Berkeley
UNIVERSITY OF CALIFORNIA

Math 228A

Partial Differential Equations

Professor: Daniel Tataru
Notes By: Yuhang Cai

2021 Fall

CONTENTS

1 Overview **5**

1.1 Syllabus . . . 5
1.2 Overview . . . 6
1.2.1 Strategies for studying PDE . . . 9
1.3 First order nonlinear scalar PDEs I . . . 11
1.3.1 Local Solution . . . 11
1.4 First order nonlinear scalar PDES II . . . 13
1.4.1 Global Solution . . . 14
1.4.2 Continuous Dependence on Data . . . 14
1.5 First order nonlinear Scalar PDEs III . . . 16
1.5.1 First order scalar PDE . . . 17
1.5.2 Linear problem . . . 17
1.6 First order nonlinear scalar PDEs IV . . . 18
1.6.1 Local solutions for linear, scalar PDEs . . . 18
1.6.2 Semilinear PDEs . . . 20
1.6.3 quasilinear PDEs . . . 20
1.7 Qusilinear and Nonlinear First Order PDEs . . . 22
1.7.1 quasilinear PDEs and conversation laws . . . 22
1.7.2 Burgers' equation . . . 23
1.7.3 Fully nonlinear problems . . . 24
1.8 Existence of Solutions to Nonlinear First Order Scalar PDEs . . . 27
1.8.1 Existence and uniqueness given initial data . . . 27
1.8.2 Problems in standard form . . . 28
1.9 Hamilton-Jacobi Equations and Calculus of Variations . . . 31
1.9.1 Recap: Hamilton-Jacobi equations . . . 31
1.9.2 Calculus of variations . . . 31
1.9.3 Connecting the Hamilton-Jacobi equations to the Euler Lagrange equations . . . 34
1.10 Hamilton-Jacobi Equations & Calculus of Variation II . . . 36
1.10.1 Recap: Connecting H-J equation to calculus of variation via Legendre transform 36
1.10.2 Existence of minimizers for the Euler-Lagrange equation . . . 37
1.10.3 E-L equation minimizers slove H-J equations . . . 39
1.11 The Hopf-Lax Solution to Hamilton-Jacobi Equations . . . 41
1.11.1 The Hamiltonian in classical mechanics . . . 41
1.11.2 The Hopf-Lax formula . . . 41
1.11.3 Properties of the Hopf-Lax solution . . . 42
1.11.4 Almost everywhere solvability of the H-J equation . . . 43
1.12 Introduction to Distribution Theory . . . 47
1.12.1 Weak solutions to PDEs . . . 47

1.12.2 Topologies on vector space . . . 48
1.12.3 Examples of distributions . . . 50
1.13 Distribution & Homogeneous Distribution . . . 52
1.13.1 Operations on distributions . . . 52
1.13.2 The support of a distribution . . . 53
1.13.3 Homogeneous distributions . . . 54
1.14 Homogeneous Distribution of Order -1, Convolution, and Fundamental Solutions . . . 57
1.14.1 Special homogeneous distributions of order -1 . . . 57
1.14.2 Convolution . . . 58
1.14.3 Fundamental solutions to PDEs . . . 59
1.15 Fundamental Solutions for 2-D PDE & Laplacian . . . 62
1.15.1 Fundamental solutions in 1 and 2 dimensions . . . 62
1.15.2 Fundamental solution for the Laplacian . . . 63
1.16 Introduction to the Fourier Transform . . . 67
1.16.1 Motivation: diagonalization for differential operators . . . 67
1.16.2 Properties of the Fourier transform . . . 67
1.17 Fourier Inversion, Plancherel's Theorem and Temper Distributions . . . 72
1.17.1 Fourier inversion . . . 72
1.17.2 Isometry properties of $\mathcal{F}$ on L^2 . . . 72
1.17.3 Temperate distributions . . . 73
1.17.4 Examples of temperate distributions . . . 74
1.17.5 The Fourier transform of δ_0 and H . . . 74
1.18 Using the Fourier Transform to Find Fundamental Solutions . . . 76
1.18.1 The Paley-Wiener theorem and the Fourier transform of even and odd functions 76
1.18.2 Using the Fourier transform to find fundamental solutions . . . 76
1.18.3 Fundamental solution of the heat equation . . . 78
1.19 Schrödinger Equation, the Uncertainty Principle, and Oscillatory Integrals . . . 80
1.19.1 Fundamental solution of the Schrödinger equation . . . 80
1.19.2 The uncertainty principle . . . 81
1.19.3 Oscillatory fintegrals and the KdV equation . . . 82
1.20 The KdV Equation and the Wave Equation . . . 84
1.20.1 Fundamental solution of the KdV equation . . . 84
1.20.2 Analysis of the wave equation . . . 85
1.21 Fundamental Solutions to the Wave Equation . . . 87
1.21.1 Lorentz invariance of fundamental solution . . . 87
1.21.2 Calculation of fundamental solutions . . . 88
1.21.3 Determination of constants for fundamental solutions . . . 90
1.21.4 Physical interpretation of solutions to the wave equation . . . 91
1.21.5 Next steps: Fourier series . . . 91
1.22 Fourier Transforms of Periodic Functions and Local Solvability of Partial Differential Operators . . . 93
1.22.1 Fourier transforms of periodic functions . . . 93
1.22.2 Local solvability of partial differential operator . . . 94
1.23 Properties of Harmonic Functions . . . 97
1.23.1 Elliptic regularity . . . 97
1.23.2 The maximum principle . . . 98
1.23.3 Liouville's thoerem . . . 100
1.23.4 Boundary value problem . . . 101
1.24 Boundary Value Problem for the Laplace Equation . . . 103
1.24.1 The Dirichlet and Neumann Problems . . . 103
1.24.2 Uniqueness concerns for the Dirichlet and Neumann problem . . . 103
1.24.3 Existence using energy type estimates . . . 104
1.24.4 Green's functions for domains with boundary . . . 106

1.25 Boundary Value Problem for the Heat Equation 109
1.25.1 Properties of the heat equation 109
1.25.2 The mean value property and the maximum principle 109
1.25.3 Energy estimates 113
1.26 Initial Value Problems and Energy Estimates for the Wave Equation 114
1.26.1 Initial value problems for the wave equation 114
1.26.2 Energy estimates for the wave equation 116
1.26.3 Finite speed of propagation via energy estimates 118

CHAPTER 1

OVERVIEW

Date: Aug 26, 2021

1.1 Syllabus

Instructor: Daniel Tataru (tataru@math.berkeley.edu)

This is the first semester of a two semester sequence in Partial Differential Equations. It is assumed you are familiar with undergraduate real analysis and have some knowlege of ordinary differential equations. For the second half of the semester some measure theory is also needed. Complex analysis is useful but not required. The course roughly follows at first Hormander (theory of distributions and Fourier Analysis) and continues with the first part and the beginning of the second part Evans but with various omissions/additions. Topics to be covered include (not necessarily in this order):

- Transport equation
- Nonlinear first order Equations
- Theory of distributions
- Fourier analysis
- The Laplace equation
- The heat equation
- The Schrodinger equation
- The Wave equation
- Sobolev Space

Time:

- Tue-Thu 8:00-9:00, 70 Evans
- Unless otherwise announced, the class will meet in person
- Backup Zoom link

Office hours:

- Wed 10:00-12:00
- Office hours will be held virtually unless otherwise announced
- OH Zoom link

Textbook:

- *Partial Differential Equations (2rd Edition)* by L.C.Evans
- *Partial Differential Equations, vol. 1* by L.Hormander
- *Guide to Distribution THeory and Fourier Transform* by Robert Strichartz

Grading:

- Homework, midterm and final, equally weighted.
- **Midterm:** Thursday, Oct 21.
- **Final:** Take home, to be assigned on Wed eve., Dec 8, and to be returned no later than Wed, Dec 15, at noon.

1.2 Overview

Firstly, we review the definition of ODE:

Definition 1.1 (ODE).
Given a function $\mu : \mathbb{R} \to \mathbb{R}$(or $\mathbb{C}$), an expression of the form:

$$F(\mu, \mu', \mu'', \ldots, \mu^{(n)}) = 0,$$

is called a n^{th}-order ordinary differential equation (ODE), where

$$F : \mathbb{R}^{n+1} \to \mathbb{R}.$$

Compared to ODE, A *partial differential equation* (PDE) is an equation involving an unknown function of two or more variables and certain of its partial derivatives.

Definition 1.2 (ODE).
Fix an integer $k \geqslant 1$ and let U denote an open subset of $\mathbb{R}^n$. An expression of the form

$$F\left(D^k u(x), D^{k-1}u(x), \ldots, Du(x), u(x), x\right) = 0 \quad (x \in U) \tag{1.1}$$

is called a k$^{\text{th}}$-order partial differential equation, where

$$F : \mathbb{R}^{n^k} \times \mathbb{R}^{n^{k-1}} \times \cdots \times \mathbb{R}^n \times \mathbb{R} \times U \to \mathbb{R}$$

is given and

$$u : U \to \mathbb{R}$$

is the unknown.

Definition 1.3 (Linear, semilinear, quasilinear and fully nonlinear PDE).

(i) The PDE (1.1) is called linear if it has the form

$$\sum_{|\alpha| \leqslant k} a_\alpha(x) D^\alpha u = f(x)$$

for given functions $a_\alpha(|\alpha| \leqslant k), f$. This linear PDE is homogeneous if $f \equiv 0$.

(ii) The PDE (1.1) is semilinear if if has the form

$$\sum_{|\alpha| = k} a_\alpha(x) D^\alpha u + a_0\left(D^{k-1}u, \ldots, Du, u, x\right) = 0.$$

(iii) The PDE (1.1) is quasilinear if it has the form

$$\sum_{|\alpha| = k} a_\alpha\left(D^{k-1}u, \ldots, Du, u, x\right) D^\alpha u + a_0\left(D^{k-1}u, \ldots, Du, u, x\right) = 0.$$

(iv) The PDE (1.1) is fully nonlinear if it depends nonlinearly upon the highest order derivatives.

A system of partial differential equations is, informally speaking, a collection of several PDE for several unknown functions.

