \right| \\ &\leq \iint_{R_{\alpha,+}} \left| (u_2 - u_1) (\mathcal{G}^{[j]} - \mathcal{G}) \frac{\partial \psi^{[j]}}{\partial x} \right| \, dy \, dx + \iint_{R_{\alpha,-}} \left| (u_2 - u_1) (\mathcal{G}^{[j]} - \mathcal{G}) \frac{\partial \psi^{[j]}}{\partial x} \right| \, dy \, dx \end{aligned}$$

Given $\epsilon > 0$, we can use (13.4.14) and (v) to choose a positive integer J such that

$$\iint_{R_{\alpha,+}} \left| (u_2 - u_1) (\mathcal{G}^{[j]} - \mathcal{G}) \frac{\partial \psi^{[j]}}{\partial x} \right| \, dy \, dx \leq 2H(Y - \alpha) P e^{CM(Y-\alpha)} \iint_{R_{\alpha,+}} |\mathcal{G}^{[j]} - \mathcal{G}| \, dx \, dy < \frac{\epsilon}{2}$$

for $j > J$. Moreover,

$$\begin{aligned} \iint_{R_{\alpha,-}} \left| (u_2 - u_1) (\mathcal{G}^{[j]} - \mathcal{G}) \frac{\partial \psi^{[j]}}{\partial x} \right| \, dy \, dx &\leq 4HM \iint_{R_{\alpha,-}} \left| \frac{\partial \psi^{[j]}}{\partial x}(x, y) \right| \, dy \, dx \\ &\leq 4HM \int_0^\alpha \left(\int_{\mathbb{R}} \left| \frac{\partial \psi_y^{[j]}}{\partial x}(x) \right| \, dx \right) \, dy = 4HM \int_0^\alpha V(\psi_y^{[j]}) \, dy \leq 4HM \alpha V(\psi_{1/N}^{[j]}) \leq \frac{\epsilon}{2} \end{aligned}$$

if we assume that $\alpha < \min\{1/N, 1/(4H M V(\psi_{1/N}^{[j]}))\}$. Thus,

$$\left| \int_0^\infty \int_{-\infty}^\infty (u_2 - u_1) \phi \, dy \, dx \right| < \epsilon .$$

Since ϵ is arbitrary, we get that

$$\int_0^\infty \int_{-\infty}^\infty (u_2 - u_1) \phi \, dy \, dx = 0 .$$

This finally completes the proof. ■

Adding some conditions to those of the previous theorem, it is also possible to prove that the strong solution of (13.2.1) depends continuously on the initial condition; namely,

$$\int_{x_1}^{x_2} |u(x, y) - \tilde{u}(x, y)| \, dx \leq \int_{x_1 - My}^{x_2 + My} |h(x) - \tilde{h}(x)| \, dx$$

for $x_1 < x_2$ in $\mathbb{R}$, where $\tilde{u}$ is the strong solution with the initial condition $\tilde{u}(x, 0) = \tilde{h}(x)$ for $x \in \mathbb{R}$. Consult [35] for the full statement and the proof of this result.

To motivate (13.4.2), we consider (13.1.1) with the initial condition $u(x, 0) = h(x)$ for $x \in \mathbb{R}$. Its solution is given implicitly by (13.1.2) as long as no x, y characteristics intersect. If we assume that $u \in C^1(\mathbb{R} \times]0, \infty[)$ and $g, h \in C^1(\mathbb{R})$ and derive both sides of (13.1.2) with respect to x , we get

$$\frac{\partial u}{\partial x} = h'(x - g(u) y) \left(1 - y g'(u) \frac{\partial u}{\partial x} \right) .$$

Solving for $\frac{\partial u}{\partial x}$, we find that

$$\frac{\partial u}{\partial x} = \frac{h'(x - g(u)y)}{1 + yg'(u)h'(x - g(u)y)}.$$

If $g'(u) \geq m > 0$ for all $u \in \mathbb{R}$ and $h'(s) > 0$ for all $s \in \mathbb{R}$, then $\frac{\partial u}{\partial x} \leq \frac{C}{y}$ with $C = 1/m$. Thus, (13.4.2) is certainly satisfied if s is small.

(13.4.2) is an **entropy condition**. It ensures that as x goes from ∞ to ∞ with y fixed, u “drops” at the points where u is discontinuous. This is an important consequence of (13.4.2). Suppose that u is a strong solution which discontinuous along the curve $\Gamma = \{(\gamma(y), y) : y \geq y_0\}$ for some $y_0 \geq 0$. Then $u^-(\mathbf{p}) \equiv \lim_{x \rightarrow \gamma(y)} u(x, y) > \lim_{x \rightarrow \gamma(y)} u(x, y) \equiv u^+(\mathbf{p})$, where $\mathbf{p} = (\gamma(y), y)$. If we assume that $G''(u) > 0$ for all u , then G' is increasing and

$$G'(u^+(\mathbf{p})) < \frac{G(u^+(\mathbf{p})) - G(u^-(\mathbf{p}))}{u^+(\mathbf{p}) - u^-(\mathbf{p})} < G'(u^-(\mathbf{p})).$$

It follows from Rankine-Hugoniot formula (13.3.6) that

$$G'(u^+(\mathbf{p})) < \gamma'(\mathbf{p}) < G'(u^-(\mathbf{p}));$$

namely, the Lax Shock Condition (13.4.1) is satisfied.

(13.4.2) is necessary to prove the convergence of the finite difference scheme in the proof of existence of a solution in Theorem 13.4.5. As we also saw, it is necessary in the proof of the uniqueness of the solution.

13.4.4 Asymptotic Behaviour

This is an important topic that we initially did not plan to include because of the size of the material associated to this topic. However, we felt that we cannot at least state one of its most non-intuitive result. The reader should consult [18, 35] for more information on this subject. We simply quote some results from [18].

Theorem 13.4.6

Suppose that $G \in C^3(\mathbb{R})$ and $G''(u) > 0$ for all $u \in \mathbb{R}$. Moreover, suppose that h has a compact support. If u is a strong solution of (13.2.1) satisfying $u(x, 0) = h(x)$ for all $x \in \mathbb{R}$, and u satisfies Lax Shock Condition, Definition 13.4.3, along all curves Γ (if any) where it is discontinuous, then $\sup_{x \in \mathbb{R}} |u(x, y)| = O(1/\sqrt{y})$ for all $y > 0$. In fact, $\sqrt{y} \sup_{x \in \mathbb{R}} |u(x, y)|$ converges to a constant as $y \rightarrow \infty$.

If u is a periodic solution of (13.2.1) of period P , we have that $\int_0^P u(x, y) dx$ is constant with respect to y because the total variation of $x \mapsto u(x, y)$ is constant on intervals of same length⁷. So, we may define the mean value of u as $M_u = \frac{1}{P} \int_0^P u(x, y) dx$ whatever $y > 0$.

⁷In fact, this is the fundamental principle used to deduce the conservation law (13.2.1). See [18].

Theorem 13.4.7

Suppose that $G \in C^3(\mathbb{R})$ and $G''(u) > k > 0$ for all $u \in \mathbb{R}$ and some constant k . If u is a strong solution of (13.2.1) satisfies Lax Shock Condition, Definition 13.4.3, along all curves Γ (if any) where it is discontinuous, and $x \mapsto u(x, y)$ is periodic of period P for all $y > 0$, then $|u(x, y) - M_u| \leq 1/(ky)$ for all $y > 0$. In fact, $|u(x, y) - M_u||y| \rightarrow 1/k$ as $y \rightarrow \infty$.

Suppose that u_1 is the solution described in Theorem 13.4.6, where the compact support of h is included in the interval $[0, P]$. Suppose that u_2 is the solution of period P described in Theorem 13.4.7, where $u_2(x, 0) = h(x)$ for $x \in [0, P]$. We have the surprising result that u_2 decays faster than u_1 despite the fact that u_2 is subject to a much larger initial disturbance than u_1 .

13.5 Systems of Conservation Laws

As the title suggests, we now consider systems of conservation laws. The following example is the simplest and standard example that still can illustrate all the complexity of solving systems of conservation laws. As we say, it is a must-have.

Example 13.5.1 (The p-system)

Consider the system

$$\begin{aligned} \frac{\partial u_1}{\partial y} - \frac{\partial u_2}{\partial x} &= 0 \\ \frac{\partial u_2}{\partial y} + \frac{\partial}{\partial x} p(u_1) &= 0 \end{aligned} \tag{13.5.1}$$

where $u_1, u_2 : \mathbb{R} \times [0, \infty[\rightarrow \mathbb{R}$ and $p(v) = kv^{-\beta}$ for some constants k and β . This is a model for isentropic⁸ gas dynamics in Lagrangian coordinates⁹. For the readers who know the full equations of conservation of mass, momentum and energy in gas dynamics, these two equations represent the conservation of mass and momentum only, while energy is added to the system and the temperature is held constant. u_1 is the specific volume (i.e. the inverse of the density), u_2 is the velocity and β is the adiabatic gas constant which is generally between 1 and 3.

We may also have initial conditions of the form $u_1(x, 0) = h_1(x)$ and $u_2(x, 0) = h_2(x)$ for $x \in \mathbb{R}$, where $h_1, h_2 : \mathbb{R} \rightarrow \mathbb{R}$ are given functions.

Assuming that the functions are sufficiently differentiable, we may rewrite (13.5.1) as a second order partial differential equation. If we derive $\frac{\partial u_2}{\partial x} = \frac{\partial u_1}{\partial y}$ with respect to y and

⁸Isentropic means that the entropy is constant. In gas dynamics, this concretely means that the pressure is an increasing function of the density.

⁹In Euclidean coordinates, the origin is a fixed point in space. In Lagrangian coordinates, the origin is associated to the position of a particle as it moves in space. We have a moving frame of coordinates.

$\frac{\partial u_2}{\partial y} + \frac{\partial}{\partial x} p(u_1) = 0$ with respect to x , we get $\frac{\partial^2 u_2}{\partial x \partial y} = \frac{\partial^2 u_1}{\partial y^2}$ and $\frac{\partial^2 u_2}{\partial y \partial x} + \frac{\partial^2}{\partial x^2} p(u_1) = 0$ respectively. Thus $\frac{\partial^2 u_1}{\partial y^2} + \frac{\partial^2}{\partial x^2} p(u_1) = 0$ with $u_1(x, 0) = h_1(x)$ and $\frac{\partial u_1}{\partial y}(x, 0) = h'_2(x)$ for $x \in \mathbb{R}$. This is a highly nonlinear partial differential equation.

Another way to reduce (13.5.1) to second order partial differential equation is to realize that the equation $\frac{\partial u_1}{\partial y} = \frac{\partial u_2}{\partial x}$ implies that $u_2 = \frac{\partial w}{\partial y}$ and $u_1 = \frac{\partial w}{\partial x}$ for some function $w : \mathbb{R} \times [0, \infty[$. If we substitute these expressions in (13.5.1), we get $\frac{\partial^2 w}{\partial y^2} + \frac{\partial}{\partial x} p\left(\frac{\partial w}{\partial x}\right) = \frac{\partial^2 w}{\partial y^2} + \frac{dp}{dv}\left(\frac{\partial w}{\partial x}\right) \frac{\partial^2 w}{\partial x^2} = 0$. This is a **nonlinear wave equation** because the speed of the wave is determined by the square root of the nonlinear function $-\frac{dp}{dv}\left(\frac{\partial w}{\partial x}\right)$. This is again a highly nonlinear partial differential equation.

Transforming (13.5.1) to second order partial differential equation does not really help us in solving (13.5.1). Moreover, remember what happened to the Burgers' equation in Example 13.3.2 after we multiplied it by u . We should expect similar complications with the two transformations above. ♣

Before diving into the analysis of the p -system (13.5.1), we need to expand the Rankine-Hugoniot formula and the Lax Shock Condition to systems of conservation laws.

We consider the system

$$\frac{\partial u}{\partial y} + \frac{\partial G(u)}{\partial x} = \frac{\partial u}{\partial y} + D G(u) \frac{\partial u}{\partial x} = \mathbf{0}, \quad (13.5.2)$$

where $u : \mathbb{R} \times [0, \infty[\rightarrow \mathbb{R}^n$ and $G : \mathbb{R}^n \rightarrow \mathbb{R}^n$ are sufficiently differentiable functions.

If we assume that u is discontinuous along the curve $\Gamma = \{(\gamma(y), y) : y \geq y_0\}$ of class C^1 for some $y_0 \geq 0$, we can proceed on each component of u as we did in Section 13.3 to obtain the **Rankine-Hugoniot formula** for systems of conservation laws

$$G(u^+(\mathbf{p})) - G(u^-(\mathbf{p})) = \frac{d\gamma}{dy}(y) (u^+(\mathbf{p}) - u^-(\mathbf{p})) \quad (13.5.3)$$

for all $\mathbf{p} = (\gamma(y), y) \in \Gamma$. This condition must be satisfied if we have a shock wave.

We now turn our attention to generalizing the Lax Shock Condition to systems of conservation laws.

From now on, we assume that (13.5.2) is a **strictly hyperbolic partial differential equation**; namely, $D G(\mathbf{u}) : \mathbb{R}^n \rightarrow \mathbb{R}^n$ has n distinct eigenvalues $\lambda_1(\mathbf{u}) < \lambda_2(\mathbf{u}) < \dots < \lambda_n(\mathbf{u})$ for all $\mathbf{u} \in \mathbb{R}^n$.

Definition 13.5.2 (Lax Shock Condition)

Suppose that the strong solution u of (13.5.2) is discontinuous along the curve $\Gamma = \{(\gamma(y), y) : y \geq y_0\}$ of class C^1 for some $y_0 \geq 0$. The strong solution u must satisfy the

Rankine-Hugoniot formula (13.5.3) and there must be a value $k \in \{1, 2, \dots, n\}$ such that

$$\begin{aligned}\lambda_{k-1}(u^-(\mathbf{p})) &< \gamma'(y) < \lambda_k(u^-(\mathbf{p})) \\ \lambda_k(u^+(\mathbf{p})) &< \gamma'(y) < \lambda_{k+1}(u^+(\mathbf{p}))\end{aligned}\tag{13.5.4}$$

for all $\mathbf{p} = (\gamma(y), y) \in \Gamma$, where $\lambda_0(\mathbf{w}) = -\infty$ and $\lambda_{n+1}(\mathbf{w}) = \infty$ for all $\mathbf{w}$.

A strong solution u that satisfies the Rankine-Hugoniot formula and the Lax Shock Condition above is called a **k-shock**.

It follows from (13.5.4) that

$$\begin{aligned}\lambda_k(u^+(\mathbf{p})) &< \gamma'(y) < \lambda_k(u^-(\mathbf{p})) \\ \lambda_{k-1}(u^-(\mathbf{p})) &< \gamma'(y) < \lambda_{k+1}(u^+(\mathbf{p}))\end{aligned}\tag{13.5.5}$$

The first condition in (13.5.5) is the condition that we will expect from (13.4.1). In the case where $n = 1$, we have that $\lambda_1(u) = \frac{dG}{du}(u)$. Hence, the definition of the Lax Shock Condition given in Definition 13.4.3 is included in the definition of the Lax Shock Condition given in Definition 13.5.2. In particular, (13.5.4) yields $\gamma'(y) < \lambda_1(u^-(\mathbf{p})) = \frac{\partial G}{\partial u}(u^-(\mathbf{p}))$ and $\frac{\partial G}{\partial u}(u^+(\mathbf{p})) = \lambda_1(u^+(\mathbf{p})) < \gamma'(y)$ for all $\mathbf{p} = (\gamma(y), y) \in \Gamma$.