Definition 1.4 (System of PDE).
An expression of the form

$$\mathbf{F}\left(D^k\mathbf{u}(x), D^{k-1}\mathbf{u}(x), \dots, D\mathbf{u}(x), \mathbf{u}(x), x\right) = \mathbf{0} \quad (x \in U) \tag{1.2}$$

is called a k^{th} -order system of partial differential equations, where

$$\mathbf{F} : \mathbb{R}^{mn^k} \times \mathbb{R}^{mn^{k-1}} \times \cdots \times \mathbb{R}^{mn} \times \mathbb{R}^m \times U \to \mathbb{R}^m$$

is given and

$$\mathbf{u} : U \to \mathbb{R}^m, \mathbf{u} = \left(u^1, \dots, u^m\right)$$

is the unknown.

Here we are supposing that the system comprises the same number m of scalar equations as unknowns $(u^1, \dots, u^m)$. This is the most common circumstance, although other systems may have fewer or more equations than unknowns. Systems are classified in the obvious way as being linear, semilinear, etc.

Following is a list of many specific partial differential equations of interest in current research. This listing is intended merely to familiarize the reader with the names and forms of various famous PDE. To display most clearly the mathematical structure of these equations, we have mostly set relevant physical constants to unity. We will later discuss the origin and interpretation of many of these PDE.

Throughout $x \in U$, where U is an open subset of $\mathbb{R}^n$, and $t \geqslant 0$. Also $Du = D_x u = (u_{x_1}, \dots, u_{x_n})$ denotes the gradient of u with respect to the spatial variable $x = (x_1, \dots, x_n)$. The variable t always denotes time.

Example 1.5.
Single partial differential equations:
a. Linear equations:

- Laplace:
$$\Delta u = \sum_{i=1}^{n} u_{x_i x_i} = 0.$$
- Helmholtz:
$$-\Delta u = \lambda u.$$
- Linear transport:
$$u_t + \sum_{i=1}^{n} b^i u_{x_i} = 0.$$
- Liouville:
$$u_t - \sum_{i=1}^{n} (b^i u)_{x_i} = 0.$$
- Heat:
$$u_t - \Delta u = 0.$$
- Schrodinger:
$$iu_t + \Delta u = 0.$$
- Kolmogorov:
$$u_t - \sum_{i,j=1}^{n} a^{ij} u_{x_i x_j} + \sum_{i=1}^{n} b^i u_{x_i} = 0.$$

- Fokker-Planck:
$$u_t - \sum_{i,j=1}^{n} (a^{ij}u)_{x_i x_j} - \sum_{i=1}^{n} (b^i u)_{x_i} = 0.$$
- Wave:
$$u_{tt} - \Delta u = 0.$$
- Klein-Gordon:
$$u_{tt} - \Delta u + m^2 u = 0.$$
- Telegraph:
$$u_{tt} + 2du_t - u_{xx} = 0.$$
- General wave:
$$u_{tt} - \sum_{i,j=1}^{m} a^{ij} u_{x_i x_j} + \sum_{i=1}^{n} b^i u_{x_i} = 0.$$
- Airy:
$$u_t + t_{xxx} = 0.$$
- Beam:
$$u_{tt} + u_{xxxx} = 0.$$

b. Nonlinear equations:

- Eikonal:
$$|Du| = 1.$$
- Nonlinear Poisson:
$$-\Delta u = f(u).$$
- p-Laplacian:
$$div(|Du|^{p-2} Du) = 0.$$
- Minimal surface:
$$\operatorname{div}\left(\frac{Du}{(1+|Du|^2)^{1/2}}\right) = 0.$$
- Monge-Ampere:
$$\det(D^2 u) = f.$$
- Hamilton-Jacobi:
$$u_t + H(Du, x) = 0.$$
- Scalar conservation law:
$$u_t + div\mathrm{F}(u) = 0.$$
- Scalar reaction-diffusion:
$$u_t - \Delta u = f(u).$$
- Porous medium:
$$u_t - \Delta(u^\gamma) = 0.$$
- Nonlinear wave:
$$u_{tt} - \Delta u + f(u) = 0.$$
- Korteweg-de Vries (KdV):
$$u_t + uu_x + u_{xxx} = 0.$$
- Nonlinear Schrodinger:
$$iu_t + \Delta u = f(|u|^2)u.$$

Example 1.6.
Systems of PDEs
a. Linear Systems

- Equilibrium equations of linear elasticity:

$$\mu\Delta u + (\lambda + \mu)D(divu) = 0.$$

- Evolution equations of linear elasticity:

$$u_{tt} - \mu\Delta u - (\lambda + \mu)D(divu) = 0.$$

- Maxwell's equation:

$$\begin{cases} \mathbf{E}_t = \operatorname{curl} \mathbf{B} \\ \mathbf{B}_t = -\operatorname{curl} \mathrm{E} \\ \operatorname{div} \mathbf{B} = \operatorname{div} \mathbf{E} = 0 \end{cases}$$

b. Nonlinear systems

- System of conservation laws:

$$\mathbf{u}_t + \operatorname{div} \mathbf{F}(\mathbf{u}) = \mathbf{0}.$$

- Reaction-diffusion system:

$$\mathbf{u}_t - \Delta \mathbf{u} = \mathbf{f}(\mathbf{u}).$$

- Euler's equations for incompressible, inviscid flow:

$$\begin{cases} \mathbf{u}_t + \mathbf{u} \cdot D\mathbf{u} = -Dp \\ \operatorname{div} \mathbf{u} = 0 \end{cases}$$

- Navier-Stokes equations for incompressible, viscous flow:

$$\begin{cases} \mathbf{u}_t + \mathbf{u} \cdot D\mathbf{u} - \Delta \mathbf{u} = -Dp \\ \operatorname{div} \mathbf{u} = 0 \end{cases}$$

1.2.1 Strategies for studying PDE

The informal notion of a well-posed problem captures many of the desirable features of what it means to solve a PDE. We say that a given problem for a partial differential equation is well-posed if

(i) the problem in fact has a solution;
(ii) this solution is unique;
(iii) the solution depends continuously on the data given in the problem.

The last condition is particularly important for problems arising from physical applications: we would prefer that our (unique) solution changes only a little when the conditions specifying the problem change a little. (For many problems, on the other hand, uniqueness is not to be expected. In these cases the primary mathematical tasks are to classify and to characterize the solutions.)

By solving a partial differential equation in the classical sense we mean if possible to write down a formula for a classical solution satisfying (i)–(iii) above, or at least to show such a solution exists, and to deduce various of its properties.

But can we achieve this? The answer is that certain specific partial differential equations (e.g. Laplace's equation) can be solved in the classical sense, but many others, if not most others, cannot. Consider for instance the scalar conservation law

$$u_t + F(u)_x = 0.$$

We will see in §3.4 that this PDE governs various one-dimensional phenomena involving fluid dynamics, and in particular models the formation and propagation of shock waves. Now a shock wave is a

curve of discontinuity of the solution u; and so if we wish to study conservation laws, and recover the underlying physics, we must surely allow for solutions u which are not continuously differentiable or even continuous. In general, as we shall see, the conservation law has no classical solutions but is well-posed if we allow for properly defined generalized or weak solutions.

The point is this: if from the outset we demand that our solutions be very regular, say k-times continuously differentiable, then we are usually going to have a really hard time finding them, as our proofs must then necessarily include possibly intricate demonstrations that the functions we are building are in fact smooth enough. A far more reasonable strategy is to consider as separate the existence and the smoothness (or regularity) problems. The idea is to define for a given PDE a reasonably wide notion of a weak solution, with the expectation that since we are not asking too much by way of smoothness of this weak solution, it may be easier to establish its existence, uniqueness, and continuous dependence on the given data. Thus, to repeat, it is often wise to aim at proving well-posedness in some appropriate class of weak or generalized solutions.

Now, as noted above, for various partial differential equations this is the best that can be done. For other equations we can hope that our weak solution may turn out after all to be smooth enough to qualify as a classical solution. This leads to the question of regularity of weak solutions. As we will see, it is often the case that the existence of weak solutions depends upon rather simple estimates plus ideas of functional analysis, whereas the regularity of the weak solutions, when true, usually rests upon many intricate calculus estimates.

Let me explicitly note here that our efforts will be largely devoted to proving mathematically the existence of solutions to various sorts of partial differential equations, and not so much to deriving formulas for these solutions. This may seem wasted or misguided effort, but in fact mathematicians are like theologians: we regard existence as the prime attribute of what we study. But unlike theologians, we need not always rely upon faith alone.

Typical Difficulties: Following are some vague but general principles, which may be useful to keep in mind:

(1) Nonlinear equations are more difficult than linear equations; and, indeed, the more the nonlinearity affects the higher derivatives, the more difficult the PDE is.
(2) Higher-order PDE are more difficult than lower-order PDE.
(3) Systems are harder than single equations.
(4) Partial differential equations entailing many independent variables are harder than PDE entailing few independent variables.
(5) For most partial differential equations it is not possible to write out explicit formulas for solutions.

None of these assertions is without important exceptions.

1.3 First order nonlinear scalar PDEs I

Date: Aug 31, 2021

In the following three weeks, we are going to consider

$$F(x, \mu, \partial\mu) = 0.$$

Firstly, we are going to consider its functional spaces:

- Which function could be a solution?
- How do we verify that a function is a solution?

We define some common functions spaces first:

Definition 1.7 (C, C^m, C^m_{loc} and C^∞ functions).
Given $\mathbb{R}^n$,

- $\mathrm{C}(\mathbb{R}^n)$ is the set of bounded continuous functions, and $\mathrm{C}_{loc}(\mathbb{R}^n)$ is the set of continuous function with $\mathrm{C}(\mathbb{R}^n) \subset \mathrm{C}_{loc}(\mathbb{R}^n)$. For $\mu \in \mathrm{C}(\mathbb{R}^n)$, the norm is defined as

$$\|\mu\|_{\mathrm{C}} = \sup_{x\in\mathbb{R}^n} |\mu(x)|.$$

- $\mathrm{C}^m(\mathbb{R}^n)$ is the set of bounded functions with m-order derivatives. For any derivative with order smaller than m, the derivative is also bounded. $\mathrm{C}^m_{loc}(\mathbb{R}^n)$ is the set of functions only m-order differentiable. The norm of $\mathrm{C}^m(\mathbb{R}^n)$ is defined as:

$$\|u\|_{\mathrm{C}^m} = \sum_{i=0}^{n} \|D^i u\|_{\mathrm{C}}.$$

- $\mathrm{C}^\infty(\mathbb{R}^n) = \cap\, \mathrm{C}^m_{loc}(\mathbb{R}^n)$.