To motivate the Lax Shock Condition above, one needs to use the fact that (13.5.2) is a strictly hyperbolic partial differential equation. Let us first consider a system of decoupled conservation laws

$$\frac{\partial v_i}{\partial y} + \lambda_i(v_i) \frac{\partial v_i}{\partial x} = 0\tag{13.5.6}$$

for $1 \leq i \leq n$, where $\lambda_1(\mathbf{v}) < \lambda_2(\mathbf{v}) < \dots < \lambda_n(\mathbf{v})$ for all $\mathbf{v} \in \mathbb{R}^n$ and $\lambda_i(\mathbf{v}) = \lambda_i(v_i)$ for all i . We assume that $v(x, 0) = \mathbf{v}^-$ for $x < 0$ and $v(x, 0) = \mathbf{v}^+$ for $x > 0$,

We have that v is possibly discontinuous along the line $\Gamma = \{(\eta y, y) : y \geq 0\}$ where η is given by the Rankine-Hugoniot formula $\lambda_i(\mathbf{v}^+) - \lambda_i(\mathbf{v}^-) = \lambda_i(v_i^+) - \lambda_i(v_i^-) = \eta(v_i^+ - v_i^-)$ for $1 \leq i \leq n$. The v_i are constant along (sections of) the x, y characteristics given by $x = \lambda_i(v_i(s, 0))y + s$ for $s \in \mathbb{R}$. Note that $\lambda_i(v) = \lambda(\mathbf{v}^-) = \lambda(v_i^-)$ on the left side of Γ and $\lambda_i(v) = \lambda(\mathbf{v}^+) = \lambda(v_i^+)$ on the right side of Γ .

Suppose that

$$\lambda_1(\mathbf{v}^+) < \lambda_2(\mathbf{v}^+) < \dots < \lambda_{k_r}(\mathbf{v}^+) < \eta < \lambda_{k_r+1}(\mathbf{v}^+) < \dots < \lambda_n(\mathbf{v}^+).\tag{13.5.7}$$

We have that v_i is completely determined by the initial condition v_i^+ on the right of Γ when $i \leq k_r$ (Figure 13.12 (a)). However, this not so for v_i with $i > k_r$ (Figure 13.12 (b)). There are $n - k_r$ conditions on Γ that need to be given.

Suppose that

$$\lambda_1(\mathbf{v}^-) < \lambda_2(\mathbf{v}^-) < \dots < \lambda_{k_l}(\mathbf{v}^-) < \eta < \lambda_{k_l+1}(\mathbf{v}^-) < \dots < \lambda_n(\mathbf{v}^-).\tag{13.5.8}$$

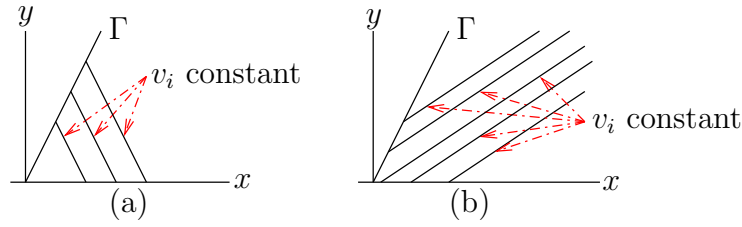


Figure 13.12: The solution for the system of conservation laws (13.5.6) on the right of Γ . The solutions for v_i with $i \leq k_r$ are represented in (a) and those for $i > k_r$ are represented in (b).

We have that v_i is completely determined by the initial condition v_i^- on the left of Γ when $i > k_l$ (Figure 13.13 (a)). However, this is not so for v_i with $i \leq k_l$ (Figure 13.13 (b)). There are k_l conditions on Γ that need to be given.

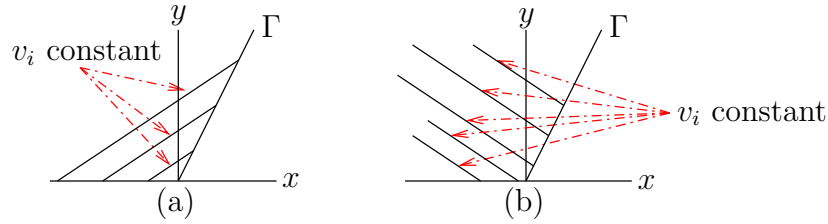


Figure 13.13: The solution for the system of conservation laws (13.5.6) on the left of Γ . The solutions for v_i with $i > k_l$ are represented in (a) and those for $i \leq k_l$ are represented in (b).

How could we select all the conditions to get a well defined shock wave? Moreover, is there a relation between k_l and k_r ? To answer these questions, we again refer to Rankine-Hugoniot Formula (13.5.3) to obtain n equations that must be satisfied; namely, $\lambda_i(\mathbf{v}^+) - \lambda_i(\mathbf{v}^-) = \eta(v_i^+ - v_i^-)$ for $1 \leq i \leq n$. We can isolate η to eliminate one of the equations. We therefore have $n - 1$ equations to produce a well defined solution. We must therefore have that $n - k_r + k_l = n - 1$. This yields $k \equiv k_r = k_l + 1$. Hence, (13.5.7) and (13.5.8) yield $\lambda_{k-1}(\mathbf{v}^-) < \eta < \lambda_k(\mathbf{v}^-)$ and $\lambda_k(\mathbf{v}^+) < \eta < \lambda_{k+1}(\mathbf{v}^+)$ as stated in (13.5.4).

To generalize the previous discussion to the strictly hyperbolic partial differential equation (13.5.2) with $n \geq 2$ requires a delicate analysis. In particular, the Lax Shock Condition is not the only approach used to address entropy in higher dimensional systems of conservation Laws. We will come back on this subject at the end of the section.

Example 13.5.3 (Example 13.5.1 continued)

The system of conservation laws (13.5.1) can be expressed in the form (13.5.2) with $G(\mathbf{u}) =$

$\begin{pmatrix} -u_2 \\ p(u_1) \end{pmatrix}$, where $p'(v) = -k\beta v^{-\beta-1} < 0$ and $p''(v) = k\beta(\beta+1)v^{-\beta-2} > 0$ for all $v \in \mathbb{R}$ ¹⁰.

We study the **Riemann problem** associated to (13.5.1). More precisely, we assume that $u(x, 0) = \mathbf{u}^-$ for $x < 0$ and $u(x, 0) = \mathbf{u}^+$ for $x > 0$.

The eigenvalues of $DG(\mathbf{u}) = \begin{pmatrix} 0 & -1 \\ p'(u_1) & 0 \end{pmatrix}$ are $\lambda_1(\mathbf{u}) = -\sqrt{-p'(u_1)}$ and $\lambda_2(\mathbf{u}) = \sqrt{-p'(u_1)}$.

According to our theory, there are two acceptable types of shock waves. As usual, let $\Gamma = \{(\eta y, y) : y \geq 0\}$ be the line along which u is discontinuous, where η is given by the Rankine-Hugoniot formula $G(\mathbf{u}^+) - G(\mathbf{u}^-) = \eta(\mathbf{u}^+ - \mathbf{u}^-)$. We may have the 1-shock

$$\eta < \lambda_1(\mathbf{u}^-) \quad , \quad \lambda_1(\mathbf{u}^+) < \eta < \lambda_2(\mathbf{u}^+)$$

or the 2-shock

$$\lambda_1(\mathbf{u}^-) < \eta < \lambda_2(\mathbf{u}^-) \quad , \quad \lambda_2(\mathbf{u}^+) < \eta .$$

Therefore, the 1-shock is given by

$$\lambda_1(\mathbf{u}^+) < \eta < \lambda_1(\mathbf{u}^-) \tag{13.5.9}$$

and the 2-shock is given by

$$\lambda_2(\mathbf{u}^+) < \eta < \lambda_2(\mathbf{u}^-) . \tag{13.5.10}$$

Since $\eta < \lambda_1(\mathbf{u}^-) < 0$ for the 1-shock (the speed of the shock wave is negative), the 1-shock is also called **back shock** because it moves backward as y increases. Similarly, since $\eta > \lambda_2(\mathbf{u}^+) > 0$ for the 2-shock (the speed of the shock wave is positive), the 2-shock is also called **front shock** because it moves forward as y increases.

We would like to know what are the possible values of $\mathbf{u}^+$ that may be connected to $\mathbf{u}^-$ by a shock wave ¹¹.

i) We first consider the 1-shock. As we did in the theory above, we have to use the Rankine-Hugoniot formula (13.5.3). We get

$$-u_2 + u_2^- = \eta(u_1 - u_1^-) \quad \text{and} \quad p(u_1) - p(u_1^-) = \eta(u_2 - u_2^-) . \tag{13.5.11}$$

It follows from $-\eta(u_1 - u_1^-) = u_2 - u_2^- = (p(u_1) - p(u_1^-))/\eta$ that $\eta^2 = (p(u_1) - p(u_1^-))/(u_1^- - u_1)$. Since $\eta < 0$, we get that

$$\eta = -\left((p(u_1) - p(u_1^-))/(u_1^- - u_1)\right)^{1/2} . \tag{13.5.12}$$

It also follows from (13.5.9) that $-\sqrt{-p'(u_1)} < -\sqrt{-p'(u_1^-)}$. Therefore, $p'(u_1) < p'(u_1^-)$. Since $p''(v) > 0$ for all $v \in \mathbb{R}$, we have that p' is increasing and we must therefore have $u_1 < u_1^-$.

¹⁰The results in this example are in fact true for all functions p that satisfy $p'(v) < 0$ and $p''(v) > 0$ for all $v \in \mathbb{R}$, not only for the specific p of the p -system.

¹¹We could have reversed the statement and ask what are the possible values of $u^-(\mathbf{p})$ that we may connected to $u^+(\mathbf{p})$ by a shock wave. This would not have changed the qualitative conclusions of the example.

Hence, the admissible values of $\mathbf{u}$ are the points on the curve $S_{1,\mathbf{u}^-}$ given by

$$u_2 - u_2^- = (u_1 - u_1^-) \left(\frac{p(u_1) - p(u_1^-)}{u_1^- - u_1} \right)^{1/2} = -((u_1^- - u_1)(p(u_1) - p(u_1^-)))^{1/2}$$

for $u_1 < u_1^-$. We have that

$$\frac{\partial u_2}{\partial u_1} = \frac{p(u_1) - p(u_1^-) - (u_1^- - u_1)p'(u_1)}{2((u_1^- - u_1)(p(u_1) - p(u_1^-)))^{1/2}} > 0$$

for all $u_1 < u_1^-$ because $p(u_1) > p(u_1^-)$ and $p'(u_1) < 0$, and

$$\frac{\partial^2 u_2}{\partial u_1^2} = \frac{(p(u_1) - p(u_1^-) + (u_1^- - u_1)p'(u_1))^2 - 2(u_1^- - u_1)^2(p(u_1) - p(u_1^-))p''(u_1)}{4((u_1^- - u_1)(p(u_1) - p(u_1^-)))^{3/2}} < 0$$

for $u_1 < u_1^-$ and u_1 sufficient close to u_1^- . A possible sketch of $S_{1,\mathbf{u}^-}$ is given in Figure 13.14.

It will be useful later to know that

$$\lim_{u_1 \rightarrow u_1^-} \frac{\partial u_2}{\partial u_1}(u_1) = \lim_{u_1 \rightarrow u_1^-} \frac{-(p(u_1) - p(u_1^-))/(u_1 - u_1^-) - p'(u_1)}{2(-(p(u_1) - p(u_1^-))/(u_1 - u_1^-))^{1/2}} = \frac{-2p'(u_1^-)}{2\sqrt{-p'(u_1^-)}} = \sqrt{-p'(u_1^-)}$$

and

$$\begin{aligned} & \lim_{u_1 \rightarrow u_1^-} \frac{\partial^2 u_2}{\partial u_1^2}(u_1) \\ &= \lim_{u_1 \rightarrow u_1^-} \frac{(u_1^- - u_1)^{-3}(p(u_1) - p(u_1^-) + (u_1^- - u_1)p'(u_1))^2 + 2p''(u_1)(p(u_1) - p(u_1^-))/(u_1 - u_1^-)}{4(-(p(u_1) - p(u_1^-))/(u_1 - u_1^-))^{3/2}} \\ &= \frac{0 + 2p''(u_1^-)p'(u_1^-)}{4(-p'(u_1^-))^{3/2}} = -\frac{p''(u_1^-)}{2\sqrt{-p'(u_1^-)}} < 0, \end{aligned}$$

where we have used the fact that $(p(u_1) - p(u_1^-) - (u_1 - u_1^-)p'(u_1)) = O(|u_1^- - u_1|^2)$ as $|u_1^- - u_1| \rightarrow 0$ to compute the limit.

ii) We now consider the 2-shock. The analysis is almost identical to the analysis for the 1-shock. We again start with the information provided by the Rankine-Hugoniot formula (13.5.3) to get (13.5.11) and from it to get $\eta^2 = (p(u_1) - p(u_1^-))/(u_1^- - u_1)$. However, we now have that $\eta > 0$. Therefore, $\eta = ((p(u_1) - p(u_1^-))/(u_1^- - u_1))^{1/2}$. It also follows from (13.5.10) that $\sqrt{-p'(u_1)} < \sqrt{-p'(u_1^-)}$. Therefore, $p'(u_1) > p'(u_1^-)$. Since $p''(v) > 0$ for all $v \in \mathbb{R}$, we have that p' is increasing and we must therefore have $u_1 > u_1^-$.

Hence, the admissible values of $\mathbf{u}$ are the points on the curve $S_{2,\mathbf{u}^-}$ given by

$$u_2 - u_2^- = -(u_1 - u_1^-) \left(\frac{p(u_1) - p(u_1^-)}{u_1^- - u_1} \right)^{1/2} = -((u_1^- - u_1)(p(u_1) - p(u_1^-)))^{1/2}$$

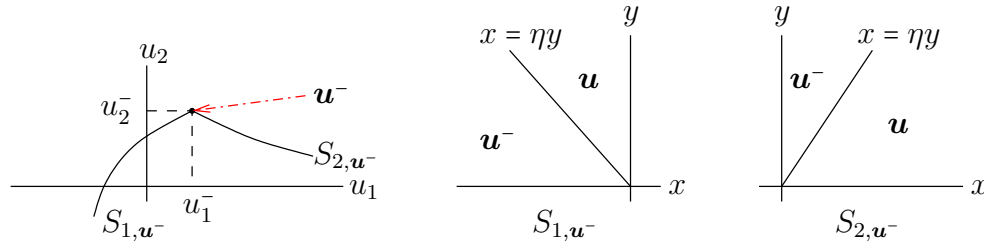


Figure 13.14: The points $\mathbf{u}$ on the curve $S_{1,\mathbf{u}^-}$ are the possible values of $\mathbf{u}^+$ that can be connected by a 1-shock to $\mathbf{u}^-$. The points $\mathbf{u}$ on the curve $S_{2,\mathbf{u}^-}$ are the possible values of $\mathbf{u}^+$ to that can be connected by a 2-shock to $\mathbf{u}^-$.

for $u_1 > u_1^-$. We have that

$$\frac{\partial u_2}{\partial u_1} = \frac{p(u_1) - p(u_1^-) - (u_1^- - u_1)p'(u_1)}{2((u_1^- - u_1)(p(u_1) - p(u_1^-)))^{1/2}} < 0$$

for all $u_1 > u_1^-$ because $p(u_1) < p(u_1^-)$ and $p'(u_1) < 0$, and

$$\frac{\partial^2 u_2}{\partial u_1^2} = \frac{(p(u_1) - p(u_1^-) + (u_1^- - u_1)p'(u_1))^2 - 2(u_1^- - u_1)^2(p(u_1) - p(u_1^-))p''(u_1)}{4((u_1^- - u_1)(p(u_1) - p(u_1^-)))^{3/2}} < 0$$

for $u_1 > u_1^-$ and u_1 sufficient close to u_1^- . A possible sketch of $S_{2,\mathbf{u}^-}$ is given in Figure 13.14.

As for (i), it will be useful later to know that

$$\lim_{u_1 \rightarrow u_1^-} \frac{\partial u_2}{\partial u_1}(u_1) = \lim_{u_1 \rightarrow u_1^-} \frac{(p(u_1) - p(u_1^-))/(u_1 - u_1^-) + p'(u_1)}{2(-(p(u_1) - p(u_1^-))/(u_1 - u_1^-))^{1/2}} = \frac{2p'(u_1^-)}{2\sqrt{-p'(u_1^-)}} = -\sqrt{-p'(u_1^-)}$$

and

$$\begin{aligned} & \lim_{u_1 \rightarrow u_1^-} \frac{\partial^2 u_2}{\partial u_1^2}(u_1) \\ &= \lim_{u_1 \rightarrow u_1^-} \frac{(u_1 - u_1^-)^{-3}(p(u_1) - p(u_1^-) + (u_1^- - u_1)p'(u_1))^2 - 2p''(u_1)(p(u_1) - p(u_1^-))/(u_1 - u_1^-)}{4(-(p(u_1) - p(u_1^-))/(u_1 - u_1^-))^{3/2}} \\ &= \frac{0 - 2p''(u_1^-)p'(u_1^-)}{4(-p'(u_1^-))^{3/2}} = \frac{p''(u_1^-)}{2\sqrt{-p'(u_1^-)}} > 0. \end{aligned}$$

iii) We would like to find out the values of $\mathbf{u}^+$ that may be connected to a given $\mathbf{u}^-$ by a rarefaction wave. We proceed as in Section 13.4.1 to address this problem.

Since $u_\mu(x, y) \equiv u(\mu x, \mu y)$ is a solution of (13.5.2) with $G(u) = \begin{pmatrix} -u_2 \\ p(u_1) \end{pmatrix}$ for all $\mu > 0$, we seek a solution of the form $u(x, y) = w(x/y)$, where $w : \mathbb{R} \rightarrow \mathbb{R}^2$ is of class C^1 . If we set

$z = x/y$ and substitute $u(x, y) = w(z)$ in (13.5.2), we have seen previously that this implies that

$$(D G)(w(z)) \frac{dw}{dz}(z) = z \frac{dw}{dz}(z) .$$

If $\frac{dw}{dz}(z) \neq \mathbf{0}$, then it is an eigenvector of $(D G)(w(z))$ associated to the eigenvalue z .

We have already seen that $D G(\mathbf{u}) = \begin{pmatrix} 0 & -1 \\ p'(u_1) & 0 \end{pmatrix}$ has the eigenvalues $\lambda_1(\mathbf{u}) = -\sqrt{-p'(u_1)}$ and $\lambda_2(\mathbf{u}) = \sqrt{-p'(u_1)}$.

We first consider the eigenvalue $\lambda_1(\mathbf{u})$. We get from

$$\begin{pmatrix} 0 & -1 \\ p'(w_1) & 0 \end{pmatrix} \begin{pmatrix} \frac{dw_1}{dz} \\ \frac{dw_2}{dz} \end{pmatrix} = -\sqrt{-p'(w_1)} \begin{pmatrix} \frac{dw_1}{dz} \\ \frac{dw_2}{dz} \end{pmatrix}$$

that $\frac{dw_2}{dz} - \sqrt{-p'(w_1)} \frac{dw_1}{dz} = 0$. Since $\frac{dw_1}{dz} \neq 0$ because we assume that $\frac{dw}{dz} \neq \mathbf{0}$ for an eigenvector, we get

$$\frac{dw_2}{dw_1} = \frac{dw_2}{dz} \left(\frac{dw_1}{dz} \right)^{-1} = \sqrt{-p'(w_1)} .$$

If we integrate from u_1^- to u_1 , we get

$$u_2 - u_2^- = \int_{u_1^-}^{u_1} \sqrt{-p'(v)} dv \quad (13.5.13)$$

Since the speed of the wave should increase along the x axis, we must have that $\lambda_1(\mathbf{u}) > \lambda_1(\mathbf{u}^-)$; namely, $-\sqrt{-p'(u_1)} > -\sqrt{-p'(u_1^-)}$. Thus, $p'(u_1) > p'(u_1^-)$. Since p' is increasing, we get that $u_1 > u_1^-$. Thus, the acceptable values of $\mathbf{u}$ are the points on the curve $R_{1, \mathbf{u}^-}$ defined by (13.5.13) for $u_1 > u_1^-$. We have that $\frac{du_2}{du_1} = \sqrt{-p'(u_1)}$ and $\frac{d^2 u_2}{du_1^2} = \frac{-p''(u_1)}{2\sqrt{-p'(u_1)}}$.

A possible sketch of the curve $R_{1, \mathbf{u}^-}$ is represented in Figure 13.16.

The rarefaction waves connecting $\mathbf{u}^-$ to $\mathbf{u}$ on the curve $R_{1, \mathbf{u}^-}$ are called **back rarefaction waves** (Figure 13.15).

We repeat the previous analysis but this time for the eigenvalue $\lambda_2(\mathbf{u})$. We get from

$$\begin{pmatrix} 0 & -1 \\ p'(w_1) & 0 \end{pmatrix} \begin{pmatrix} \frac{dw_1}{dz} \\ \frac{dw_2}{dz} \end{pmatrix} = \sqrt{-p'(w_1)} \begin{pmatrix} \frac{dw_1}{dz} \\ \frac{dw_2}{dz} \end{pmatrix}$$

that $\frac{dw_2}{dz} + \sqrt{-p'(w_1)} \frac{dw_1}{dz} = 0$. Since $\frac{dw_1}{dz} \neq 0$ because we assume that $\frac{dw}{dz} \neq \mathbf{0}$ for an eigenvector, we get

$$\frac{dw_2}{dw_1} = \frac{dw_2}{dz} \left(\frac{dw_1}{dz} \right)^{-1} = -\sqrt{-p'(w_1)} .$$

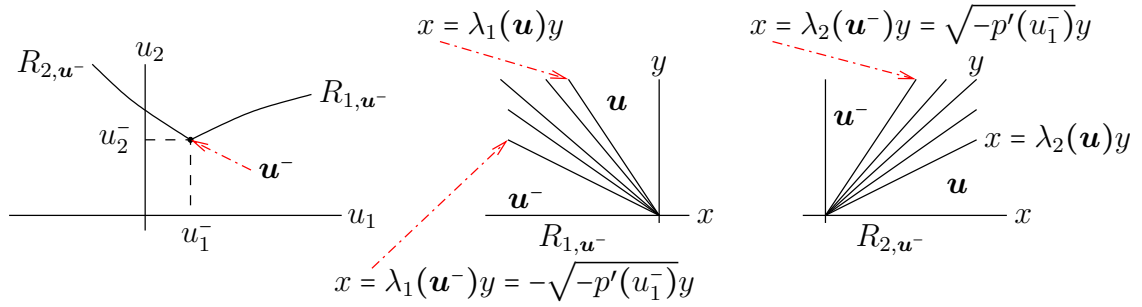


Figure 13.15: The points $\mathbf{u}$ on the curve $R_{1,\mathbf{u}^-}$ are the possible values of $\mathbf{u}^+$ that can be connected by a back rarefaction wave to $\mathbf{u}^-$. The points $\mathbf{u}$ on the curve $R_{2,\mathbf{u}^-}$ are the possible values of $\mathbf{u}^+$ that can be connected by a front rarefaction wave to $\mathbf{u}^-$.

If we integrate from u_1^- to u_1 , we get

$$u_2 - u_2^- = - \int_{u_1^-}^{u_1} \sqrt{-p'(v)} dv \quad (13.5.14)$$

Since the speed of the wave should increase along the x axis, we must have that $\lambda_2(\mathbf{u}) > \lambda_2(\mathbf{u}^-)$; namely, $\sqrt{-p'(u_1)} > \sqrt{-p'(u_1^-)}$. Thus, $p'(u_1) < p'(u_1^-)$. Since p' is increasing, we get that $u_1 < u_1^-$. Thus, the acceptable values of $\mathbf{u}$ are the points on the curve R_2 defined by (13.5.13) for $u_1 < u_1^-$. We have that $\frac{du_2}{du_1} = -\sqrt{-p'(u_1)}$ and $\frac{d^2u_2}{du_1^2} = \frac{p''(u_1)}{2\sqrt{-p'(u_1)}}$.

A possible sketch of the curve $R_{2,\mathbf{u}^-}$ is drawn in Figure 13.15. The rarefaction waves connecting $\mathbf{u}^-$ to $\mathbf{u}$ on the curve $R_{2,\mathbf{u}^-}$ are called **front rarefaction waves**.

The curves $S_{1,\mathbf{u}^-}$ and $R_{1,\mathbf{u}^-}$ intersect at $\mathbf{u}^-$ where they have matching first and second order derivatives as it is confirmed by our computations at the end of (i) (Figure 13.16). Likewise, the curves $S_{2,\mathbf{u}^-}$ and $R_{2,\mathbf{u}^-}$ intersect at $\mathbf{u}^-$ where they have matching first and second order derivatives as it is confirmed by our computations at the end of (ii).

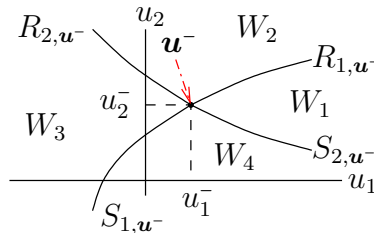


Figure 13.16: We have combined the graphs of $S_{1,\mathbf{u}^-}$, $S_{2,\mathbf{u}^-}$, $R_{1,\mathbf{u}^-}$ and $R_{2,\mathbf{u}^-}$.

iv) We divide the x, y plane into four regions (at least in an open neighbourhood of $\mathbf{u}^-$). W_1 is the region bounded by $S_{2,\mathbf{u}^-}$ and $R_{1,\mathbf{u}^-}$, W_2 is the region bounded by $R_{1,\mathbf{u}^-}$ and $R_{2,\mathbf{u}^-}$,

W_3 is the region bounded by R_{2,u^-} and S_{1,u^-} , and W_4 is the region bounded by S_{1,u^-} and S_{2,u^-} as in Figure 13.16.

We know from (i), (ii) and (iii) how to use a shock wave or a rarefaction wave to connect u^- to u if u is on one of S_{1,u^-} , S_{2,u^-} , R_{1,u^-} or R_{2,u^-} . We now explain how to do it when u is not on one of these four curves.

For instance, suppose that $u \in W_2$ as in Figure 13.17. Suppose that $u \in R_{2,u_{bf}}$ with $u_{bf} \in R_{1,u^-}$. We can connect u_{bf} to u with a front rarefaction wave and u^- to u_{bf} with a back rarefaction wave. There may be another way to connect u^- to u . Suppose that $u \in R_{1,u_{fb}}$ with $u_{fb} \in R_{2,u^-}$. We can connect u_{fb} to u with a back rarefaction wave and u^- to u_{fb} with a front rarefaction wave.

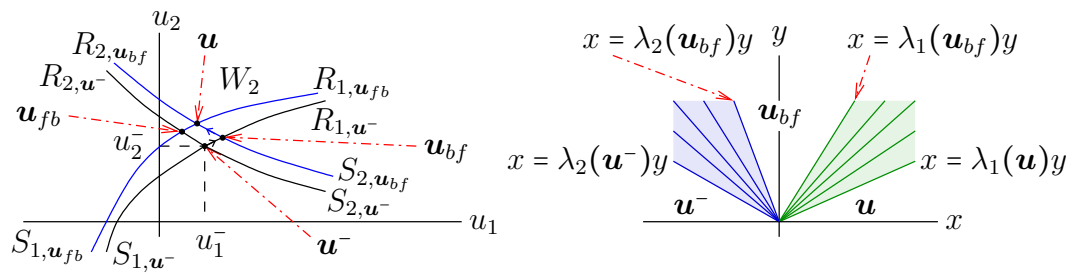


Figure 13.17: The figure used to explain how to connect u^- to $u \in W_2$ using rarefaction and shock waves.

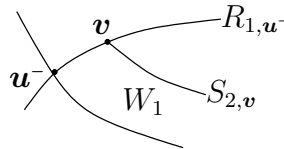


Figure 13.18: The region W_1 is covered by the curves $S_{2,v}$ for $v \in R_{1,u^-}$.

v) The procedure formally presented in (iv) to connect u to u^- with shock and rarefaction waves is valid in a small neighbourhood of the origin but may fail globally. It is proved in [35] that the procedure presented in (iv) is valid without restriction in the regions W_1 , W_3 and W_4 . Namely, $\bigcup_{v \in R_{1,u^-}} S_{2,v}$ will cover W_1 without any intersection between the $S_{2,v}$ curves (Figure 13.18). We have a similar result for W_3 and W_4 . However, it is not valid in the entire region W_2 . To understand why it fails, we use Figure 13.19. Suppose $P = \int_{u_1^-}^{\infty} \sqrt{-p'(v)} dv < \infty$. The reader should verify that this last assumption is true in the important case where $p(v) = kv^{-\beta}$ and $\beta > 1$.

Given $\mathbf{v} \in R_{1,u^-}$, we have from (13.5.13) that $v_2 - u_2^- = \int_{u_1^-}^{v_1} \sqrt{-p'(v)} dv$. So, $\lim_{v_1 \rightarrow \infty} v_2 = u_2^- + P$. In particular, $v_2 \leq u_2^- + P$ for all $\mathbf{v} \in R_{1,u^-}$.

Given $\mathbf{w} \in R_{2,v}$ with $w_1 = u_1^-$, we have from (13.5.14) that $w_2 - v_2 = - \int_{v_1}^{w_1} \sqrt{-p'(v)} dv$. So, $w_2 \leq v_2 + P \leq u_2^- + 2P$ for all $\mathbf{v} \in R_{1,u^-}$. Therefore, there is no curve $R_{2,v}$ that contains $\mathbf{u}$ if $u_1 = u_1^-$ and $u_2 > u_2^- + 2P$.

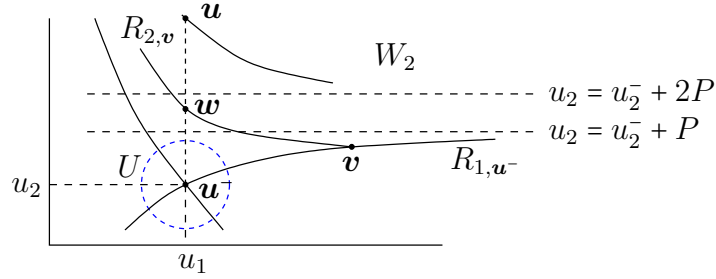


Figure 13.19: Illustration of the problem to connect $\mathbf{u} \in W_2$ to $\mathbf{u}^-$ using a back rarefaction wave first.

If U is a small enough open neighbourhood of $\mathbf{u}^-$ as in Figure 13.19, then it can be shown that $\mathbf{u} \in R_{2,v}$ for some $\mathbf{v} \in R_{1,u^-}$ for $\mathbf{u} \in U$. ♣

13.5.1 Basic Concepts

The previous example illustrate what may be expected in systems of conservation laws in higher dimensions; namely, when $n > 2$. However, more conditions have to be imposed on the systems of conservation laws to be able to get valuable results.