For C_{loc}, we cannot define norms but we can consider a series of interval $I_N = [-N, N]$, and

$$\|\mu\|_{C[I_N]} = \sup_{x\in I_N} |\mu(x)|.$$

We say $\mu_n \to \mu$ in C_{loc}, if $\|\mu_n - \mu\|_{C[I_N]}$ converges to 0 for all N. We call space with such a series of norms the locally convex space, which is the extension of norm space.

Now we review some knowledge in ODE. Given ODE,

$$\mu : \mathbb{R} \to \mathbb{R}^n, \begin{cases} \mu'(x) = F(x, \mu(x)) \\ \mu(0) = \mu_0 \end{cases}$$

We may ask about?

- The existence of solution
- The uniqueness of solution
- This depends on the initial data
- Local vs global equations

1.3.1 Local Solution

At a minimum, we may ask that F is continuous and look for a C^1 local solution h.

Theorem 1.8 (Peano).
If f continuous, then a local C^1 solution exists.

Example 1.9.
Solution of $u'(x) = \sqrt{u}, u(0) = 0$ is not unique. For any $c > 0$,

$$f(x) = \begin{cases} 0 & x \leqslant c \\ (x-c)^2/4 & x > c \end{cases}$$

is a solution.

Definition 1.10 (Lipschitz continuous function).
Give a function f from a Banach space to $\mathbb{R}$, we call it Lipschitz, if

$$|F(x) - F(y)| \leqslant L|x - y|, \quad \forall x, y$$

where L is a constant in $\mathbb{R}$.

Remark 1.11.
$\mathrm{C}^1 \subset$ Lip but Lip is not a subset of C^1. An example is $f(x) = |x|$.

Definition 1.12 (Holder continuous function).
We call a function f is holder continuous if

$$|F(x) - F(y)| \leqslant M|x - y|^s,$$

for some constant M and $0 < s < 1$.

Theorem 1.13.
If f is locally Lipschitz, then a local solution exists and is unique.

Theorem 1.14.
Given $f : \mathbb{R} \to \mathbb{R}$ and $|f'| < 1$, f has an unique fixed point.

Theorem 1.15.
Given a Lipschitz function $f : B \to B$, with constant < 1, f has a unique fixed point.

Theorem 1.16.
Given a Lipschitz function $f : D \subset B \to D$ where D is a closed set and the Lipschitz

1.4 First order nonlinear scalar PDES II

Theorem 1.17.
Suppose f is locally Lipschitz, then the ODE

$$\begin{cases} u' = F(x,u) \\ u(0) = u_0 \end{cases}$$

has an unique local solution $u \in \mathrm{C}^1[0,T]$.

Proof.
We can rewrite the ODE as:

$$N(u)(x) = u_0 + \int_0^x F(y,u(y))\,dy$$

Let $B = \mathrm{C}[0,T]$, we will figure out what is T later.

Let $D = \{u \in C[0,T], \|u - u_0\|_C \leqslant R\}$. We will find an R such that N is a contraction.

- N maps from D to D:
  For $u \in B(u_0,R), x \in [0,T]$, we have

$$\begin{aligned} |F(x,u)| &\leqslant |F(x,u_0)| + |F(x,u) - F(x,u_0)| \\ &\leqslant |F(u_0)| + L \cdot R. \end{aligned}$$

Suppose $R \leqslant 1$, we have

$$\begin{aligned} |N(u)(x) - u_0| &\leqslant \int_0^X |F(y,u(y))|\,dy \\ &\leqslant T \cdot \Big(|F(u_0)| + LR \Big) \end{aligned}$$

Let T be small enough, we have

$$|N(u)(x) - u_0| \leqslant \frac{R}{2}.$$

- N is a contraction:
  We have

$$\begin{aligned} \|N(u) - N(v)\|_C &\leqslant \int_0^T \big|F(y,u(y)) - F(y,u(y))\big|\,dy \\ &\leqslant \int_0^T L \cdot |u(y) - v(y)|\,dy \\ &\leqslant LT \cdot \|u - v\|_C \end{aligned}$$

Hence by contraction principle, there exists a unique u for the integral equation in D. If we solve the integral equation, since RHS is continuous, we have $u \in \mathrm{C}^1[0,T]$.

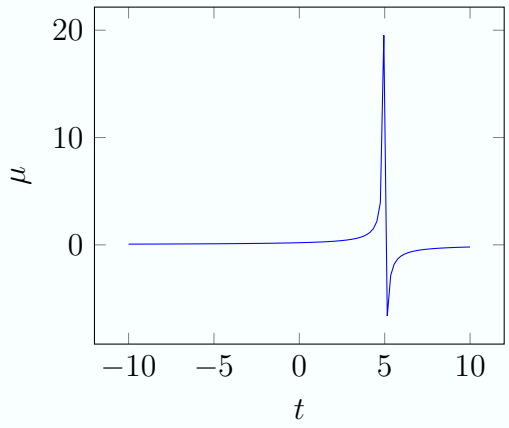
For uniqueness, we may want to ask is there any other solution which exists outside of $B(\mu_0,R)$. If such solution u_1 exists, we can consider smaller interval, where the other solution doesn't go over $\mu_0 + R$. Assume that $u_1(x) \in B(\mu_0,R)$ for $x \in [0,T_0]$. Then we can get that u_1 and u are the same in $[0,T_0]$, since $N(u)$ is also a contraction in $[0,T_0]$. Therefore, $|u_1 - u_0| \leqslant \frac{R}{2}$, which contradicts with the maximum of T_0. □

Remark 1.18.
Here we use the Bootstrap argument. We consider D first, and then prove the solution must be in D.

1.4.1 Global Solution

Now we may consider the global solutions. Before, the global solution, we may consider the maximal solution, i.e., a solution cannot be extended further.

Example 1.19.
Here we give an example for which the global solution doesn't exist.



Consider the ODE:

$$u' = u^2, u(0) = u_0 > 0.$$

The solution is $U(t) = \frac{1}{T-t}, T = \frac{1}{u_0}$. Hence it cannot be extended over $\frac{1}{u_0}$.

Suppose $u_1 : [0, T_1] \to \mathbb{R}^n$ and $u_2 : [0, T_2] \to \mathbb{R}^n$ are two solutions and $T_1 \leqslant T_2$. Choose T maximal $u_1 = u_2$ in $[0, T]$. If $T < T_1$, by local well posedness, $u_1 = u_2$ in $[T, T + \varepsilon]$ which contradict to the maximality of T.

Conclusion: As long as both two solutions exists, one must be the extension of another one. This leads to that the set of solutions is ordered. Based on Zorn's lemma, the maximal solution exists.

Proposition 1.20.
Assume u is a maximal solution on $[0, T)$, then we have

$$\lim_{t \to T} |u(t)| = \infty.$$

Proof.
If not, there exists a series $t_n \to T$ and a constant M such that

$$|u(t_n)| \leqslant M.$$

Hence all these t_n can extend ε based on the bound M and local Lipschitz condition and there exists t_i such that $t_i + \varepsilon > T$. □

1.4.2 Continuous Dependence on Data

Given ODE,

$$\begin{cases} u' = F(x, u) \\ u(0) = u_0 \end{cases}$$

u is the solution for initial data u_0. v is the solution for initial data v_0. We usually call $u : [0, T] \to \mathbb{R}^n$ the reference solution.

Theorem 1.21.

- if $|v_0 - u_0|$ is small enough, then v exists on $[0, T]$ and satisfies $|v - u|_\infty \leqslant 1$.
- If $v_0 \to u_0$, then $v \to u$ in $\mathrm{C}[0, T]$.

We will try try to track

$$\frac{d}{dt}|u - v|^2 = 2(u - v) \cdot \frac{d}{dt}(u - v) = 2(u - v)(F(u) - F(v)) \leqslant 2|u - v| \cdot L|u - v| = 2L|u - v|^2,$$

with Gronwall's inequality.

1.5 First order nonlinear Scalar PDEs III

Date: Sep 7 Given two ODEs:

$$\begin{cases} u' = F(t,u) \\ u(0) = u_0 \end{cases} \quad \text{and} \quad \begin{cases} v' = F(t,v) \\ v(0) = v_0 \end{cases}$$

we are going to show that:

Theorem 1.22.
Suppose that the solution u exists on $[0,T]$. Then there exists $\varepsilon > 0$ such that if $|v_0 - u_0| \leqslant \varepsilon$, then v exists on $[0,T]$ and $\|u - v\|_c \leqslant c|u_0 - v_0|$. In other words, the map $u_0 \to u_{[0,T]}$ is locally Lipschitz.

Proof.
We compute that

$$\frac{d}{dt}|u-v|^2 = 2(u-v)\cdot\frac{d}{dt}(u-v) = 2(u-v)(F(u)-F(v)) \leqslant 2|u-v|\cdot L|u-v| = 2L|u-v|^2.$$

Hence with Gronwall's inequality,

$$f(t) \leqslant f(0)e^{2Lt}.$$

The proof is finished except that we don't know that V exists up to full $[0,T]$. Let

$$D_1 = \{v \in \mathrm{C}[0,T], |v-u| \leqslant 1\}$$
$$D_2 = \{v \in \mathrm{C}[0,T], |v-u| \leqslant 2\}$$

Suppose we know $v \in D_2$, then v is defined on $[0,T]$and satay in a compact set, so the above argument applies.

Suppose this is not true, we use T_2 to denote the time the function exists from D_2. Then we can use Gronwall up to T_2.

$$|u(t)-v(t)|^2 \leqslant |u_0 - v_0|^2 \cdot e^{2LT_2} \leqslant \varepsilon^2 e^{2LT}.$$

When we set ε small enough, we have $|u(t)-v(t)| \leqslant 1, \forall t \in [0,T_2]$. Hence v doesn't exist D_1. Contradiction. □

Remark 1.23.
Here we also use Bootstrap assumption. The key ideal is that we use a weaker assumption to prove a stronger results. Here we use $|u-v| \leqslant 2$ to get $|u-v| \leqslant 1$.

Now we assume $F \in \mathrm{C}^1$ and u_0^h takes a one parameter family of data. When h close to 0, u_0^h is differentiable in h. From data u_0^0, we get u^0 and from u_0^h, we get u^h.

Question: How does u_h depend on h? We already know that

$$|u_0^h - u_0^0| \leqslant h \Rightarrow |u^h - u^0| \leqslant he^{2Lt}.$$

But we can give a more formal computation for $v^h = \frac{d}{dh}u^h$. We can get a equation

$$\dot{v}^h = DF(t,u^h)v^h, \quad v^h(0) = \frac{d}{dh}u_0^h.$$

This equation is called the Linearized equation. This is because:

$$\frac{d}{dt}(u^h - u^0) = F(t,u^h(t)) - F(t,u^0(t)) = DF(t,u^0(t))u^h(t) - u^0(t)) + o(u^h(t) - u^0(t))^2.$$

1.5.1 First order scalar PDE

The general form of first order scalar PDE is:

$$u : \mathbb{R}^n \to \mathbb{R} \quad F(x, u, \partial u) = 0.$$

We can classify them by degree of difficulty: linear, semilinear, quasilinear and fully nonlinear PDEs. THe initial data is given on a surface. Besides, it can also be divided into static PDEs and dynamic PDEs.

1.5.2 Linear problem

The general form of linear problem is:

$$\sum A_j(x)\partial_j u = bu + f.$$

We need curves which are tangent to $A = (A_1(x), ..., A_n(x))$ at each point. Hence we are going to solve two ODEs:

$$\begin{cases} \dot{x}(t) = A(x(t)) \\ \frac{d}{dt}u(x(t)) = \nabla u x(t) = A \cdot \nabla(t) = bu(x(t)) + f. \end{cases}$$

1.6 First order nonlinear scalar PDEs IV

Date: Sep 9, 2021

1.6.1 Local solutions for linear, scalar PDEs

As for the linear case, we are considering PDE like:

$$\underbrace{\sum A_j(x)\partial_j u}_{\text{directional derivative}} + bu = f. \tag{1.3}$$

Definition 1.24 (Initial curves or Characteristics).
The initial curves of (1.3) are the solutions to the ODE

$$\dot{x} = A(x), \quad x(0) = x_0.$$

Along the integral curves, the PDE looks like:

$$\frac{d}{dt}u(x(t)) + b(x(t))u(x(t)) = f(x(t)).$$

With chracteristics we can solve the PDE with two ODEs.

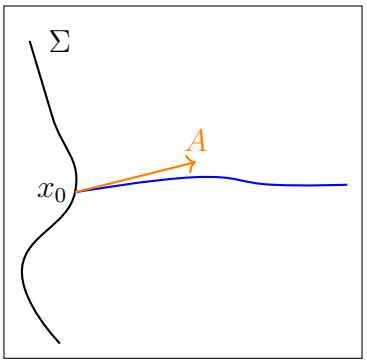
If we assume $A \in C^1$, then the chracteristics $x(t, x_0) \in C^1$. We want these chracteristics to locally foliate $\mathbb{R}^n$; in other words, we want them to cover the domain. One issue is that: What if $A(x_0) = 0$? Then $x(t) = x_0$ for all t!

Example 1.25.
Consider A that gives by:

$$\dot{x_1} = x_2, \quad \dot{x_2} = -x_1.$$

Then the integral curves will be circles, so $A(0) = 0$.

The fix for this problem is to assume that $A(x) \neq 0$ for any x.



Now suppose we have initial data $u(x) = u_0(x)$ on a curve Σ. If we swe start an x_0 on the curve of surface Σ, we can look at the integral starting from x_0. From $x_0 \in \Sigma$ and $t \in [-\varepsilon, \varepsilon]$, we can construct $x(t, x_0)$. Once we know $u(x_0)$, we can solve the second ODE to get $u(x(t, x_0))$, where $x(t, x_0) \in C^1$. By our ODE theorem, we will get $u \in C^1$.

However, we still may have bad cases.

- The integral may intersect Σ twice. But we might still get a local solution if we look at a small enough neighborhood of x_0.

Figure 1.1: Bad case I

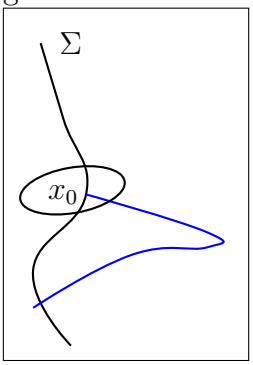
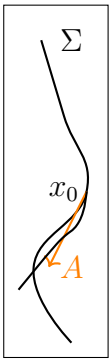


Figure 1.2: Bad case II



- A may be tangent to Σ , and re-intersection can happen arbitrarily close. Even if re-intersection is not arbitrarily close, there may be a more subtle issue with the solution not being C^1.

Here is how we avoid this issue.

Definition 1.26.
We say that Σ is noncharacteristic for out PDE (1.3) if $A\dot{N} \neq 0$ on Σ, where N is the normal to Σ.

This says that A is not tangent to Σ at any point.

Theorem 1.27 (Local inversion theorem).
Let $F : \mathbb{R}^n \to \mathbb{R}^n \in \mathrm{C}^1$. If $\det dF(x_0) \neq 0$¡ then F is a local diffeomorphism.

Theorem 1.28.
Assume $a, b, f, \Sigma, u_0 \in \mathrm{C}^1$, and suppose that Σ is noncharacteristic. Then the equation

$$\sum_j A_j \partial_j u + bu = f$$

with initial data u_0 has a unique C^1 local solution.

Proof.
We use 3 steps for this proof:

- Step 1: For $x_0 \in \Sigma$, solve for the characteristic $\Sigma \times [-\varepsilon, \varepsilon] \ni (x_0, t) \to x(x_0, t)$.
- Step 2: Solve the ODE
$$\frac{d}{dt} u(x(t)) + b(x(t))u(x(t)) = f(x(t))$$
along the characteristics to get $u(x(t, x_0))$, which is C^1 in t and x_0.
- Step 3: Show that our characteristics foliate a neighborhood of Σ. What does this mean? Looking at the map $(x_0, t) \to x(t, x_0)$, we want this to be a local diffeomorphism, i.e., a C^1 map with a C^1 inverse. Recall we have local inversion theorem. We would like to change coordinates so that Σ is a hyperplane. Since Σ is C^1, locally, Σ is the graph of a C^1 function, $x_n = f(x'), x' = (x_1, ..., x_{n-1})$ with $f \in \mathrm{C}^1$. The new coordinates are $y = (x', x_n - f(x'))$. To check that this is local diffeomorphism, we have
$$\frac{\partial y}{\partial x} = \begin{bmatrix} \frac{\partial y'}{\partial x'} & \frac{\partial y'}{\partial x_n} \\ \frac{\partial y_n}{\partial x'} & \frac{\partial y_n}{\partial x_n} \end{bmatrix} = \begin{bmatrix} I_{n-1} & 0 \\ df & 1 \end{bmatrix}$$
which has determinant 1. We can also check that the coefficient remain C^1 after changing coordinates.

In the new coordinates, $\Sigma = \{y_n = 0\}, y' = (y_1, ..., y_{n-1})$, We are looking at the equation $\dot{y} = A(y)$. Here, $y = y(t, y_0')$. Look at $\frac{\partial y}{\partial (y_0', t)}$ at $t = 0$. When $t = 0, y(y_0', 0) = (y_0', 0)$. So

$$\frac{\partial (y', y_n)}{(y_0', t)} = \begin{bmatrix} \frac{\partial y'}{\partial y_0'} & \frac{\partial y'}{\partial t} \\ \frac{\partial y_n}{\partial y_0'} & \frac{\partial y_n}{\partial t} \end{bmatrix} = \begin{bmatrix} I_{n-1} & 0 \\ A' & A_n \end{bmatrix}$$

So $\det \frac{\partial y}{\partial (y_0', t)} = A_n = \dot{y}_n \neq 0$, precisely from our noncharacteristic surface property.

□

Remark 1.29.
In the above proof, we reduced the situation to the case where Σ is a hyper plane. Let's use this to model the noncharacteristic case. Using coordinates (x, t), we can write $\Sigma = \{t = 0\}$.

Our equation looks like

$$A_t \cdot \partial_t u + A_1 \cdot \partial_1 u + \cdots + A_n \cdot \partial_n u + bu = f.$$

where $A_t \neq 0$ by the noncharacteristic assumption. So we may divide by it and just look at euqations of the form

$$\cdot \partial_t u + A_1 \cdot \partial_1 u + \cdots + A_n \cdot \partial_n u + bu = f$$

This is only a local modelling, however, not necessarily a global one.

1.6.2 Semilinear PDEs

Now we move on to solving semilinear PDEs, of the from

$$\begin{cases} \sum_j A_j(x) \partial_j u + b(u, x) = 0 \\ u|_\Sigma = 0 \end{cases}$$

The characteristic are still $\dot{x} = A(x)$, and our noncharacteristic initial surface condition is still $A \cdot N \neq 0$ on Σ. The evolution along the characteristics is

$$\frac{d}{dt} u(x(x_0, t)) = -b(u(x(x_0, t)), x(x_0, t)).$$

The difference from before is that our second equation is a nonlinear ODE, so it may have finite time blow-up. So local well-posedness is identical to the linear case, but gobal well-posedness may fail because the second ODE blows up.

Question: Why linear ODE won't blow up?

1.6.3 quasilinear PDEs

Now we look at the quasilinear problem

$$\begin{cases} \sum_j A_j(x, u) \partial_j(u) + b(x, u) = 0 \\ u|_\Sigma = u_0. \end{cases} \tag{1.4}$$

Our characteristics now look like $\dot{x} = A(x, u)$. We cannot solve this because we do not know what u is outside of Σ. The second equation would read $\dot{u} = b(x, u)$. These two ODEs would be true if we already had a solution, but we cannot solve them. What if we put these two equations together into a system?