We should start by generalizing the concept of x, y characteristics for the system of conservation laws (13.5.2). A x, y characteristic is a curve $\Sigma = \{(\rho(y), y) : y \geq 0\}$ such that

$$\frac{d}{dy} u(\rho(y), y) = \mathbf{0}$$

for a solution u of (13.5.2). If Σ is a x, y characteristic, we therefore have that

$$\begin{aligned} \mathbf{0} &= \frac{\partial u(\rho(y), y)}{\partial y} + \rho'(y) \frac{\partial u(\rho(y), y)}{\partial x} = -(\mathbf{D}G)(u(\rho(y), y)) \frac{\partial u(\rho(y), y)}{\partial x} + \rho'(y) \frac{\partial u(\rho(y), y)}{\partial x} \\ &= -((\mathbf{D}G)(u(\rho(y), y)) - \rho'(y) \text{Id}) \frac{\partial u(\rho(y), y)}{\partial x}. \end{aligned}$$

Thus, if we assume that (13.5.2) is a strictly hyperbolic differential equation, then $\rho'(y)$ must be an eigenvalue of $(\mathbf{D}G)(u(\rho(y), y))$ and $\frac{\partial u(\rho(y), y)}{\partial x}$ an eigenvector of $(\mathbf{D}G)((\rho(y), y))$ associated to $\rho'(y)$ for all y .

Since they will play an important role in studying strictly hyperbolic systems of conservation laws, we assume from now on that $\lambda_1(\mathbf{u}) < \lambda_2(\mathbf{u}) < \dots < \lambda_n(\mathbf{u})$ are the eigenvalues of $DG(\mathbf{u})$ for $\mathbf{u} \in \mathbb{R}^n$, and $q_i(\mathbf{u}) \in \mathbb{R}^n$ is a (right) eigenvector of $DG(\mathbf{u})$ associated to the eigenvalue $\lambda_i(\mathbf{u})$ for $1 \leq i \leq n$.

Hence, the x, y characteristics associated to the system of conservation laws (13.5.2) are given by $\rho'(y) = \lambda_i(u(\rho(y), y))$ for $1 \leq i \leq n$.

It is also going to be useful to assume that $p_i(\mathbf{u}) \in \mathbb{R}^n$ is a left eigenvector of $DG(\mathbf{u})$ associated to the eigenvalue $\lambda_i(\mathbf{u})$ for $1 \leq i \leq n$; namely, $(p_i(\mathbf{u}))^\top DG(\mathbf{u}) = \lambda_i(\mathbf{u})(p_i(\mathbf{u}))^\top$. It is well known that $(p_i(\mathbf{u}))^\top q_j(\mathbf{u}) = 0$ for $i \neq j$ and $(p_i(\mathbf{u}))^\top q_j(\mathbf{u}) \neq 0$ for $i = j$. So, after scaling, we may assume that $(p_i(\mathbf{u}))^\top q_i(\mathbf{u}) = 1$ for $1 \leq i \leq n$.

Definition 13.5.4

The pair (λ_i, q_i) is called the i^{th} **characteristic field**.

The first property which is useful to analyze the Riemann problem associated to (13.5.2) is given in the next definition. It is particularly useful for the study of rarefaction waves connecting two states $\mathbf{u}^-$ and $\mathbf{u}^+$.

Definition 13.5.5

We say that the i^{th} characteristic field of the system of conservation laws (13.5.2) is **genuinely nonlinear** in U if $\nabla \lambda_i(\mathbf{u}) \cdot q_i(\mathbf{u}) \neq 0$ for all $\mathbf{u} \in U$ ¹². If $\nabla \lambda_i(\mathbf{u}) \cdot q_i(\mathbf{u}) \neq 0$ for all $\mathbf{u} \in U$ and $1 \leq i \leq n$, we say that the system of conservation laws (13.5.2) is **genuinely nonlinear**.

Example 13.5.6 (Example 13.5.3 continued)

For the p -system, we have found that the eigenvalues of the matrix $DG(\mathbf{u}) = \begin{pmatrix} 0 & -1 \\ p'(u_1) & 0 \end{pmatrix}$ are $\lambda_1(\mathbf{u}) = -\sqrt{-p'(u_1)}$ and $\lambda_2(\mathbf{u}) = \sqrt{-p'(u_1)}$. An eigenvector associated to $\lambda_1(\mathbf{u})$ is $q_1(\mathbf{u}) = \begin{pmatrix} 1 \\ \sqrt{-p'(u_1)} \end{pmatrix}$ and an eigenvector associated to $\lambda_2(\mathbf{u})$ is $q_2(\mathbf{u}) = \begin{pmatrix} -1 \\ \sqrt{-p'(u_1)} \end{pmatrix}$.

We have that

$$\nabla \lambda_1(\mathbf{u}) \cdot q_1(\mathbf{u}) = \left(\frac{p''(u_1)}{2\sqrt{-p'(u_1)}}, 0 \right) \cdot \left(1, \sqrt{-p'(u_1)} \right) = \frac{p''(u_1)}{2\sqrt{-p'(u_1)}}$$

and

$$\nabla \lambda_2(\mathbf{u}) \cdot q_2(\mathbf{u}) = \left(\frac{-p''(u_1)}{2\sqrt{-p'(u_1)}}, 0 \right) \cdot \left(-1, \sqrt{-p'(u_1)} \right) = \frac{p''(u_1)}{2\sqrt{-p'(u_1)}}.$$

Thus, by assuming that $p''(v) > 0$ for all v , we have that the p -system is genuinely nonlinear. ♣

¹²When using the dot product, we assume that $q_i(\mathbf{u})$ is treated as a row vector, not a column vector. This is the perennial dilemma between the geometric and algebraic interpretation of vectors.

Example 13.5.7

To appreciate the importance of this definition, we consider the system of conservation laws (13.5.2) with $u(x, y) = \mathbf{u}^-$ for $x < y\lambda_i(\mathbf{u}^-)$. We seek $\mathbf{u}^+$ such that $\mathbf{u}^-$ could be connected to $\mathbf{u}^+$ by a rarefaction wave.

Let us assume that $(\nabla \lambda_i)(\mathbf{u}) \cdot q_i(\mathbf{u}) \neq 0$ in a neighbourhood U of $\mathbf{u}^-$. By scaling the eigenvector q_i , we may assume that $(\nabla \lambda_i)(\mathbf{u}) \cdot q_i(\mathbf{u}) = 1$ for $\mathbf{u} \in U$.

Let w be the solution of

$$\frac{dw}{dz} = q_i(w)$$

with the initial condition $w(z^-) = \mathbf{u}^-$ where $z^- = \lambda_i(\mathbf{u}^-)$. The unique solution w is well defined and continuously differentiable (assuming that G is continuously differentiable) on an interval $[z^-, z^+]$.

Since

$$\frac{d}{dz} \lambda_i(w(z)) = (D \lambda_i)(w(z)) \frac{dw}{dz}(z) = (D \lambda_i)(w(z)) q_i(w(z)) = (\nabla \lambda_i)(w(z)) \cdot q_i(w(z)) = 1$$

for all $z \in [z^-, z^+]$. we have that $\lambda_i(w(z)) = z + C$ for all $z \in [z^-, z^+]$. Moreover, since $w(\lambda_i(\mathbf{u}^-)) = \mathbf{u}^-$, we have that $C = 0$.

Let $\mathbf{u}^+ = w(z^+)$. The **ith-rarefaction wave** connecting $\mathbf{u}^-$ to $\mathbf{u}^+$ is defined by

$$u(x, y) = \begin{cases} \mathbf{u}^- & \text{if } x \leq yz^- \\ w(x/y) & \text{if } yz^- < x < yz^+ \\ \mathbf{u}^+ & \text{if } x \geq yz^+ \end{cases}$$

See Figure 13.20.

Since $z = \lambda_i(w(z))$ for all $z^- \leq z \leq z^+$, we have that $\lambda_i \circ w$ is increasing along the wave; in particular, $\lambda_i(w(z^+)) = z^+ > z^- = \lambda_i(w(z^-))$.

We have that $u(x, y) = w(x/y)$ is a classical solution of (13.5.2) in the wedge W defined by $yz^- < x < yz^+$. Let $z = x/y$ for $(x, y) \in W$. Then

$$\begin{aligned} \frac{\partial u}{\partial y}(x, y) + (D G)(u(x, y)) \frac{\partial u}{\partial x}(x, y) &= \frac{\partial w}{\partial z}(z) \frac{\partial z}{\partial y} + (D G(w(z))) \frac{\partial w}{\partial z}(z) \frac{\partial z}{\partial x} \\ &= -\frac{x}{y^2} q_i(w(z)) + \frac{1}{y} (D G)(w(z)) q_i(w(z)) = -\frac{x}{y^2} q_i(w(z)) + \frac{1}{y} \lambda_i(w(z)) q_i(w(z)) \\ &= -\frac{x}{y^2} q_i(w(z)) + \frac{x}{y^2} q_i(w(z)) = 0. \end{aligned}$$

♣

Definition 13.5.8

Let $w : I \rightarrow \mathbb{R}^n$ be a solution of $\frac{dw}{dz} = q_i(w)$, where I is an open interval containing the origin. The orbit $\{w(z) : z \in I\}$ is called an **integral curve** associated to q_i .

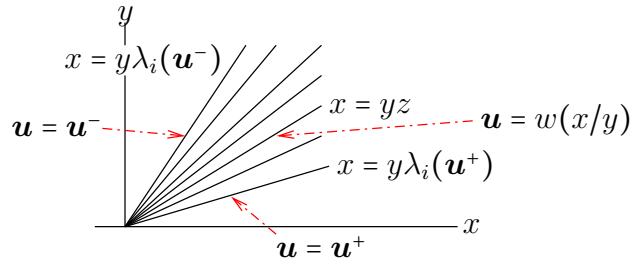


Figure 13.20: Example of an i^{th} -rarefaction wave for a system of conservation laws in higher dimension. We assume that $u(x, y) = \mathbf{u}^-$ for $x < y\lambda_i(\mathbf{u}^-)$. The lines where u is constant are given by $x = yz$ for $\lambda_i(\mathbf{u}^-) = z^- \leq z \leq z^+ = \lambda_i(\mathbf{u}^+)$.

Definition 13.5.9

A **i -Riemann invariant** on an open set $U \subset \mathbb{R}^n$ is continuously differentiable function $v : U \rightarrow \mathbb{R}$ such that $\nabla v(\mathbf{u}) \cdot \mathbf{q}_i(\mathbf{u}) = 0$ for all $\mathbf{u} \in U$.

It can be shown that we can always select $n - 1$ i -Riemann invariants $v_j : U \rightarrow \mathbb{R}$ for $1 \leq j \leq n - 1$ such that $\nabla v_j(\mathbf{u})$ for $1 \leq j \leq n - 1$ are linearly independent for all $\mathbf{u} \in U$. At least, if U is small enough.

Example 13.5.10 (Example 13.5.6 continued)

To find the 1-Riemann invariants v of the p -system, we have to solve

$$\nabla v(\mathbf{u}) \cdot \mathbf{q}_1(\mathbf{u}) = \left(\frac{\partial v}{\partial u_1}(\mathbf{u}), \frac{\partial v}{\partial u_2}(\mathbf{u}) \right) \cdot \left(1, \sqrt{-p'(u_1)} \right) = \frac{\partial v}{\partial u_1}(\mathbf{u}) + \sqrt{-p'(u_1)} \frac{\partial v}{\partial u_2}(\mathbf{u}) = 0.$$

It is easy to verify that the solutions are given by $v(\mathbf{u}) = u_2 + F(u_1)$, where $F(v)$ is a primitive of $-\sqrt{-p'(v)}$. Similarly, we find that the 2-Riemann invariants v of the p -system are given by $v(\mathbf{u}) = u_2 - F(u_1)$. $\clubsuit$

The i -Riemann invariants are constant along the integral curves associated to $\mathbf{q}_i$ because

$$\begin{aligned} \frac{d}{dz} v(w(z)) &= (Dv)(w(z)) \frac{dw}{dz} = (Dv)(w(z)) \mathbf{q}_i(w(z)) \\ &= (\nabla v)(w(z)) \cdot \mathbf{q}_i(w(z)) = 0 \end{aligned} \tag{13.5.15}$$

for the solution w of $\frac{dw}{dz} = \mathbf{q}_i(w)$ and any i -Riemann invariants $v : U \rightarrow \mathbb{R}$. In the previous argument, we obviously assume that we consider the values of z for which $w(z) \in U$ to be able to compose w with v .

Example 13.5.11 (Example 13.5.10 continued)

Let us prove for the p -system that the i -Riemann invariants are constant along the integral curves associated to $\mathbf{q}_i$.

The integral curve associated to q_1 is given by

$$\frac{dw}{dz}(z) = q_1(w(z)) = \left(\frac{1}{\sqrt{-p'(w_1(z))}} \right)$$

Thus $w(z) = \left(\frac{z+C}{\tilde{F}(z)} \right)$, where C is a constant and $\tilde{F}$ is a primitive of $\sqrt{-p'(z+C)}$.

Let $v(\mathbf{u}) = u_2 + F(u_1)$ be the 1-Riemann invariants defined in Example 13.5.10. We have to show that $v(w(z)) = \tilde{F}(z) + F(z+C)$ is constant. This is clear because

$$\frac{d}{dz}v(w(z)) = \sqrt{-p'(z+C)} - \sqrt{-p'(z+C)} = 0$$

for all z . A similar computation shows that the 2-Riemann invariants are constant along the integral curves associated to q_2 . ♣

Remark 13.5.12

Suppose that (13.5.2) is a strictly hyperbolic system of conservation laws in $\mathbb{R}^2$. Moreover, suppose that $U \subset \mathbb{R}^2$ is an open and connected set.

Let $v_i : U \rightarrow \mathbb{R}$ denotes the i -Riemann invariants for $i = 1, 2$. We have that v_i is constant along integral curves associated to right eigenvector q_j if $i \neq j$. Since the right eigenvectors $q_1(\mathbf{u})$ and $q_2(\mathbf{u})$ are linearly independent for all $\mathbf{u} \in \mathbb{R}^2$, the integral curves associated to q_1 intersect transversely the integral curves associated to q_2 . Hence, using standard results in ordinary differential equation, we can show that

$$\begin{aligned} \Phi : U &\rightarrow \Phi(U) \\ \mathbf{u} &\mapsto (v_1(\mathbf{u}), v_2(\mathbf{u})) \end{aligned}$$

defines an isomorphism and $\Phi(U) \subset \mathbb{R}^2$ is open and connected.

Since the right and left eigenvectors satisfy $p_i(\mathbf{u}) \cdot q_j(\mathbf{u}) = \delta_{i,j}$ for all $\mathbf{u}$, we have that $Dv_1(\mathbf{u})$ is a left eigenvector associated to $\lambda_2(\mathbf{u})$ and $Dv_2(\mathbf{u})$ is a left eigenvector associated to $\lambda_1(\mathbf{u})$. Therefore

$$\begin{aligned} 0 &= (Dv_1(u)) \left(\frac{\partial u}{\partial y} + (DG(u)) \frac{\partial u}{\partial x} \right) = (Dv_1(u)) \frac{\partial u}{\partial y} + (Dv_1(u))(DG(u)) \frac{\partial u}{\partial x} \\ &= (Dv_1(u)) \frac{\partial u}{\partial y} + \lambda_2(u)(Dv_1(u)) \frac{\partial u}{\partial x} = \frac{\partial v_1(u)}{\partial y} + \lambda_2(u) \frac{\partial v_1(u)}{\partial x}. \end{aligned}$$

Using the v_1, v_2 coordinate system, it is the equation $\frac{\partial v_1}{\partial y} + \tilde{\lambda}_2(v_1, v_2) \frac{\partial v_1}{\partial x} = 0$, where $\tilde{\lambda}_2(v_1, v_2) = \lambda_2(u_1, u_2)$. Similarly, we find that $\frac{\partial v_2}{\partial y} + \tilde{\lambda}_1(v_1, v_2) \frac{\partial v_2}{\partial x} = 0$. This seems to be an interesting approach so solve our original system of conservation laws. Unfortunately, this system is not completely decoupled and is no longer in the form of a system of conservation laws; namely, cannot be written in the form (13.5.2). This is another example where a transformation of a system of conservation laws may lead to undesirable consequences. ♠

Example 13.5.13 (Example 13.5.7 continued)

The fact that i -Riemann invariants are constant along the integral curves associated to q_i implies that these i -Riemann invariants are constant on $u(W)$, where u is the solution defined in Example 13.5.7 and $W \subset \mathbb{R} \times [0, \infty[$ is the wedge where the rarefaction wave is defined. In fact, we have from (13.5.15) that the i -Riemann invariants are constant on $\{w(z) : z \in u(W) = [z^-, z^+]\}$ where w , used to define u in Example 13.5.7, yields an integral curve associated to q_i . This property will become the base for the definition of i^{th} -rarefaction wave. ♣

We generalize our concept of rarefaction wave.

Definition 13.5.14

Suppose that u is a classical solution of (13.5.2) with $u(x, y) \in U$, where U is an open subset of $\mathbb{R}^n$. We will say that u is a **i^{th} -rarefaction wave** in U if all i -Riemann invariant on U are constant.

Proposition 13.5.15

Suppose that u is an i^{th} -rarefaction wave on U . Then the x, y characteristic curves given by $\rho'(y) = \lambda_i(u(\rho(y), y))$ are straight lines and u is constant on each line.

Proof.

Let $w_j : U \rightarrow \mathbb{R}$ for $1 \leq j \leq n-1$ be $n-1$ linearly independent i -Riemann invariants associated to q_i .

Let I be a subset of the domain of ρ such that $u(\rho(y), y) \in U$ for $y \in I$. We need to prove that $\frac{\partial}{\partial y} u(\rho(y), y) = 0$ for all $y \in I$ to prove that u is constant on the x, y characteristic curves. At the same time, this proves that $\lambda_i(u(\rho(y), y))$ is constant. So, the characteristic curves are given by $x = \rho(y) = zy + C$ with $z = \lambda_i(u(\rho(y_0), y_0))$ for some $y_0 \in I$ arbitrary but fixed and C a constant¹³.

Since $\frac{\partial u}{\partial y} + DG(u) \frac{\partial u}{\partial x} = \mathbf{0}$ for $u \in U$, we have that

$$\frac{\partial u}{\partial y}(\rho(y), y) + (DG)(u(\rho(y), y)) \frac{\partial u}{\partial x}(\rho(y), y) = \mathbf{0}$$

for $y \in I$. Hence,

$$\begin{aligned} (p_i(u(\rho(y), y)))^\top \frac{\partial}{\partial y} u(\rho(y), y) &= (p_i(u(\rho(y), y)))^\top \left(\frac{\partial u}{\partial y}(\rho(y), y) + \lambda_i(u(\rho(y), y)) \frac{\partial u}{\partial x}(\rho(y), y) \right) \\ &= (p_i(u(\rho(y), y)))^\top \frac{\partial u}{\partial y}(\rho(y), y) + \lambda_i(u(\rho(y), y)) (p_i(u(\rho(y), y)))^\top \frac{\partial u}{\partial x}(\rho(y), y) \\ &= (p_i(u(\rho(y), y)))^\top \frac{\partial u}{\partial y}(\rho(y), y) + (p_i(u(\rho(y), y)))^\top (DG)(u(\rho(y), y)) \frac{\partial u}{\partial x}(\rho(y), y) \end{aligned}$$

¹³For the Riemann problems that we have considered so far, $C = 0$ because the rarefaction waves were based at $(x, y) = (0, 0)$. This does not need to be.

$$= (p_i(u(\rho(y), y)))^\top \left(\frac{\partial u}{\partial y}(\rho(y), y) + (D G)(u(\rho(y), y)) \frac{\partial u}{\partial x}(\rho(y), y) \right) = 0$$

for $y \in I$.

Moreover, since w_j is constant on U , we have that

$$(D w_j)(u(\rho(y), y)) \frac{\partial}{\partial y} u(\rho(y), y) = \frac{\partial}{\partial y} w_j(u(\rho(y), y)) = 0$$

for $1 \leq j \leq n-1$ and $y \in I$.

Therefore, $\frac{\partial}{\partial y} u(\rho(y), y)$ is orthogonal to $p_i(u(\rho(y), y))$ and $(D w_j)^\top(u(\rho(y), y))$ with $1 \leq j \leq n-1$ for all $y \in I$. However, the vectors $(D w_j)^\top(u(\rho(y), y))$ for $1 \leq j \leq n-1$ are linearly independent and orthogonal to $q_i(u(\rho(y), y))$. Since $(q_i(u(\rho(y), y)))^\top p_i(u(\rho(y), y)) \neq 0$, the vector $p_i(u(\rho(y), y))$ is not orthogonal to $q_i(u(\rho(y), y))$ and so is linearly independent of $(D w_j)^\top(u(\rho(y), y))$ for $1 \leq j \leq n-1$.

Thus, $\frac{\partial}{\partial y} u(\rho(y), y)$ is orthogonal to a basis of $\mathbb{R}^n$ for all $y \in I$. This is possible only if $\frac{\partial}{\partial y} u(\rho(y), y) = \mathbf{0}$ for all $y \in I$. ■

The result of the previous proposition confirm what we have observed in Example 13.5.7.

The analysis that we have performed for the p -system in Example 13.5.3 can be repeated for systems of conservation laws in higher dimensions. There are two main results supporting this analysis.

Theorem 13.5.16

Suppose that (13.5.2) is a strictly hyperbolic system of conservation laws which is genuinely linear on an open set $U \subset \mathbb{R}^n$. For each $1 \leq i \leq n$ and $\mathbf{u}^- \in U$, if we select the eigenvectors $q_i(\mathbf{u})$ such that $\nabla \lambda_i(\mathbf{u}) \cdot q_i(\mathbf{u}) = 1$ for all $\mathbf{u} \in U$, then there exist an interval $[0, r_i]$ and a smooth function $R_i : [0, r_i] \rightarrow \mathbb{R}^n$ such that $R_i(0) = \mathbf{u}^-$ and $\mathbf{u}^-$ can be connected to $R_i(r)$ by an i -rarefaction wave for all $r \in [0, r_i]$.

The main idea for the proof of this theorem was introduced in Example 13.5.7; namely, one considers the ordinary differential equation $\frac{dw}{dz} = q_i(w)$ with $w(\lambda_i(\mathbf{u}^-)) = \mathbf{u}^-$. The proof can be found in Section B of Chapter 17 of [35].

The choice of R_i for the name of the function is not random. We obviously want to indicate that the curves $R_i([0, r_i]) \subset \mathbb{R}^n$ are the generalizations of the curves $R_{1, \mathbf{u}^-}$ and $R_{2, \mathbf{u}^-}$ in Example 13.5.3 about the p -system.

The generalization to higher dimensions of rarefaction waves can also be done for the shock waves. The main results are summarized in the next proposition

Theorem 13.5.17

Suppose that (13.5.2) is a genuinely linear and strictly hyperbolic system of conservation laws. For each $1 \leq i \leq n$ and $\mathbf{u}^-$, there exist an interval $[s_i, 0]$ and a smooth function $S_i : [s_i, 0] \rightarrow \mathbb{R}^n$ such that $S_i(0) = \mathbf{u}^-$, $\frac{dS_i}{ds}(0) = q_i(\mathbf{u}^-)$, $\frac{d^2 S_i}{ds^2}(0) = \frac{d}{ds} q_i(S_i(s)) \Big|_{s=0}$ and $\mathbf{u}^-$ can be connected to $S_i(s)$ by a i -shock for all $s \in [s_i, 0]$. Moreover, we have that $G(S_i(s)) - G(\mathbf{u}^-) = \eta(s)(S_i(s) - \mathbf{u}^-)$ with $\eta(0) = \lambda_i(\mathbf{u}^-)$ and $\eta'(0) = 1/2$.

The proof of the previous theorem can be found in Section B of Chapter 17 in [35]. The curve $\Gamma \subset \mathbb{R} \times [0, \infty[$ where the strong solution u is discontinuous is obtained simultaneously as we obtain the curves S_i . For y fixed, the value of $\gamma'(y)$ for the parametric representation of the curve Γ that we have used before is given by $\eta(s)$. So, $\gamma'(y)$ may change as s varies. In Example 13.5.3, η is given in (13.5.12) as a function of u_1 and the curves S_i and R_i were represented by u_2 as functions of u_1 .

As we did for the previous theorem, we want to indicate that the curves $S_i([s_i, 0]) \subset \mathbb{R}^n$ are the generalizations of the curves $S_{1,\mathbf{u}^-}$ and $S_{2,\mathbf{u}^-}$ in Example 13.5.3 about the p -system. The curve $S_i([s_i, 0])$ is called the **i^{th} Hugoniot curve** emanating from $\mathbf{u}^-$.

As we have for the p -system, one can show that for each i the curves $S_i([s_i, 0])$ and $R_i([0, r_i])$ have matching first and second order derivatives at the point $\mathbf{u}^-$ where they intersect.

We should mention that there are more than standard k -shocks in systems of conservation laws in higher dimensions. The systems may satisfy the following condition.

Definition 13.5.18

Suppose that $\lambda_1(\mathbf{u}) < \lambda_2(\mathbf{u}) < \dots < \lambda_n(\mathbf{u})$ are the eigenvalues of $DG(\mathbf{u})$ for $\mathbf{u} \in U \subset \mathbb{R}^n$, and $q_i(\mathbf{u}) \in \mathbb{R}^n$ is an eigenvector of $DG(\mathbf{u})$ associated to the eigenvalue $\lambda_i(\mathbf{u})$ for $1 \leq i \leq n$. We say that the i^{th} characteristic field is **linearly degenerate** at $\mathbf{u} \in U$ if $\nabla \lambda_i(\mathbf{u}) \cdot q_i(\mathbf{u}) = 0$, and **linearly degenerate** on U if $\nabla \lambda_i(\mathbf{u}) \cdot q_i(\mathbf{u}) = 0$ for all $\mathbf{u} \in U$.

Suppose that i^{th} characteristic field is linearly degenerate in an open set $U \subset \mathbb{R}^n$, then λ_i is constant an i -Riemann invariant. It follows that λ_i is along the integral curves associated to q_i because i -Riemann invariants are constant along the integral curves associated to q_i . If we consider the integral curve given by $\frac{dw}{ds}(s) = q_i(w(s))$ with the initial condition $w(0) = \mathbf{u}^- \in U$, we have that $\lambda_i(w(s)) = \lambda_i(w(0)) = \lambda_i(\mathbf{u}^-)$ for $|s| < \epsilon$, where ϵ is such that $w(s) \in U$ for $|s| < \epsilon$. It follows that $q_i(w(s)) = q_i(\mathbf{u}^-)$ for $|s| < \epsilon$. Hence $q_i(\mathbf{u}^-)$ is an eigenvector associated to the eigenvalue $\lambda_i(\mathbf{u}^-)$ of $(DG)(w(s))$ for $|s| < \epsilon$.

For $|s| < \epsilon$, let $z^- = \lambda_i(\mathbf{u}^-)$ and

$$u(x, y) = \begin{cases} \mathbf{u}^- & \text{if } x < yz^- \\ w(s) & \text{if } x \geq yz^- \end{cases}$$

(Figure 13.21). The function u is discontinuous along the line $\Gamma = \{(\gamma(y), y), y > 0\}$ where $\gamma(y) = yz^-$. We have that u is a strong solution of (13.5.2) satisfying the initial condition $u(x, 0) = \mathbf{u}^-$ for $x < 0$ and $u(x, 0) = w(s)$ for $x \geq 0$ because u obviously satisfies the initial condition and is also satisfies (13.5.2) outside Γ since it is constant outside Γ .

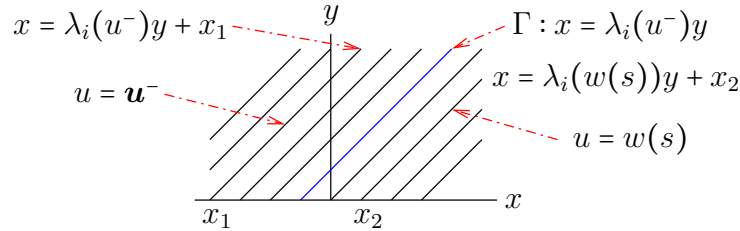


Figure 13.21: Example of a contact discontinuity.

Moreover, u satisfies the Rankine-Hugoniot formula for a shock wave as we now prove. Since

$$\begin{aligned} \frac{d}{ds} G(w(s)) - \lambda_i(\mathbf{u}^-) w(s) &= (D G)(w(s)) \frac{dw}{ds}(s) - \lambda_i(\mathbf{u}^-) \frac{\partial w}{\partial s}(s) \\ &= ((D G)(w(s)) - \lambda_i(\mathbf{u}^-) \text{Id}) q_i(w(s)) = ((D G)(w(s)) - \lambda_i(\mathbf{u}^-) \text{Id}) q_i(\mathbf{u}^-) = 0 \end{aligned}$$

for $|s| < \epsilon$, we have that $s \mapsto G(w(s)) - \lambda_i(\mathbf{u}^-) w(s)$ is constant. Therefore, $G(w(s)) - \lambda_i(\mathbf{u}^-) w(s) = G(w(0)) - \lambda_i(\mathbf{u}^-) w(0) = G(\mathbf{u}^-) - \lambda_i(\mathbf{u}^-) \mathbf{u}^-$. It follows that

$$G(w(s)) - G(\mathbf{u}^-) = \lambda_i(\mathbf{u}^-) (w(s) - \mathbf{u}^-) = \gamma'(y) (w(s) - \mathbf{u}^-)$$

as required by Rankine-Hugoniot formula.

Solutions where the shock speed equals the characteristic speed on one side are called **contact discontinuities**. This is the situation above where $\gamma'(y)$ is equal to $\lambda_i(\mathbf{u}^-)$ with $\mathbf{u}^- = u^-(\mathbf{p})$ and $\mathbf{p} = (\gamma(y), y)$.

As we did with the p -system, we can (in theory) combine i -rarefaction waves, i -shocks and contact discontinuities to create connection from $\mathbf{u}^-$ to other values of $\mathbf{u}$ that are not on the curves $S_i([s_i, 0])$ and $R_i([0, r_i])$. In fact, suppose that (13.5.2) is strictly hyperbolic system of conservation laws on an open set U and each characteristics fields is either genuinely linear or nonlinear, then at most n i -rarefaction waves, i -shocks and contact discontinuities may be used to connect $\mathbf{u}^-$ to $\mathbf{u}^+$ if it is closed enough to $\mathbf{u}^-$.

13.5.2 The Glimm Difference Scheme

Most of the discussion so far was about i -rarefaction waves and i -shocks based at $(x, y) = (0, 0)$. Obviously, using a translation, everything that we have said is also valid for i -rarefaction waves and i -shocks based at any point of $\mathbb{R} \times [0, \infty[$. In particular, we may extend the Riemann problem to system of conservation laws with piecewise constant initial

condition at $y = 0$. From there, we may also consider systems of conservation laws with any continuous initial condition at $y = 0$. The **Glimm Difference Scheme** is one of the methods that can be used to prove the existence of solutions to systems of conservation laws with continuous initial condition at $y = 0$.