Definition 1.30 (Characteristic system).
Given a quasilinear PDE (1.4),

$$\begin{cases} \dot{x} = A(x,u) \\ \dot{u} = b(x,u) \end{cases}$$

We call this a **characteristic system**. The initial data for the characteristic system is

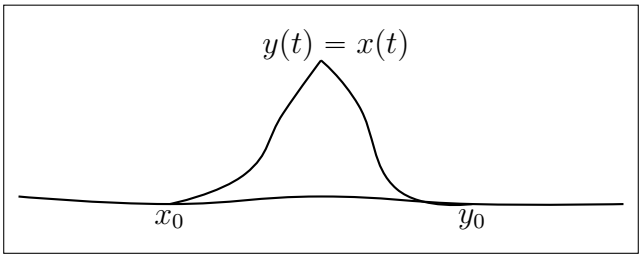
$$\begin{cases} x(0) = x_0 \in \Sigma \\ u(0) = u(x(0)) = u_0(x_0) \end{cases}$$

where the second initial condition depends on u_0. In this situation, our noncharacteristic Σ condition is

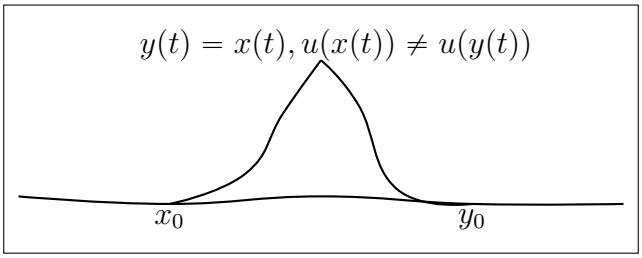
$$A(x_0, u_0(x_0)) \cdot N \neq 0.$$

Our local well-posedness theorem is identical: if Σ is noncharacteristic and $u_0 \in C^1$, then there exists a unique local C^1 solution u.

The key difference is that the characteristics may now intersect. In the semilinear case, suppose two characteristics were to intersect. Then the characteristic equation would have the same data, so the two chracteristics must be same.



In the quasilinear case, the initial data is both the location and the value of the function. Intersection means $x(t) = y(t)$, but it does not necessarily mean $u(x(t)) = u(y(t))$. So we cannot say that the two characteristics must be the same.



Next time, we will talk about what might make characteristics intersect and what to do about it.

1.7 Qusilinear and Nonlinear First Order PDEs

Date: Sep 14, 2021

1.7.1 quasilinear PDEs and conversation laws

Last time, we were looking at first order, quasilinear, scalar PDEs

$$\sum_j A_j(x,u)\partial_j u + b(x,u) = 0.$$

We saw that our characteristics have to consider both x and u. We need to solve the characteristic system

$$\begin{cases} \dot{x} = A(x,u) \\ \dot{u} = -b(x,u) \end{cases}$$

to get a local solution. Because characteristics carry information about x and u, there was no prohibition against characteristics intersecting. If our initial data is $u|_\Sigma = u_0$, then the noncharacteristic condition becomes

$$A(x_0, u_0(x_0)) \cdot N \neq 0 \quad \text{on } \Sigma.$$

Remark 1.31.
The condition of being noncharacteristic depends both on the surface and on the initial data on the surface. So the *problem* is noncharacteristic, rather than the surface (until we have fixed set of initial data).

Now since we can just consider $\Sigma = \{t = 0\}$, the model problem is

$$\begin{cases} \partial_t u + \sum_j A_j(x,u)\partial_j u + b(x,u) = 0 \\ u|_{t=0} = u_0 \end{cases}$$

Since we already using t, let's use s as the parameter along the characteristics. We have

$$\begin{cases} \dot{s} = 1 \\ \dot{x} = A(x,u) \\ \dot{u} = -b(x,u) \end{cases}$$

The first equation tells us that we can choose $s = t$. This corresponds to a dimensionality reduction of our problem.

Example 1.32 (Conservation Law).
A special case of this is what we call conservation laws:

$$u_t + \sum_j \partial_j F_j(u) = 0.$$

We can equivalently write this as

$$u_t + \sum_j F_j'(u)\partial_j u = 0.$$

Using the first form is not important for scalar equations, but it is for scalar systems because it is not always the case that we can write the second version with a divergence term.

The first version is called **density flux notation**. THis is because the u_t tells how the density of some quantity changes in time, and the flux term, $\partial_j F_j(u)$, tells you how the mass is moving with velocity $F_j'(u)$.

1.7.2 Burgers' equation

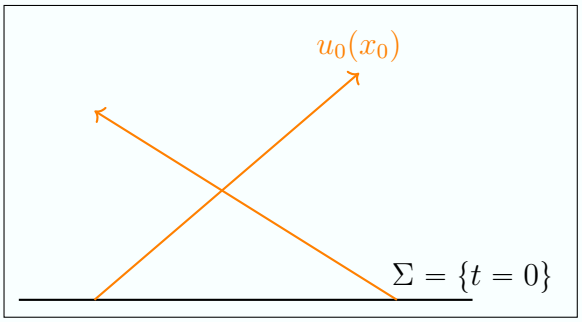
Example 1.33 (Burgers' equation).
The simplest quasilinear problem is the Burgers' equation

$$\begin{cases} u_t + uu_x = 0 \\ u|_{t=0} = u_0. \end{cases}$$

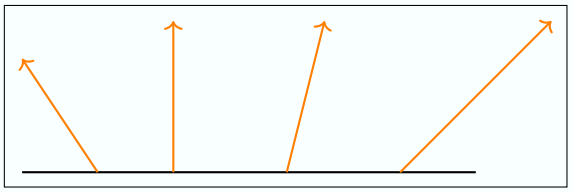
This equation seems simple, but it ends up being a model problem for more complicated equations. Here are the characteristics:

$$\begin{cases} \dot{x} = u \\ \dot{u} = 0. \end{cases}$$

Thus the characteristics are $x(t) = x_0 + tu_0(x_0)$. Hence, the characteristics may intersect as follows:



How would we choose our dat aso the lines don't intersect? If u_0 is increasing , the picture looks like this:



So we get a global solution forward in time, but we don't get a global solutions backward in time. So the only global solutions are constant.

Remark 1.34.
In physics, we expect there to be *causality*. That is, we expect the future to be determined by the past but not the past to be determined by the future. Later we will see what we will do after the point where characteristics intersect.

Example 1.35.
Let' give an equation for u_x:

$$u_{tx} + uu_{xx} + u_x^2 = 0.$$

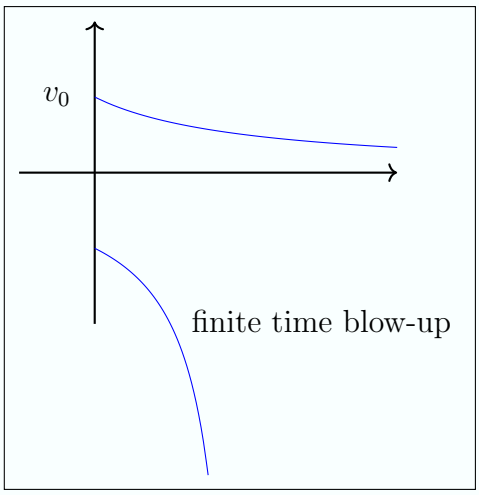
If we write $u_x = v$, then this equation is just talking about the derivative along the characteristics:

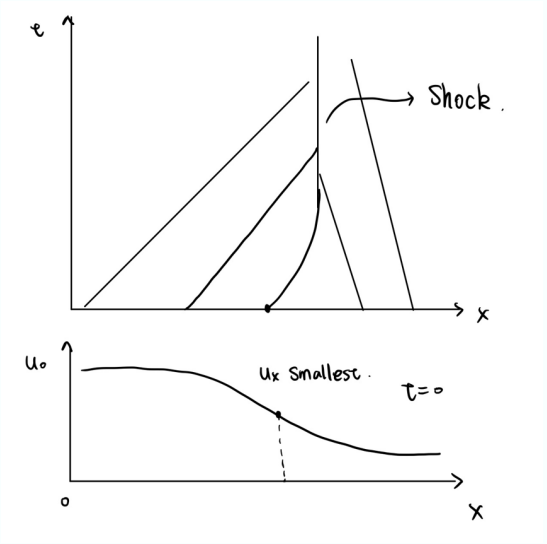
$$(\partial_t + u\partial_x)v + v^2 = 0.$$

We may also write this as

$$\dot{v} + v^2 = 0,$$

where the dot is the derivative along the characteristic. This equation tells us how the slope of the solution is evolving. If $v_0 > 0$, the slope decreases toward 0. However, if $v_0 < 0$, we get finite time blow-up.





The smallest slope means the fastest blow-up. Suppose the initial data u_0 is decreasing, so we will get intersections of characteristics.Suppose the initial data u_0 is decreasing, so we will get intersections of characteristics. Because things intersect, there is no unique way to continue the equation. Here, we have a **shock**, or a jump discountinuity. We will see later how to find what the equation fo the shock curve looks like.

Conservation laws is still a very active area, with a number of hard problems.

1.7.3 Fully nonlinear problems

We now look at PDEs of the form

$$\begin{cases} F(x, u, \partial u) = 0 \\ u|_\Sigma = u_0, \end{cases}$$

where the dependence on ∂u is nonlinear. Where do we start? Before, we had a vector field that let us interpret the equation using a directional derivative.

Let's look at the linearized equation: Suppose we have not just a solution but a 1- parameter family of solutions u^h to our problem with solution v^h to our linearized given by

$$\frac{d}{dh} u^h = v^h.$$

Differentiate the equation with respect to h to get the linearized equation:

$$0 = \frac{\partial}{\partial h} F\left(x, u^h, \partial u^h\right) = F_u \cdot v^h + \sum_j F_{p_j} \cdot \partial_j v^h,$$

where we write $F = F(x, u, p)$ and $p = (p_1, \ldots, p_n)$. This linearized equation is linear transport equation. So we get a vector field $A_j = F_{p_j}(x, u, \partial u)$. We should try to use this vector field to find characteristics. Our equation looks like

$$\begin{cases} \dot{x}_j = F_{p_j}(x, u, \partial u) \\ \dot{u} = \ldots \end{cases}$$

The first equation depends on ∂u, so we may try to add an equation $\dot{\partial} u = \ldots$. But then we would get $\partial^2 u$ in this equation, and we would be in the same situation. How do we get past this issue?

This is trivial since

$$\dot{u} = \sum_j \partial_j u \cdot \dot{x}_j = \sum_j F_{p_j}(x, u, \partial u) \cdot \partial_j u.$$

We also have

$$\begin{aligned} 0 &= \partial_{x_j} F(x,u,\partial u) \\ &= F_{x_j}(x,u,\partial u) + F_u(x,u,\partial u)\partial_j u + \underbrace{\sum_k F_{p_k}(x,u,\partial u)\partial_k \partial_j u}_{=\sum_k \dot{x}_k \partial_k \partial_j u = \dot{\partial_j u}} \end{aligned}$$

Therefore, the characteristic equations are:

$$\begin{cases} \dot{x}_j = F_{p_j}(x,u,\partial u) \\ \dot{u} = \sum_j F_{p_j}(x,u,\partial u)\cdot \partial_j u \\ \dot{\partial_j u} = -F_{x_j}(x,u,\partial u) - F_u(x,u,\partial u)\cdot \partial_j u. \end{cases}$$

Or we can write it as:

$$\begin{cases} (a)\, \dot{x}_j(t) = F_{p_j}(x,z,p) \\ (b)\, \dot{z}(t) = \sum_j F_{p_j}(x,z,p)\cdot p^j u \\ (c)\, \dot{p}_j(t) = -F_{x_j}(x,z,p) - F_z(x,z,p)\cdot p_j \end{cases} \tag{1.5}$$

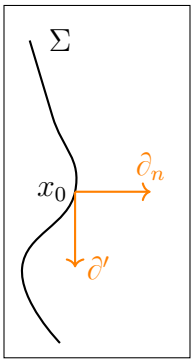
In fact, we have proved that

Theorem 1.36 (Structure of chracteristic ODE).
Let $u \in C^2(U)$ solve the nonlinear, first order PDE in U. Assume $x(\cdot)$ solves the 1.(a), where $p(\cdot) = Du(x(\cdot)), z(\cdot) = u(x(\cdot))$. Then $p(\cdot)$ solves the 1.(c) and $z(\cdot)$ solves the 1.(b).

Now we may ask what is the initial data for this system (1.5)? We had $x(0) = x_0$ and $u(0) = u_0$ before, but now we have

$$\begin{cases} x(0) = x_0 \\ u(0) = u_0 \\ \partial u(0) = ? \end{cases}$$

We need the information of all the derivatives of u at x_0. In particular, we need both $n-1$ tangential derivatives to Σ and 1 normal partial derivative to Σ. If we frame this in the tangnet space, we wan t the tangent derivative $\partial' = (\partial_1, \partial_2, \ldots, \partial_{n-1})$ and the normal derivative ∂_n. We know ∂', but what about ∂_n?



We know that

$$F(x_0, u_0, \partial' u_0, \partial_n u) = 0,$$

so we would like to solve for $\partial_n u$. This tells us that

$$\partial_n u = G(x_0, u_0, \partial' u_0),$$

for some function G. We can do this if

$$F_{p_n}(x_0, u_0, \partial' u_0, p_n) \neq 0.$$

If we didnt put our equation in this special frame, this condition reads as

$$F_p(x_0, u_0, p)\cdot N \neq 0,$$

the condition that the equation is noncharacteristic.

Remark 1.37.
What if this question has more than 1 solution? We may not get uniqueness; the answer may depend on our choice here of initial data.

1.8 Existence of Solutions to Nonlinear First Order Scalar PDEs

Date: Sep 14, 2021

1.8.1 Existence and uniqueness given initial data

Last time, we were looking at fully nonlinear equations

$$\begin{cases} F(x, u, \partial u) = 0 \\ u = u_0 \text{ on } \Sigma. \end{cases}$$

If u solves this equation and x solve the 1.(a) ODE, then $(x, u, \partial_j u)$ solves the characteristic system

$$\begin{cases} \dot{x} = F_p(x, z, p) \\ \dot{z} = F_p(x, z, p) \cdot p \\ \dot{p} = -F_x(x, z, p) - F_z(x, z, p) \cdot p \end{cases}$$

The initial data for the characteristic system on Σ is

$$\begin{cases} x(0) = x_0 \\ z(0) = u_0(x_0) \\ p(0) = p_0 \end{cases}$$

where p_0 has a tangential component $\partial_\tau u_0$ and a normal component given by solving $F(x_0, u_0, p_0)$. In this last part, we had a local solvability condition $F_p \cdot N \neq 0$, where N is the normal to Σ. This is the same as the noncharacteristic condition. Our objective is to turn this into an existence proof.

Theorem 1.38.
Assume that F is of clsss C^2, Σ is C^2, $\mu_0 \in C^2$, and the problem is noncharacteristic, i.e., there exists p_0 on Σ such that $F_{p_0} \cdot N \neq 0, F(x_0, u_0, p_0) = 0$, and $(p_0)_\tau = \partial_\tau u_0$. Then there exists a unique local solution $u \in C^2$ near Σ such that $u|_\Sigma = u_0$ and $\partial u|_\Sigma = p_0$.

Proof.
Firstly, an outline is that:

- Step 1: Solve the characteristic system with initial data (x_0, u_0, p_0) on Σ. This gives us

$$(x(s, x_0), u(s, x_0), p(s, x_0)),$$

  which we can solve by using ODE theory.

Step 2: Show that the map

$$\Sigma \times [-\varepsilon, \varepsilon] \ni (x_0, s) \mapsto x(x_0, s) \in \mathbb{R}^n$$

is a local diffeomorphism with inverse

$$x \to (x_0, s).$$

Step 3: Define

$$u(x(s, x_0)) = (s, x_0).$$

This is true if a solution u exists.

The main difficulty is that at the end of our construction, we get the function

$$z(s, x_0) = u(x), \quad x = x(s, x_0), \quad p_j(s, x_0) \stackrel{?}{=} \partial_j z(x).$$

Our final goal is to prove that $p_j(s, x_0) = \partial_j z(s, x_0)$. By construction of our initial data ,we know this is true at $s = 0$. Ideally, we might want to show $\frac{\partial}{\partial s}(p_j - \partial_j z) = 0$. Instead, we will have a weaker version that works:

$$\frac{\partial}{\partial s}\left(p_j - \partial_j z\right) = \text{coeff} \cdot \left(p_j - \partial_j z\right)$$

which is a linear ODE for $p_j - \partial_j z$.

Our preliminary step is to show that $F(x, z, p) = 0$. This is true on Σ, i.e., when $s = 0$. Compute

$$\frac{d}{ds}F(x, z, p) = F_x \cdot \dot{x} + F_z \cdot \dot{z} + F_p \cdot \dot{p} = 0.$$

Next, compute $\frac{\partial}{\partial s}(p_j - \partial_j z)$. We have

$$\frac{\partial p_j}{\partial s} = -F_{x_j} - F_z \cdot p_j,$$

but to calculate $\frac{\partial}{\partial s}\partial_j z$, we need to use $\dot{z} = F_p \cdot p$. $\frac{\partial}{\partial s} = \sum_k F_{p_k} \cdot \frac{\partial}{\partial x_k}$, where F_{p_k} has variable coefficients. So the derivatives do onot commute. We can explicitly compute

$$\frac{d}{ds}F(x, z, p) = \sum_k \sum_k F_{p_k}\partial_k \partial_j z.$$

Since

$$\partial_j \dot{z} = \partial_j \left(\sum_k F_{p_k}\partial_k z\right) = \sum_k F_{p_k}\partial_j \partial_k z + \sum_k \partial_j F_{p_k}, \cdot \partial_k z$$

which gives

$$\frac{\partial}{\partial s}\partial_j z = \partial_j \dot{z} - \sum_k \partial_j F_{p_k} \cdot \partial_k z.$$

So we get

$$\begin{aligned}
\frac{\partial}{\partial s}\left(p_j - \partial_j z\right) &= -F_{x_j} - F_z \cdot p_j - \partial_j \dot{z} + \sum_k \partial_j \left(F_{p_k}\right) \cdot \partial_k z \\
&= -F_{x_j} - F_z \cdot p_j - \sum_k \partial_j \left(F_{p_k} \cdot p_k\right) + \sum_k \partial_j \left(F_{p_k}\right) \partial_k z \\
&= -F_{x_j} - F_z \cdot p_j - \sum_k F_{p_k}\partial_j p_k \\
&\underbrace{- \sum_k p_k \left(F_{x_j p_k} + F_{z p_k}\partial_j z + F_{p_\ell p_k}\partial_j p_\ell\right) + \sum_k \partial_k z (F_{x_j p_k} + F_{z p_k}\partial_j z + F_{p_\ell p_k}\partial_j p_\ell)}_{\sum_k -(p_k - \partial_k z)\cdot \partial_j F_{p_k}} \\
&= -F_{x_j} - F_z \cdot p_j - \sum_k F_{p_k}\partial_j p_k + \sum_k -\left(p_k - \partial_k z\right) \cdot \partial_j F_{p_k}
\end{aligned}$$

We also have

$$F_{x_j} + F_z \cdot \partial_j z + \sum_k F_{p_k}\partial_j p_k = 0$$

by taking $\frac{\partial}{\partial x_j}$ of our earlier computation. This last term $F_{p_k} \cdot \partial_j p_k$ is the same term in the above expression. If we substitute, we get

$$\frac{\partial}{\partial_s}\left(p_j - \partial_j z\right) = -F_z \cdot \left(p_j - \partial_j z\right) - \sum_k \partial_j F_{p_k} \cdot \left(p_k - \partial_k z\right),$$

which is a linear system. Therefore, z is the solution to our equation, and we are done. □ □

1.8.2 Problems in standard form

Example 1.39.
Begin with the equation

$$u_t + F(t, x, u, \partial u) = 0$$

We will label u_t as τ, u as z and ∂u as p. So we get the equation:

$$\widetilde{F}(t, x, z, \tau, p) = \tau + F(t, x, z, p) = 0,$$

and the system

$$\begin{cases} \dot{t} = 1 (\text{so } s = t) \\ \dot{x} = F_p \\ \dot{z} = F_p \cdot p + \tau = F_p \cdot p - F \\ \dot{p} = -F_x - -F_z \cdot p \\ \dot{\tau} = -F_t - F_z \cdot \tau \end{cases}$$

In the midlle 3 euqations, we have no τ terms, so we can discard the last equation. Another way to think of this is that $\widetilde{F} = 0$, so τ is already given as $-F$. So we get smaller system

$$\begin{cases} \dot{x} = F_p \\ \dot{z} = F_p \cdot p - F \\ \dot{p} = -F_x - F_z \cdot p \end{cases}$$

The price we pay is the extra F term in the second equation, compared to before.