We sketch the main idea behind the Glimm Difference Scheme. As usual, we assume that (13.5.2) is a genuinely linear and strictly hyperbolic system of conservation laws. We also assume that the total variation of the initial condition is **sufficiently small**; namely, that

$$\text{TV}(h) \equiv \sup_{\substack{x_1 < x_2 < \dots < x_m \\ m > 0}} \sum_{j=1}^{m-1} |h(x_{j+1}) - h(x_j)|$$

is sufficiently enough. We consider the following mesh on $\mathbb{R} \times [0, \infty[$.

$$M_{\Delta x, \Delta y} \equiv \{(j\Delta x, k\Delta y) : j \in \mathbb{Z} \text{ and } k \in \mathbb{N}\}.$$

Suppose that $U \subset \mathbb{R}^n$ is a **sufficiently small** open neighbourhood of the origin and

$$C \equiv \sup_{\substack{1 \leq j \leq n \\ \mathbf{u} \in \bar{U}}} |\lambda_j(\mathbf{u})| < \infty.$$

We assume that $\Delta x / \Delta y > C$. We select values $\mathbf{u}_{j,k} \in \mathbb{R}^n$ such that $k + j = 0 \pmod{2}$ as it follows. We set $\mathbf{u}_{j,0} = h((j+1)\Delta x + \theta_{j,0}\Delta x)$ for $j \in \mathbb{Z}$ even and $\theta_{j,0} \in [-1, 1]$ randomly selected. Suppose that we have $\mathbf{u}_{j,k} \in \mathbb{R}^n$ for $j \in \mathbb{Z}$ with $k + j = 0 \pmod{2}$ and $k \in \mathbb{N}$ fixed. For each $j \in \mathbb{Z}$ with $j + k \pmod{2}$, we solve the Riemann problem

$$\frac{\partial v}{\partial y} + D G(v) \frac{\partial v}{\partial x} = \mathbf{0}$$

on $[(j-1)\Delta x, (j+1)\Delta x] \times [k\Delta y, (k+1)\Delta y]$ with initial condition $v(x, k\Delta y) = \mathbf{u}_{j-2,k}$ for $x \leq j\Delta x$ and $v(x, k\Delta y) = \mathbf{u}_{j,k}$ for $x > j\Delta x$. Let $v_{j,k}$ be the solution of this problem. We set $\mathbf{u}_{j-1,k+1} = v_{i,j}(\mathbf{a}_{j-1,k+1})$, where $\mathbf{a}_{j-1,k+1} = (j\Delta x + \theta_{j-1,k+1}\Delta x, (k+1)\Delta y)$ and $\theta_{j-1,k+1} \in [-1, 1]$ has been randomly selected (Figure 13.22).

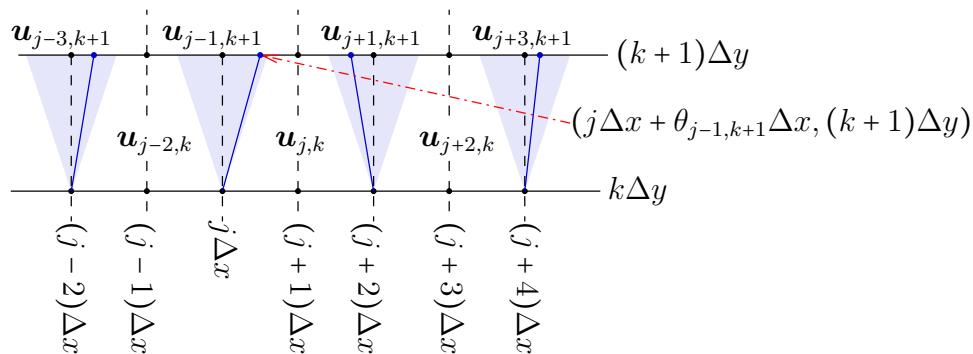


Figure 13.22: Diagram used to illustrate the Glimm Difference Scheme

The condition $\Delta x/\Delta y > C$ is essential as well as Δx , and so Δy , **sufficiently small**. The previous condition ensures that there is no interaction between the waves that we find for each value of j ; namely, the areas in blue in Figure 13.22 do not intersect. These areas contain x, y characteristics used for the connections associated to each Riemann problem. It is traditional to take $\Delta y = c\Delta x$ with $c > C$ fixed.

Let $I = \{(j, k) \in \mathbb{Z} \times \mathbb{N} : j + k = 0 \pmod{2}\}$ and $\tilde{R} = \prod_{(j,k) \in I} [-1, 1]$. For each choice of Δx and $\Theta \in \tilde{R}$, we define the following approximation $u_{\Theta, \Delta x}$ of the solution u of (13.5.2) with the initial condition $u(x, 0) = h(x)$. Set $u_{\Theta, \Delta x}(x, y) = v_{j,k}(x, y)$ for $(x, y) \in [(j-1)\Delta x, (j+1)\Delta x] \times [k\Delta x, (k+1)\Delta y]$ and $j + k = 0 \pmod{2}$. The function $u_{\Theta, \Delta x}$ is likely to be discontinuous along the lines $y = k\Delta y$.

Figure 13.23 illustrates what happens when Δx , and so Δy , converge to 0. Because of the strong law of large numbers, the segments associated to the $\mathbf{a}_{j,k}$ with $j + k = 0 \pmod{2}$ provide an approximation of the exact solution.

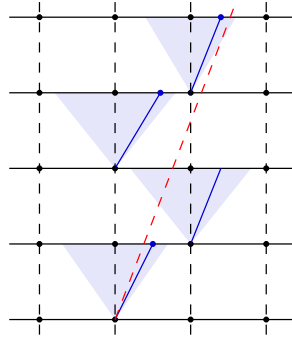


Figure 13.23: The exact x, y characteristic for the system of conservation laws with a general initial condition is the dashed line in red.

Let $R = \prod_{(j,k) \in I} [j\Delta x, (j+2)\Delta x] \times \{k\Delta y\}$. We have that $\Phi : \tilde{R} \rightarrow R$ defined by $\Phi(\Theta) = \prod_{(j,k) \in I} (j\Delta x + \theta_{j,k}\Delta x, k\Delta y)$ for all $\Theta \in \tilde{R}$ is an homeomorphism for any Δx .

One shows that there exist a set $\tilde{N}$ of measure zero in $\tilde{R}$ and a sequence $\{\Delta x_i\}_{i=1}^{\infty}$ converging to 0 such that $\{u_{\Theta, \Delta x_i}\}_{i=1}^{\infty}$ converges in $L^1_{loc}(\mathbb{R} \times [0, \infty[)$ to a solution of (13.5.2) with the initial condition $u(x, 0) = h(x)$ if $\Theta \in \tilde{R} \setminus \tilde{N}$. Recall that $\mathbf{a}_{j,k} = ((j+1)\Delta x + \theta_{j,k}\Delta x, k\Delta y) \in [j\Delta x, (j+2)\Delta x] \times \{k\Delta y\}$ and $\theta_{j,k} \in [-1, 1]$ has been randomly selected.

There are a lot of technical details that need to be proved to support the claims in the sketch of the proof. One has to properly define what is met by **sufficiently small**; in particular for the total variation $\text{TV}(h)$. This is a major constrain. A comprehensive version of Grimm's proof is given by [35]. Another reference is [19].

13.5.3 Entropy Pairs

Another issue that we should address is the entropy condition associated to a shock wave. We have used so far the Lax Shock Condition. Other, more convenient and sometime more general, entropy conditions are used for system of conservation laws in higher dimension. Under specific conditions on G in (13.5.2), they are often equivalent to the Lax Shock Condition. We mention a couple of possible extensions of the entropy condition.

Definition 13.5.19

Suppose that $G : \mathbb{R}^n \rightarrow \mathbb{R}^n$ in (13.5.2) is of class C^1 , and that $E, F : U \rightarrow \mathbb{R}$ are two continuously differentiable functions defined on an open set $U \subset \mathbb{R}^n$. If $(DF)(\mathbf{u}) = (DE)(\mathbf{u})(DG)(\mathbf{u})$ for all $\mathbf{u} \in U$, then (E, F) is called an **entropy pair** on U . The function E is called the **entropy** and the function F is called the **entropy-flux**.

Example 13.5.20

It is not always possible to find an entropy pair for a system of conservation laws in dimension higher than 2. There is however one type of systems of conservation laws for which we can always find a entropy pair.

If G is the gradient of a function $g : \mathbb{R}^n \rightarrow \mathbb{R}$, then the system of conservation laws is

$$\frac{\partial u}{\partial y}(x, y) + \frac{\partial}{\partial x} G(u(x, y)) = \frac{\partial u}{\partial y}(x, y) + \frac{\partial}{\partial x} (Dg)(u(x, y)) = 0. \quad (13.5.16)$$

We claim that $E(\mathbf{u}) = \|\mathbf{u}\|^2/2$ and $F(\mathbf{u}) = \mathbf{u}^\top (Dg)(\mathbf{u}) - g(\mathbf{u})$ define an entropy pair for the (13.5.16) on $\mathbb{R}^n$. This follows from

$$(DE)(\mathbf{u})(DG)(\mathbf{u}) = \mathbf{u}^\top (D^2 g)(\mathbf{u}) = (Dg)(\mathbf{u}) + \mathbf{u}^\top (D^2 g)(\mathbf{u}) - (Dg)(\mathbf{u}) = (DF)(\mathbf{u})$$

for all $\mathbf{u}$. ♣

If $u : \mathbb{R} \times]0, \infty[\rightarrow U$ is a classical solution of (13.5.2), we have that

$$\begin{aligned} \frac{\partial}{\partial y} E(u(x, y)) + \frac{\partial}{\partial x} F(u(x, y)) &= (DE)(u(x, y)) \frac{\partial u}{\partial y}(x, y) + (DF)(u(x, y)) \frac{\partial u}{\partial x}(x, y) \\ &= -(DE)(u(x, y))(DG)(u(x, y)) \frac{\partial u}{\partial x}(x, y) + (DF)(u(x, y)) \frac{\partial u}{\partial x}(x, y) \\ &= ((DF)(u(x, y)) - (DE)(u(x, y))(DG)(u(x, y))) \frac{\partial u}{\partial x}(x, y) \end{aligned}$$

for all $(x, y) \in \mathbb{R} \times]0, \infty[$. Therefore, if (E, F) is an entropy pair, then

$$\frac{\partial}{\partial y} E(u(x, y)) + \frac{\partial}{\partial x} F(u(x, y)) = 0 \quad (13.5.17)$$

for all $(x, y) \in \mathbb{R} \times]0, \infty[$.

Definition 13.5.21

A strong solution $u : \mathbb{R} \times]0, \infty[\rightarrow \mathbb{R}^n$ of (13.5.2) is called an **entropy solution** if there exists an entropy pair (E, F) on $\mathbb{R}^n$ such that

$$- \int_{\mathbb{R}} \int_{]0, \infty[} \left(E(u(x, y)) \frac{\partial \psi}{\partial y}(x, y) + F(u(x, y)) \frac{\partial \psi}{\partial x}(x, y) \right) dy dx \leq 0 \quad (13.5.18)$$

for all function $\psi \in \mathcal{D}(\mathbb{R} \times]0, \infty[)$ with $\psi \geq 0$ on $\mathbb{R} \times]0, \infty[$. The inequality (13.5.18) is called the **entropy inequality associated to the entropy pair (E, F)** .

It follows from (13.5.17) that we have equality to 0 in (13.5.18) if u is a classical solution of (13.5.2). However, we may have a strict inequality if u is a strong solution.

Lemma 13.5.22

Suppose that $u : \mathbb{R} \times [0, \infty[\rightarrow \mathbb{R}^n$ is a strong solution of (13.5.2) which is discontinuous along the curve $\Gamma = \{(\gamma(y), y) : y \geq y_0\}$ of class C^1 for some $y_0 \geq 0$. If u is an entropy solution associated to the entropy pair (E, F) ¹⁴, then

$$\gamma'(y) (E(u^+(\mathbf{p})) - E(u^-(\mathbf{p}))) - (F(u^+(\mathbf{p})) - F(u^-(\mathbf{p}))) \geq 0 \quad (13.5.19)$$

for all $\mathbf{p} = (\gamma(y), y) \in \Gamma$.

Proof.

The proof is very similar to the proof of the Rankine-Hugoniot formula in Section 13.3. If we use all the setup in Section 13.3, and consider

$$\begin{aligned} & - \iint_{V_l} \left(E(u(x, y)) \frac{\partial \psi}{\partial y}(x, y) + F(u(x, y)) \frac{\partial \psi}{\partial x}(x, y) \right) dy dx \\ & - \iint_{V_r} \left(E(u(x, y)) \frac{\partial \psi}{\partial y}(x, y) + F(u(x, y)) \frac{\partial \psi}{\partial x}(x, y) \right) dy dx \leq 0 \end{aligned} \quad (13.5.20)$$

obtained from (13.5.18). Recall that $\psi \in \mathcal{D}(V)$ and V is split in two by Γ with $V_l = \{(x, y) \in V : x < \gamma(y)\}$ and $V_r = \{(x, y) \in V : x > \gamma(y)\}$.

We may use the Divergence Theorem as in (13.3.2) to evaluate the first integral in (13.5.20). We get

$$\begin{aligned} & \iint_{V_l} \left(E(u(x, y)) \frac{\partial \psi}{\partial y}(x, y) + F(u(x, y)) \frac{\partial \psi}{\partial x}(x, y) \right) dy dx \\ & = \iint_{V_l} \left(\frac{\partial}{\partial y} E(u) + \frac{\partial}{\partial x} F(u) \right) \psi dx dy + \int_{\Gamma \cap \partial V_l} (F(u^-) \psi, E(u^-) \psi) \cdot \boldsymbol{\nu} ds, \end{aligned}$$

where $\boldsymbol{\nu}(\mathbf{p})$ is the unit normal to Γ at $\mathbf{p} \in \Gamma$ pointing to the right of Γ , so outside V_l , and

$$u^-(\mathbf{p}) = \lim_{t \rightarrow 0^-} u(\mathbf{p} + t \boldsymbol{\nu}(\mathbf{p}))$$

¹⁴As usual, we assume that u is of class C^1 on both sides of Γ .

for $\mathbf{p} \in \Gamma$. Since $\frac{\partial}{\partial y} E(u) + \frac{\partial}{\partial x} F(u) = 0$ in V_l because u is of class C^1 on the left side of Γ , we get

$$\begin{aligned} & \iint_{V_l} \left(E(u(x, y)) \frac{\partial \psi}{\partial y}(x, y) + F(u(x, y)) \frac{\partial \psi}{\partial x}(x, y) \right) dy dx \\ &= \int_{\Gamma \cap \partial V_l} (F(u^-) \psi, E(u^-) \psi) \cdot \boldsymbol{\nu} ds. \end{aligned} \quad (13.5.21)$$

Proceeding as we just did but using the second integral in (13.5.20), we get

$$\begin{aligned} & \iint_{V_r} \left(E(u(x, y)) \frac{\partial \psi}{\partial y}(x, y) + F(u(x, y)) \frac{\partial \psi}{\partial x}(x, y) \right) dy dx \\ &= - \int_{\Gamma \cap \partial V_r} (F(u^+) \psi, E(u^+) \psi) \cdot \boldsymbol{\nu} ds, \end{aligned} \quad (13.5.22)$$

where

$$u^+(\mathbf{p}) = \lim_{t \rightarrow 0^+} u(\mathbf{p} + t\boldsymbol{\nu}(\mathbf{p}))$$

for $\mathbf{p} \in \Gamma$. The minus sign in front of the previous line integral comes from the counter-clockwise direction of the contour integral along ∂V_r while the normal $\boldsymbol{\nu}$ is pointing inside V_r .