Remark 1.40.
Solutions are local, near Σ, until characteristics may intersect. There is no way to continue solutions in general past this intersection of characteristics. For specific classes of problems, however, there is hope.

Example 1.41 (Hamiton flow).
Suppose we have an equation

$$H(x, \partial u) + u_t = 0$$

which does not depend directly on u. Then we get

$$\begin{cases} \dot{x} = H_p \\ \dot{p} = -H_x \\ \dot{z} = H_p \cdot p - H \end{cases}$$

The first two equations do not depend on z, so we can discard the last equation, solve the first two questions first, and integrate the last equation at the end.

This type of system is called a **Hamilton Flow**. Hamilton flows play a role in symplectic geometry. Many PDEs can be interpreted as Hamiltonian flows. The Hamilton-Jacobi euqations are of the form

$$u_t + H(x, \partial u) = 0.$$

Next time, we will do a bit of variational calculus to not only solve Hamilton-Jacobi equations but to also see how we may extend a solution past a point where characteristics intersect. In a Hamilton flow, the characteristics only depend on (x, p). When characteristics intersect, they may have the same x but different $p = \partial u$. We will try to continue the solution in a way such that ∂u has a jump discontinuity.

Example 1.42.
Consider the equation

$$\begin{cases} u_t + \frac{1}{2}|\partial_x u|^2 = 0 \\ u(0) = u_0. \end{cases}$$

Here, $H(p) = \frac{1}{2}p^2$, and we get the system

$$\begin{cases} \dot{x} = p \\ \dot{p} = 0 \end{cases}$$

Here, the characteristics are straight lines, with $p(0) = \partial_x u_0$.

Example 1.43 (Eikonal equation).
The equation

$$|u_t|^2 - |\partial_x u|^2 = 0.$$

is not int the form we have talked about already. This gives

$$u_t = \pm|\partial_x u|,$$

so we will get 2 solutions.

1.9 Hamilton-Jacobi Equations and Calculus of Variations

Date: Sep 21, 2021

1.9.1 Recap: Hamilton-Jacobi equations

Last time, we started talking about Hamilton-Jacobi equations, as an example of first order PDEs:

$$\begin{cases} u_t + H(x, Du) = 0 \\ u(0) = u_0 \end{cases}$$

The characteristics for this system were given by

$$\begin{cases} \dot{x} = H_p(x, p) \\ \dot{p} = -H_x(x, p) \\ \dot{z} = H_p(x, p) \cdot p - H(x, p) \end{cases}$$

with initial data

$$\begin{cases} x(0) = x_0 \\ z(0) = u_0 \\ p(0) = \partial_x u_0. \end{cases}$$

The equation for $\dot{u}$ and $\dot{p}$ are called the Hamilton equations. We noticed that we only need to solve them first to get the characteristics, and then we can integrate the $\dot{z}$ equation to solve it after the fact.

1.9.2 Calculus of variations

Today, we will be looking at the calculus of variations. Here is the setup: We have a function $L(x, q)$ we call the **Lagrangian**, and to each function $x : [0, T] \to \mathbb{R}$, we associate to this function an **action functional**

$$\mathcal{L}(x) = \int_0^T L(x, \dot{x})\, dt.$$

The question we want to ask is: what are the minimizers of $\mathcal{L}$? We are looking for

$$\min_{x:[0,T]\to\mathbb{R}} \mathcal{L}(x).$$

We can think of $\mathcal{L}$ giving the cost of the trajectory x. So we want to find the most efficient trajectory x.

If we were just minimizing a function in $\mathbb{R}^n$, we would look for critical points. In particular, for $f : \mathbb{R}^n \to \mathbb{R}$, a minimum point in a critical point if $\nabla f = 0$. How do we do this in the case of our functional? We can talk in terms of directional derivatives. Replace x by $x + hy$ and look at the map $h \to \mathcal{L}(x + hy)$, where $h = 0$ is minimum point. Assume that our perturbation y is compactly supported. In this case $h = 0$, we have

$$\begin{aligned} 0 &= \frac{d}{dh}\mathcal{L}(x + hy) \\ &= \frac{d}{dh}\int_0^T L(x + hy, \dot{x} + h\dot{y})\, dt \\ &= \int_0^T L_x(x, \dot{x}) \cdot y + L_q(x, \dot{x}) \cdot \dot{y}\, dt \end{aligned}$$

where we are using q as a placeholder for the second variable, as we did with p before. This holds for all $y \in \mathrm{C}_0^\infty([0, T])$. To deal with $\dot{y}$ term, we integrate by parts (using the compact support assumptions):

$$= \int_0^T y(L_x(x, \dot{x}) - \frac{d}{dt}L_q(x, \dot{x}))\, dt$$

when integrated against any function with compact support, the part inside the parentheses gives 0. So it must equal 0, thus, we have actually proven a theorem:

Theorem 1.44 (Euler-Lagrange equation).
x is a critical point for $\mathcal{L}$ if and only if it solves

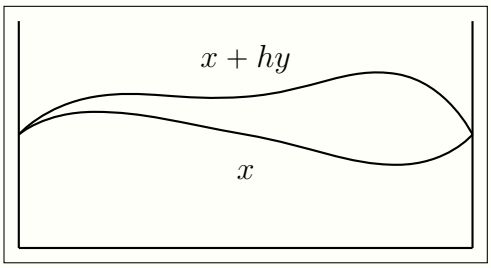
$$L_x(x, \dot{x}) - \frac{d}{dt} L_q(x, \dot{x}) = 0.$$

Remark 1.45.

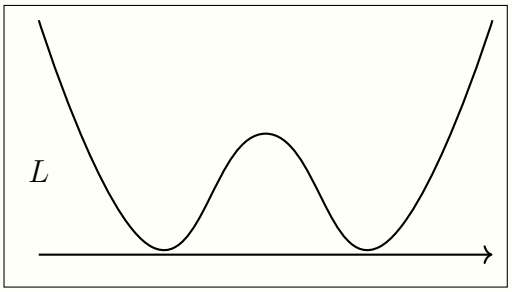
- The PDE analogue takes a function $u : \mathbb{R}^n \to \mathbb{R}$ and gives the Euler-Lagrange equation

  $$L_x(u, \partial u) - \partial_j L_{q_j}(u, \partial u) = 0,$$

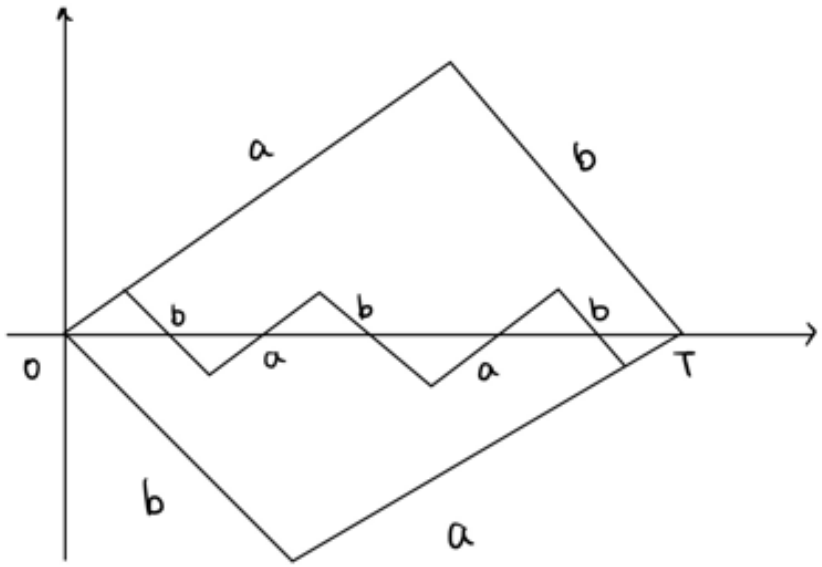
  which is a second order PDE.
- Our perturbation does not change the values at the endpoints $x(0), x(T)$, so it gives critical points in a context where $x(0)$ and $x(T)$ are fixed.



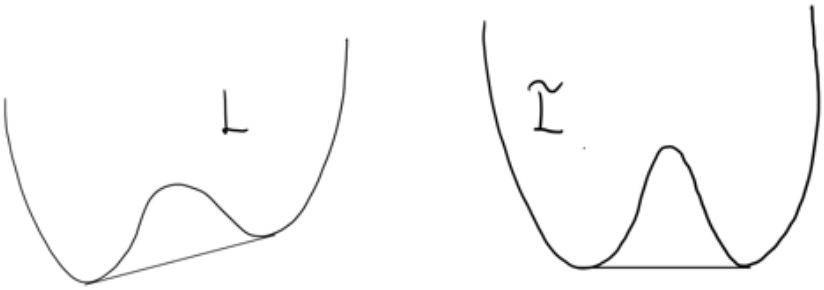
- Suppose $L = L(\dot{x})$ is the following "double well potential".