It follows from (13.5.20) that

$$\begin{aligned} & - \int_{\Gamma \cap \partial V_l} (F(u^-) \psi, E(u^-) \psi) \cdot \boldsymbol{\nu} ds + \int_{\Gamma \cap \partial V_r} (F(u^+) \psi, E(u^+) \psi) \cdot \boldsymbol{\nu} ds \\ &= \int_{\Gamma \cap \partial V} \psi (F(u^+) - F(u^-), E(u^+) - E(u^-)) \cdot \boldsymbol{\nu} ds \leq 0. \end{aligned}$$

Since the last inequality is true for all $\psi \in \mathcal{D}(V)$ with $\psi \geq 0$ on V , and open sets V which are split by Γ , we get

$$(F(u^+) - F(u^-), E(u^+) - E(u^-)) \cdot \boldsymbol{\nu} \leq 0$$

on Γ . Therefore,

$$\begin{aligned} 0 &\geq (F(u^+(\mathbf{p})) - F(u^-(\mathbf{p}))) + \frac{\nu_y(\mathbf{p})}{\nu_x(\mathbf{p})} (E(u^+(\mathbf{p})) - E(u^-(\mathbf{p}))) \\ &= (F(u^+(\mathbf{p})) - F(u^-(\mathbf{p}))) - \frac{d\gamma}{dy}(y) (E(u^+(\mathbf{p})) - E(u^-(\mathbf{p}))) \end{aligned}$$

for all $\mathbf{p} = (\gamma(y), y) \in \Gamma$. Note that $\nu_x(\mathbf{p}) > 0$ because $\boldsymbol{\nu}$ is the unit normal to Γ at $\mathbf{p} \in \Gamma$ pointing to the right of Γ . ■

We say that a function $f : U \rightarrow \mathbb{R}$, where U is an open subset of $\mathbb{R}^n$, is **strictly convex** on U if $D^2 f(\mathbf{u})$ (the Hessian matrix of f) is definite positive for all $\mathbf{u} \in U$.

There is a relation between the entropy jump condition (13.5.19) and the Lax Shock Condition.

Theorem 13.5.23

Suppose that (13.5.2) is a strictly hyperbolic and genuinely nonlinear system of conservation laws, and that (E, F) is an entropy pair on $\mathbb{R}^n$ with E strictly convex on $\mathbb{R}^n$. In particular, we assume that G , E and F are sufficiently differentiable. If u is a strong solution of (13.5.2) which is discontinuous along the curve $\Gamma \subset \mathbb{R} \times [0, \infty[$ of class C^1 where the Rankine-Hugoniot formula (13.5.3) is satisfied with $\|u^+(\mathbf{p}) - u^-(\mathbf{p})\|$ small enough for $\mathbf{p} \in \Gamma$, then the Lax Shock Condition (13.5.4) is satisfied if and only if the entropy jump condition (13.5.19) is strictly satisfied.

Proof.

We have from Theorem 13.5.17 that $\mathbf{u}^- = u^-(\mathbf{p})$ is connected to $\mathbf{u}^+ = u^+(\mathbf{p})$ by an i -shock (and therefore the Lax Shock Condition is satisfied) if and only if $u^-(\mathbf{p}) = S_i(0)$ and $u^+(\mathbf{p}) = S_i(s)$ for some $0 \leq i \leq 1$ and $s < 0$ ¹⁵. Moreover, we have that $G(S_i(s)) - G(S_i(0)) = \eta(s)(S_i(s) - S_i(0))$ with $\eta(0) = \lambda_i(S_i(0))$ and $\eta'(0) = 1/2$.

Let

$$H(s) = \eta(s)(E(S_i(0)) - E(S_i(s))) - (F(S_i(0)) - F(S_i(s))) .$$

We have that (13.5.19) is strictly satisfied if and only if $H(s) < 0$. We now prove that $H(s) < 0$ if and only if $s < 0$ is close enough to 0. To prove this last statement, we prove that $H(0) = \frac{dH}{ds}(0) = \frac{d^2H}{ds^2}(0) = 0$ and $\frac{d^3H}{ds^3}(0) > 0$. Hence, for $s < 0$ close to 0, the sign of $H(s)$ will be the sign of s^3 .

We obviously have that $H(0) = 0$.

If we differentiate $G(S_i(s)) - G(S_i(0)) = \eta(s)(S_i(s) - S_i(0))$ with respect to s , we get

$$(DG)(S_i(s)) \frac{dS_i}{ds}(s) = \eta'(s)(S_i(s) - S_i(0)) + \eta(s) \frac{dS_i}{ds}(s) .$$

Hence

$$\begin{aligned} \frac{dH}{ds}(s) &= \eta'(s)(E(S_i(0)) - E(S_i(s))) - \eta(s)(DE)(S_i(0)) \frac{dS_i}{ds}(s) + (DF)(S_i(s)) \frac{dS_i}{ds}(s) \\ &= \eta'(s)(E(S_i(0)) - E(S_i(s))) - \eta(s)(DE)(S_i(s)) \frac{dS_i}{ds}(s) \\ &\quad + (DE)(S_i(s))(DG)(S_i(s)) \frac{dS_i}{ds}(s) \\ &= \eta'(s)(E(S_i(0)) - E(S_i(s))) - \eta(s)(DE)(S_i(s)) \frac{dS_i}{ds}(s) \\ &\quad + (DE)(S_i(s)) \left(\eta'(s)(S_i(s) - S_i(0)) + \eta(s) \frac{dS_i}{ds}(s) \right) \\ &= \eta'(s)(E(S_i(0)) - E(S_i(s))) + \eta'(s)(DE)(S_i(s))(S_i(s) - S_i(0)) . \end{aligned}$$

¹⁵As stated, Theorem 13.5.17 does not exactly imply this. A closer look at the proof of this theorem is required to conclude that it is true if $\mathbf{u}^+$ is close enough to $\mathbf{u}^-$. This is certainly what is illustrated with the p -system.

It follows that $\frac{dE}{ds}(0) = 0$. Moreover, after some computations, we get that

$$\begin{aligned} \frac{d^2 H}{ds^2}(s) &= \eta''(s)(E(S_i(0)) - E(S_i(s))) + \eta''(s)(DE)(S_i(s))(S_i(s) - S_i(0)) \\ &\quad + \eta'(s)\left(\frac{dS_i}{ds}(s)\right)^\top ((D^2 E)(S_i(s)))(S_i(s) - S_i(0)) . \end{aligned}$$

Again, it is easy to see that $\frac{d^2 E}{ds^2}(0) = 0$. Finally, after a bit more computations, we get that

$$\frac{d^3 H}{ds^3}(s) = \eta'(s)\left(\frac{dS_i}{ds}(s)\right)^\top ((D^2 E)(S_i(s)))\frac{dS_i}{ds}(s) + \dots$$

where the $\dots$ are terms that have a factor $(E(S_i(0)) - E(S_i(s)))$ or $(S_i(0) - S_i(s))$. Hence,

$$\frac{d^3 H}{ds^3}(0) = \eta'(0)\left(\frac{dS_i}{ds}(0)\right)^\top ((D^2 E)(S_i(0)))\frac{dS_i}{ds}(0) > 0$$

because $\eta'(0) = 1/2$ and E is strictly convex. ■

It is proved in [35] that, as a consequence of the previous theorem, solutions of a strictly hyperbolic and genuinely nonlinear system of conservation laws obtained from the Glimm Difference Method satisfy (13.5.4) if there is an entropy pair (E, F) with E strictly convex.

13.5.4 Viscosity Method

Definition 13.5.21 is motivated by the physical assumption that solutions of the system of conservation laws (13.5.2) are limits (in some sense to be determined later) as $\epsilon \rightarrow 0$ of solutions of a system of partial differential equation of the form

$$\frac{\partial u^{[\epsilon]}}{\partial y} + \frac{\partial G(u^{[\epsilon]})}{\partial x} = \epsilon A \frac{\partial^2 u^{[\epsilon]}}{\partial x^2} , \quad (13.5.23)$$

where A is a positive definite $n \times n$ matrix. The term $\epsilon A \frac{\partial^2 u^{[\epsilon]}}{\partial x^2}$ represent some dissipation which in gas dynamics is associated to viscosity. The system (13.5.23) is a perturbation of the system of conservation laws (13.5.2). From an experimental point of view, (13.5.23) may be a more realistic system governing the physical phenomenon which is investigated. To simplify the discussion that follows, we also assume that A is constant.

Suppose that (E, F) is an entropy pair on an open subset $U \subset \mathbb{R}^n$ and that $u^{[\epsilon]}$ is a classical solution of (13.5.23). If we multiply both sides of (13.5.23) by $(DE)(u^{[\epsilon]})$, we get

$$\begin{aligned} (DE)(u^{[\epsilon]}) \frac{\partial u^{[\epsilon]}}{\partial y} + (DE)(u^{[\epsilon]}) \frac{\partial G(u^{[\epsilon]})}{\partial x} &= \epsilon (DE)(u^{[\epsilon]}) A \frac{\partial^2 u^{[\epsilon]}}{\partial x^2} \\ \Rightarrow (DE)(u^{[\epsilon]}) \frac{\partial u^{[\epsilon]}}{\partial y} + (DE)(u^{[\epsilon]}) (DG)(u^{[\epsilon]}) \frac{\partial u^{[\epsilon]}}{\partial x} &= \epsilon (DE)(u^{[\epsilon]}) A \frac{\partial^2 u^{[\epsilon]}}{\partial x^2} \end{aligned}$$

$$\begin{aligned}
&\Rightarrow (D E)(u^{[\epsilon]}) \frac{\partial u^{[\epsilon]}}{\partial y} + (D F)(u^{[\epsilon]}) \frac{\partial u^{[\epsilon]}}{\partial x} = \epsilon (D E)(u^{[\epsilon]}) A \frac{\partial^2 u^{[\epsilon]}}{\partial x^2} \\
&\Rightarrow \frac{\partial}{\partial y} E(u^{[\epsilon]}) + \frac{\partial}{\partial x} F(u^{[\epsilon]}) = \epsilon \frac{\partial}{\partial x} \left((D E)(u^{[\epsilon]}) A \frac{\partial u^{[\epsilon]}}{\partial x} \right) - \epsilon \left(\frac{\partial u^{[\epsilon]}}{\partial x} \right)^\top (D^2 E)(u^{[\epsilon]}) A \frac{\partial u^{[\epsilon]}}{\partial x} .
\end{aligned}$$

If we assume that $(D^2 E)(u) A$ is positively definite for all $u \in U$, and express the previous equation in the sense of distributions, we have

$$\begin{aligned}
& - \int_{\mathbb{R}} \int_{[0, \infty[} E(u^{[\epsilon]}) \frac{\partial \psi}{\partial y} dy dx - \int_{\mathbb{R}} \int_{[0, \infty[} F(u^{[\epsilon]}) \frac{\partial \psi}{\partial x} dy dx \\
& = -\epsilon \int_{\mathbb{R}} \int_{[0, \infty[} \left((D E)(u^{[\epsilon]}) A \frac{\partial u^{[\epsilon]}}{\partial x} \right) \frac{\partial \psi}{\partial x} dy dx \\
& \quad + \epsilon \int_{\mathbb{R}} \int_{[0, \infty[} \underbrace{\left(\left(\frac{\partial u^{[\epsilon]}}{\partial x} \right)^\top (D^2 E)(u^{[\epsilon]}) A \frac{\partial u^{[\epsilon]}}{\partial x} \right)}_{>0} \psi dy dx \tag{13.5.24} \\
& \leq -\epsilon \int_{\mathbb{R}} \int_{[0, \infty[} \left((D E)(u^{[\epsilon]}) A \frac{\partial u^{[\epsilon]}}{\partial x} \right) \frac{\partial \psi}{\partial x} dy dx
\end{aligned}$$

for all $\psi \in \mathcal{D}(\mathbb{R} \times [0, \infty[)$ with $\psi(x, y) \geq 0$ for all $(x, y) \in \mathbb{R} \times [0, \infty[$.

As usual, we assume that E and F are at least of class $C^2(\mathbb{R}^n)$. Moreover, suppose that E is such that $\int_{\mathbb{R}} \int_{[0, \infty[} \left((D E)(u^{[\epsilon]}) A \frac{\partial u^{[\epsilon]}}{\partial x} \right) \frac{\partial \psi}{\partial x} dy dx$ is bounded independently of ϵ . If $u^{[\epsilon]} \rightarrow u$ in the sense of distributions as $\epsilon \rightarrow 0$, where u is solution of (13.5.2), then (13.5.24) yields the entropy inequality (13.5.18) associated to the pair (E, F) as $\epsilon \rightarrow 0$. Solutions of (13.5.18) obtained as limits of solutions of (13.5.2) are called **viscosity solutions**.

The reader may find a nice formal introduction to the viscosity method in [32] to find discontinuous solutions of (13.5.2).

13.6 Addendum

We have provided a very brief introduction to the vast subject of conservation laws. To use the classical formula, we have only touched the tip of the iceberg.

For an extensive study of conservation laws and shock waves, readers should consult [6, 19, 35]. In particular, [6] provides extensive notes at the end of each chapter to direct the reader to further reading and the source of the material in the chapter. Moreover, a brief “early history of hyperbolic conservation laws” can be found in [6]. Likewise, [19] provides very detailed bibliographic notes that may guide the reader to further study of conservation laws.

The reader can find in Chapter VIII of [6] and Section 1 of Chapter III1 of [19] several results comparing different entropy conditions: Lax shock condition, entropy inequality associated to an entropy pair (E, F) , **Liu entropy condition** and **Oleinik entropy condition**.

If we refer to the conservation law (13.2.1), the Oleinik entropy condition states that

$$\frac{G(u^+(\mathbf{p})) - G(u)}{u^+(\mathbf{p}) - u} \leq \frac{G(u^+(\mathbf{p})) - G(u^-(\mathbf{p}))}{u^+(\mathbf{p}) - u^-(\mathbf{p})} \leq \frac{G(u) - G(u^-(\mathbf{p}))}{u - u^-(\mathbf{p})}$$

for all u between $u^-(\mathbf{p})$ and $u^+(\mathbf{p})$, where $\mathbf{p}$ is a point on the curve $\Gamma \subset \mathbb{R} \times [0, \infty[$ where u is discontinuous and a shock wave is generated. Oleinik entropy condition implies Lax shock condition (13.4.3). The converse is not generally true.

There are other methods to prove the existence of solutions for conservation laws than the Grimm difference method. Some of the other methods are the calculus of variations and the viscosity method.

Be aware that [35] contains more typos than is expected in a mathematical textbook. Despite that, it is an excellent book.

13.7 Exercises

Some of the problems below refer to weak solutions of partial differential equations. Using our terminology, they are solutions in the sense of distribution or, what we call, strong solutions. The expression “weak solutions” is reserved for the solutions of variational problems as we have seen in Chapter 12.

Suggested exercises:

- In [16]: numbers 3 and 6 in Section 1.6.
- In [22]: all the numbers in Sections 1.2.
- In [38]: numbers 2, 3, 5, 7, 8, 10 and 12 in Section 14.1.

Bibliography

- [1] R. Abraham, J.E. Marsden and T. Ratiu, **Manifolds, Tensor Analysis, and Applications**, Addison-Wesley Publishing Company, Inc., 1983.
- [2] R. A. Adams and J. Fournier, **Sobolev Spaces, 2nd Edition**, Academic Press, 2003.
- [3] H. Brezis, **Analyse fonctionnelle: Theorie et applications**, Masson, 1983.
- [4] A. J. Chorin and J. E. Marsden, **A Mathematical Introduction to Fluid Mechanics**, Springer-Verlag, Universitext, 1979.
- [5] J. B. Conway, **Functions of One Complex Variable**, Springer-Verlag, 1978.
- [6] C. M. Dafermos, **Hyperbolic Conservations Laws in Continuum Physics, 4th Edition**, Springer-Verlag, 2016.
- [7] W. Fleming, **Functions of Several Variables**, Springer-Verlag, 1977.
- [8] G. D. Folland, **Real Analysis**, John-Wiley, 1984.
- [9] G. D. Folland, **Advanced Calculus**, Printice Hall, 2002.
- [10] G. D. Folland, **An Introduction to Partial Differential Equations, 2nd Edition**, Princeton Academic Press, 1995.
- [11] Jack K. Hale, **Ordinary Differential equations**, Robert E. Krieger Publishing Company, 1969.
- [12] D. Henry, **Geometric Theory of Semilinear Parabolic Equations**, Lecture Notes in Mathematics, Springer-Verlag, 1974.
- [13] F. Hoppensteadt, **Mathematical Theories of Populations: Demographics, Genetics and Epidemics**, SIAM, 1975.
- [14] L. Hörmander, **Linear Partial Differential Operators**, Springer-Verlag, 1963.
- [15] A. Iserles, **A First Course in the Numerical Analysis of Differential Equations**, Cambridge University Press, 1996.
- [16] F. John, **Partial Differential Equations, 4th edition**, Springer-Verlag, 1982.

- [17] P. D. Lax, **Hyperbolic Systems of Conservation Laws and the Mathematical Theory of Shock Wave**, CBMS no. 11, SIAM, 1973.
- [18] P. D. Lax, **The Formation and Decay of Shock Wave**, Amer. Math. Monthly, 79, pp 227-241, 1972.
- [19] P. G. LeFloch, **Hyperbolic System of Conservation Laws: The Theory of Classical and Nonclassical Shock Waves**, Springer Basel AG, 2002.
- [20] J. L. Lions and E. Magenes, **Non-Homogeneous Boundary Value Problems and Applications: Vol. I**, Springer-Verlag, 1972.
- [21] J. L. Lions and E. Magenes, **Non-Homogeneous Boundary Value Problems and Applications: Vol. II**, Springer-Verlag, 1972.
- [22] R. C. McOwen, **Partial Differential Equations: Methods and Applications, 2nd Edition**, Prentice Hall, 2003.
- [23] J. R. Munkres, **Topology, A First Course**, Prentice Hall, 1975.
- [24] Y. Pinchover and J. Rubinstein, **An Introduction to Partial differential Equations**, Cambridge University Press, 2005.
- [25] M. Protter and H. Weinberger, **Maximum Principle in Differential Equations**, Prentice-Hall, 1967.
- [26] E. D. Rainville, **Special Functions**, Chelsea Publishing Company, 1971.
- [27] M. Reed and B. Simon, **Methods of Modern Mathematical Physics, Vol. I: Functional Analysis**, Academic Press, 1980.
- [28] M. Renardy and R. C. Rogers, **An Introduction to Partial Differential Equations**, Springer-Verlag, 1993.
- [29] W. Rudin, **Principles of Mathematical Analysis**, McGraw-Hill, 1976.
- [30] W. Rudin, **Real and Complex Analysis**, McGraw-Hill, 1974.
- [31] W. Rudin, **Functional Analysis**, McGraw-Hill, 1973.
- [32] S. Salsa, **Partial Differential Equations in Action: From Modelling to Theory**, Springer-Verlag, 2008.
- [33] G. F. Simmons, **Differential Equations with Applications and Historical Notes**, McGraw-Hill, 1972.
- [34] G. D. Smith, **Numerical Solution of Partial Differential Equations, 2nd Edition**, Oxford University Press, 1978.
- [35] J. Smoller, **Shock Waves and Reaction-Diffusion Equations, 2nd Edition**, Springer-Verlag, 1994.

-
- [36] S. L. Sobolev, **Partial Differential Equations of Mathematical Physics, 3rd Edition**, Pergamon Press, Addison-Wesley Publishing Company, Inc, 1964.
 - [37] M. Spivak, **A Comprehensive Introduction to Differential Geometry, Volume One, 3rd Edition**, Publish or Perish, Inc, 2005.
 - [38] W. A. Strauss, **Partial Differential Equations: an Introduction, 2nd Edition**, John-Wiley & Sons, 2008.
 - [39] F. Trèves, **Basic Linear Partial Differential Equations**, Academic Press, Inc., 1975.

Index

- 2-norm, [6](#)
- k -norm, [6](#)
- Adjoint Variational Problem, [486](#)
 - Self Adjoint, [487](#)
- Analytic Field, [71](#)
- Approximate Identity, [153](#)
- Associated Legendre Equation, [128](#)
- Associated Legendre Polynomials, [129](#)
- Banach Space, [75](#)
- Barrier Condition, [282](#)
- Bessel Differential Equation, [130](#)
- Bessel Functions
 - Of the First Kind, [130](#)
 - Of the Second Kind, [131](#)
- Bounded Continuous Functions, [7](#)
- Bounded Inverse Theorem, [92](#)
- Bounded Uniformly Continuous Functions, [7](#)
- Cauchy Problem, [60](#), [69](#)
- Characteristic Curve, [48](#), [61](#)
- Characteristic Form, [63](#)
- Characteristic Surface, [63](#)
- Classical Fourier Expansion, [82](#)
- Classical Fourier expansion, [85](#)
- Classical Solution, [140](#)
- Closed Graph Theorem, [92](#), [97](#)
- Complete Orthogonal Set, [78](#)
- Complete Orthonormal Set, [78](#)
- Conjugate Exponent, [150](#)
- Conjugate-Linear mapping, [98](#)
- Connected Set, [238](#)
- Conservation Laws, [555](#)
 - i^{th} Characteristic Field, [590](#)
 - i^{th} Hugoniot Curve, [596](#)
 - i^{th} -Rarefaction Wave, [591](#), [594](#)
 - i -Riemann Invariant, [592](#)
- k -Shock, [581](#)
- x, y Characteristics, [557](#)
- Back Rarefaction Wave, [586](#)
- Back Shock, [583](#)
- Burgers' Equation, [556](#)
- Contact Discontinuity, [597](#)
- Entropy, [600](#)
- Entropy Condition, [569](#), [578](#)
- Entropy Inequality, [568](#)
- Entropy Inequality Associated to the Entropy Pair (E, F) , [601](#)
- Entropy Pair, [600](#)
- Entropy Solution, [601](#)
- Entropy-Flux, [600](#)
- Expansive Wave, [565](#)
- Front Rarefaction Wave, [587](#)
- Front Shock, [583](#)
- Genuinely Linear, [568](#)
- Genuinely Nonlinear, [568](#), [590](#)
- Glimm Difference Scheme, [598](#)
- Integral Curve, [591](#)
- Lax Shock Condition, [567](#), [581](#)
- Linearly Degenerate, [596](#)
- Liu Entropy Condition, [605](#)
- Nonlinear Wave Equation, [580](#)
- Oleinik Entropy Condition, [605](#)
- Rankine-Hugoniot Formula, [561](#), [580](#)
- Rarefaction Wave, [565](#)
- Riemann Problem, [566](#), [583](#)
- Signal, [557](#)
- Signal Speed, [557](#)
- Strong Solution, [558](#)
- Viscosity Solution, [605](#)
- Continuous Kernel of Order α , [302](#)
- Continuous Linear Functional, [77](#)
- Convex Space, [76](#)
- Dirac Delta Function, [142](#)

- Dirichlet Form, 469
- Dirichlet kernel, 83
- Dirichlet Problem
 - Classical Solution, 468
 - Strong Solution, 468
 - Weak Solution, 470
- Distribution, 141
 - Convergence, 143
 - Convolution, 153, 162
 - Fourier Transform, 162
 - Partial Derivative, 144
 - Product, 142
 - Support, 143
 - Tempered, 161
- Divergence, 7
- Divergence Theorem, 232
- Dual Space, 93
- dual space, 77
- Eikonal Equation, 31
- Elliptic Partial Differential Equation
 - Free Boundary Conditions, 475
 - Natural Boundary Conditions, 475
 - Second Order, 50
 - Standard Form, 53
- Elliptic Regularity Theorem, 526
- Euclidean Norm, 6
- Euler's Relation, 271
- Exterior Normal Derivative, 285
- First Order Partial Differential Equation
 - Cauchy Problem, 9, 20
 - Characteristic Curve, 10, 25
 - Characteristic Equations, 10, 25
 - Characteristic Strip, 25
 - Complete Integral, 35
 - Envelope, 18
 - General Solution, 16
 - Integral Surface, 9, 20
 - Linear, 9
 - Monge Cone, 21
 - Quasi-Linear, 9
- Fourier Cosine Series, 88
- Fourier Sine Series, 88
- Functions
 - Convolution, 151, 156
 - Fourier Transform, 155
 - Rapidly Decreasing Functions, 159
 - Schwartz Class, 159
 - Vanish at Infinity, 158
- Fundamental Solution, 167, 373
- Gårding's Inequality, 483
- Gamma Function, 129
- Generalized Fourier Series, 79
- Generalized Young's Inequality, 150
- Gradient, 7
- Green's Function, 248, 256, 262, 263, 373
- Green's Identities, 232
- Harmonic Functions, 231, 232
- Heat Equation, 369
 - Gauss-Weierstrass Kernel, 375
 - Gaussian Kernel, 375
 - Heat Kernel, 375
- Hilbert Space, 76
- Hilbert-Schmidt Kernel, 101
- Hilbert-Schmidt Theorem, 100
- Hyper-Surface, 62
- Hyperbolic Partial Differential Equation
 - Second Order, 50
 - Standard Form, 54
- Hölder Continuous Functions of Exponent μ , 438
- Indicial Equation, 110
- Induce Norm, 92
- Inner Product, 75
- Inner Product Space, 75
- Interior Normal Derivative, 285
- Internal Regularity, 519
- Kelvin Transform, 288
- Kernel of Order 0, 302
- Kernel of Order α , 302
- Laplace Equation, 231
 - Variational Formulation, 502
- Laplace Operator, 209
- Laplacian, 7
 - Eigenfunction, 360

- Eigenvalue, 360
- Eigenvector, 360
- Spherical Harmonics, 368
- Surface Harmonics, 366
- Lax-Milgram Theorem, 106
- Layer Potentials, 293
- Legendre Differential Equation, 124
- Legendre Polynomials
 - Rodrigue's Formula, 124
 - Standardized, 124
- Linear Functional, 91
- Linear Mapping, 91, 96
 - Adjoint, 94, 98
 - Bounded, 92, 96
 - Closed, 92, 97
 - Closure, 97
 - Compact, 94, 98
 - Compression Spectrum, 97
 - Continuous Spectrum, 97
 - Discrete Spectrum, 97
 - Domain, 96
 - Eigenspace, 93, 97
 - Eigenvalue, 93, 97
 - Eigenvectors, 93, 97
 - Extension, 96
 - Graph, 92, 97
 - Hilbert Space Adjoint, 99
 - Kernel, 91, 96
 - Point Spectrum, 97
 - Range, 91, 96
 - Residual Spectrum, 97
 - Resolvent, 93, 97
 - Resolvent Set, 92, 97
 - Self Adjoint, 100
 - Spectral Radius, 93
 - Spectrum, 92, 97
 - Symmetric, 99
- Linear Operator, 91, 96
- Linear Partial Differential Operator
 - Adjoint, 140
 - Divergence Form, 467
 - Elliptic, 64, 467
 - Hyperbolic, 64
 - Hypoelliptic, 169
 - Locally Solvable, 459
 - Parabolic, 64
 - Strongly Elliptic, 65, 467
 - Ultrahyperbolic, 64
 - Uniformly Elliptic, 65
- Linear Second Order Partial Differential Equation, 45
- Liouville's Theorem, 236
- Local Regularity, 519
- Locally Finite Open Cover, 430
- Malgrange-Ehrempreis' Theorem, 168
- Mean Value Theorem, 233
- Method of Descent, 220
- Method of Difference Quotient, 507
- Method of Translations, 507
- Negative Definite, 51
- Neumann Problem
 - Classical Solution, 470
 - Strong Solution, 471
 - Weak Solution, 471
- Newton's Law of Heating, 231
- Newton's Law of heating, 369
- Norm Induced, 76
- Normal Set of Differential Operators, 473
- Open Mapping Theorem, 92
- Ordinary Differential Equation, 10
 - Orbit, 11
- Orthogonal Basis, 78
- Orthogonal Set, 77
- Orthogonal Vectors, 77
- Orthonormal Basis, 78
- Orthonormal Set, 77
- Parabolic Partial Differential Equation
 - Second Order, 50
 - Standard Form, 59
- Partial Differential Equation, 9
- Partition of Unity, 297, 298
- Perpendicular Vectors, 77
- Piecewise Continuous, 80
- Piecewise Differentiable, 82
- Plancherel's Theorem, 159
- Poincaré's Inequality, 441
- Poisson Equation, 232

- Dirichlet Problem, 245
- Neumann Problem, 246
- Poisson Kernel, 256, 258, 264
- Poisson Summation Formula, 156
- Positive Definite, 51
- Positive Semi-Definite, 542
- Potential Equation, 231
- Potential Functions, 231
- Precompact Set, 94
- Principal Symbol, 49, 59, 62, 63

- Reflexive Space, 98
- Refraction Index, 31
- Regular Point, 110
- Regular Singular Point, 110
- Rellich's Theorem, 439
- Riesz Representation Theorem, 77

- Scalar Product, 75
- Schwarz Inequality, 76
- Second Order Semi-Linear Partial Differential Equation, 59
- Segment Property, 401
- Sequentially Compact Set, 94
- Sequilinear Form
 - Bounded, 102
 - Coercive, 103
 - Continuous, 102
 - Symmetric, 102
- Sobolev Space, 140
 - $H^{k,p}(\Omega)$, 420
 - $H_0^{k,p}(\Omega)$, 420
 - $H_0^{s,2}(\Omega)$, 422
 - $W_0^{k,p}(\Omega)$, 416
 - $W^{s,2}(\mathbb{R}^n)$, 400
 - $W_0^{s,2}(\Omega)$, 416
 - $W^{k,p}(\Omega)$, 397
 - $(\mathbf{k}, \mathbf{p})$ -Extension Operator, 427
 - $(\mathbf{k}, \mathbf{p})$ -Extension Property, 427
 - Dual Space, 423
 - Localized, 413
 - Strong $\mathbf{k}$ -Extension Operator, 427
 - Strong $\mathbf{k}$ -Extension Property, 427
 - Total Extension Operator, 427
 - Trace Operator, 444, 446, 448, 451
- Sobolev's Lemma, 436
- Solution in the Sense of Distribution, 140, 163
- Source Function, 373
- Spherical Mean, 212
- Stampacchia's Theorem, 104
- Strictly Convex, 602
- Strictly Hyperbolic Partial Differential Equation, 580
- Strong Solution, 140, 163
- Sturm-Liouville Problem, 119
 - Eigenfunctions, 119
 - Eigenvalue, 119
 - Regular, 119
 - Regular Singular, 119
 - Singular, 119
- Subharmonic Functions, 237
- Supremum Norm of a Function, 6
- System of Linear First Order Partial Differential Equations, 60

- Test Functions, 140
 - Convergence, 141
- Theory of Distributions, 140
- Topological Vector Space, 154
- Triangle Inequality, 76

- Uniform $\mathbf{C}^k$ -Regularity Condition, 430

- Wave Equation, 173, 205
 - n^{th} -Harmonic, 189, 229
 - n^{th} -Normal Mode, 189, 229
 - Consistency Conditions, 180
 - Domain of Determinacy, 175, 178
 - Energy, 207
 - Fundamental Harmonic, 189, 229
 - Modal Lines, 230
 - Non-Homogeneous, 180
 - Propagation Speed, 173
 - Range of Influence, 175
 - Space-Like, 205
 - Transmission Condition, 186
- Wave Number, 31
- Wavefronts, 31
- Wavelength, 31
- Weakly Coercive Bounded Sequilinear Mapping, 476

Young's Inequality, [151](#)