Suppose also that $x(0) = X(T)$. We want to minimize $\int_0^T L(\dot{x})\, dt \geqslant 0$. Can we achieve 0? We can make a line with slope a and then a line with slope b to get 0 as the minimum (notice that this is not differentiable!), Alternatively, we can alternate between lines of slope a and b in any number of ways as follows:



So we get the infimum is 0, and the minimum is 0 if we allow for any Lipschitz function x. In fact, all trajectory with slopes between $[a, b]$ are limiting minimizers. This means we are actually dealing with an **effective Lagrangian** L_{eff} with the hump between a and b flattened out. The effective Lagrangian L_{eff} is the convex envelope of L.
If we had another Lagrangian like the following, could we again look at the convex envelope?



Suppose we add a linear constant to get $\widetilde{L}(q) = L + c \cdot q$ and we can get $\widetilde{L}$ as this. So the effective Lagrangian must be convex as a function of q. For PDEs, convexity is no longer required. Instead, we require **rank one convexity**, which is given by convexity in one variable at a time.

Example 1.46.
Here is an example that comes from classical mechanics. Suppose we have a particle with trajectory $x(t)$ moving in a conservative force field $F = \nabla\phi$, where ϕ is the potential. Then we have the Lagrangian

$$L(x, q) = \underbrace{\frac{1}{2}mq^2}_{\text{kinetic energy}} - \underbrace{\phi(x)}_{\text{potential energy}},$$

where we have $\phi(x) = \frac{d}{dt}(m\dot{x})$ (This comes from the minimizer condition), which we can write as $m \cdot \ddot{x} = F(x)$, which is Newton's law.

1.9.3 Connecting the Hamilton-Jacobi equations to the Euler Lagrange equations

Returning to Hamilton-Jacobi equations, we have x, p with the function H, and we want to relate this to the $x, q = \dot{x}$ and L in the Euler-Lagrange equation. We can think of the Euler-Lagrange equation as a system for x and q via

$$\begin{cases} \dot{x} = q \\ \frac{d}{dt} L_q(x, q) = L_x. \end{cases}$$

We want to let $p = L_q(x, q)$. For this to make sense, we need $q \to L_q(x, q)$ to be a diffeomorphism from $\mathbb{R}^n \to \mathbb{R}^n$ for fixed x.

Proposition 1.47.
If $L(\cdot) := L(x, \cdot) : \mathbb{R}^n \to \mathbb{R}$ is strictly convex and coercive ($\lim_{q\to\infty} \frac{L(x,q)}{|q|} = \infty$), then $p \to L_q(x, p)$ is a diffeomorphism.

Proof.
Injectivity: L is strictly convex, so the graph of L is above its tangent lines at points of nonintersection:

$$L(y) > L(x) + (y - x)DL(x), \quad y \neq x.$$

We can use this to write

$$(y - x)(DL(y) - DL(x)) > 0, \quad y \neq x.$$

This gives injectivity. Surjectivity: We want to minimize $L(x, q) - p \cdot q$. If a minimum exists, then the gradient must equal 0:

$$L_q(x, q) = p,$$

which is our surjectivity. Why mut the minimum exist? This is because $\lim_{q\to\infty} L(x, q) - p \cdot q = \infty$ by coercivity. To check that this is a local diffeomorphism, the differential of $q \to L_q(x, q)$ is $L_{qq} \geqslant 0$. In fact, by strict convexity, this is > 0. □ □

So we have $p = L_q(x, q)$, we will define $H(x, p) = max_q p \cdot q - L(x, q)$. Note that this is the same quantity we dealt with in the above proof. The functions $p \cdot q - L(x, q)$ are linear in p, so this maximum is convex.

Proposition 1.48.
H is convex and coercive.

Proof.
This comes from the strict convexity and coercivity of L. □ □

Proposition 1.49.

$$q = H_p(x, p).$$

Proof.
Since H is defined by maximum, we have $H(x, p) + L(x, q) - pq \leqslant 0$ and this holds with equality if $p = L_q(x, q)$. Now we can fix q and vary p! Then p is a maximum point for this expression when the derivative $H_p(p) - q = 0$. □ □

Now let's change our variables: The Euler-Lagrange equations say

$$L_x(x,q) - \frac{d}{dt}\underbrace{L_q(x,q)}_{p} = 0$$

So we get

$$\begin{cases} \dot{p} = L_x(x,q) \stackrel{?}{=} -H_x(x,p) \\ \dot{x} = q = H_p(x,p) \end{cases}$$

We also have

$$H(x,p) + L(x,q) - pq \leqslant 0.$$

If we think of $p = p(x,q)$, we can take $\frac{d}{dx}$ to get

$$\begin{aligned} &H_x(x,p) + H_p(x,p)\frac{\partial p}{\partial x} + L_x(x,q) - q \cdot \frac{\partial p}{\partial x} \\ =&H_x(x,p) + +L_x(x,q) + \underbrace{(H_p(x,p) - q)}_{=0} \cdot \frac{\partial p}{\partial x} = 0 \end{aligned}$$

And this tells us that

$$L_x(x,q) = -H_x(x,p).$$

1.10 Hamilton-Jacobi Equations & Calculus of Variation II

Date: Sep 23, 2021

1.10.1 Recap: Connecting H-J equation to calculus of variation via Legendre transform

Last time, we wanted to compare Hamilton-Jacobi equations to calculus of variations. The Hamilton-Jacobi equations are of the form

$$\begin{cases} u_t + H(x, \partial u) = 0 & \text{in } \mathbb{R} \times \mathbb{R}^n \\ u(0) = u_0 & \text{in } \mathbb{R}. \end{cases}$$

The characteristics given to this equation are

$$\begin{cases} \dot{x} = H_p \\ \dot{p} = -H_x \\ \dot{z} = p \cdot H_p - H \end{cases}$$

with initial data

$$\begin{cases} x(0) = x_0 \\ z(0) = u_0 \\ p(0) = \partial_x u_0. \end{cases}$$

The first two equations of characteristic system are called the **Hamilton flow**. In calculus of variations, we have a Lagrangian $L : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, and we want to minimize an action functional

$$\min_{x \in \mathcal{A}} \mathcal{L}(x) = \min_{x \in \mathcal{A}} \int_0^T L(x, \dot{x})\, dt,$$

where $\mathcal{A} = \{x : [0, T] \to \mathbb{R} \mid x(0) = x_0, x(T) = x_T, x \text{ Lipschitz}\}$. Minimizers satisfy the **Euler-Lagrange equation**

$$L_x(x, \dot{x}) - \frac{d}{dt} L_q(x, \dot{x}) = 0.$$

Last time, we connected these two setups. We saw that L is strictly convex and coercive if and only if H is strictly convex and coercive. And the relation between H and L is

$$H(x, p) = \max_{q \in \mathbb{R}^n} -L(x, q) + p \cdot q,$$

which is maximized at $p = L_q(x, q)$. This relation gives that

$$H(x, p) + L(x, q) \geqslant p \cdot q$$

with equality when $p = L_q(x, q)$. This expression is symmetric in p and q, so it allows us to cast q in terms of $p : q = H_p(x, p)$. This relationship is known as the **Legendre tranform**.

Remark 1.50.
The Legendre transform is well-defined and is an involution, only assuming convexity.

Example 1.51.
If we remove strict convexity and coercivity, we can get functions which are not defined everywhere. For example, take

$$\begin{cases} L(0) = 0 \\ L(q) = -\infty \quad q \neq 0 \end{cases}$$

What is H in this case?

We have not incorporated the initial data of the Hamilton-Jacobi equations into our calculus of variations. We will do this by adding $u_0(x_0)$ to the minimization problem (so when $T = 0$, we get $u_0(x_0)$ and removing the condition $x(0) = x_0$ from our set $\mathcal{A}$. So we are minimizing

$$\min_{x\in\mathcal{A}} \int_0^T L(x, \dot{x})dt + u_0(x_0) = u(T, x_T), \tag{1.6}$$

with $\mathcal{A} = \{x : [0, T] \to \mathbb{R} \text{ Lipschitz} | x(T) = x_T\}$.

1.10.2 Existence of minimizers for the Euler-Lagrange equation

We want to prove:

Theorem 1.52.
The minimal value function $u(T, x_T)$ in teh calculus of variations is the solution to the H-J equations

First, we should ask: Does a minimum solution to the Euler-Lagrange equation exist? The answer is yes, as long as L is convex, coercive, and Lipschitz in x and if $u_0 \in$ Lip. However, there is no guarantee of uniqueness. We will not prove this, but here is some intuition:

Here is the trivial case:

Proposition 1.53.
SUppose we have a continuous function $F : K \to \mathbb{R}$ with K compact. Then $\min F$ is attained.

Proof.
Just consider the campactness of K. □ □

What if we try to apply this to calculus of variations? Suppose we have a minimizing sequence $x_n : [0, T] \to \mathbb{R}^n$. Then $\mathcal{L}(x_n) \to u(T, x_T)$ but in what topology? Is x_n in a bounded set? We know that $\mathcal{L}(x_n)$ is bounded. If$L(x, q) = q^2$, for example, we could conclude that $\int_0^T (\dot{x}_n)^2 \leqslant c$. Then $\dot{x}_n$ is bounded in $L^2([0, T])$. This would imply that x_n is bounded in $C^{1/2}$ using Holder's inequality: $(|x_n(t)-x_n(s)| \leqslant c|t-s|^{1/2})$. This implies that x_n is equicontinuous (and equibounded by the $x(T) = x_T$ assumption). So the Arzela-Ascoli theorem says that $x_n \to x$ uniformly (x_n is a subsequence). Then

$$\lim_{n\to\infty} \mathcal{L}(x_n) = \lim_{n\to\infty} \int_0^T L(x_n, \dot{x}_n)dt + u(x_n, 0)$$

We can pas to the limit without a problem for x, but the convergence with respect to $\dot{x}$ is trouble. The limit of the intergral may not exist, but maybe we can hope for

$$\int_0^T L(x, \dot{x})dt \leqslant \liminf_{n\to\infty} \int_0^T L(x_n, \dot{x}_n)\, dt$$

This is lower seminicountinuity for the map $x \to \mathcal{L}(x)$.The key observation is that convexity of $\mathcal{L}$ implies lower semicontinuity of $\mathcal{L}$.

Proposition 1.54.
The minimum solution to the Euler-Lagrange equation (1.6) exists.

Proof.
The convexity inequality tells us that

$$L(\dot{x}_n) \geqslant L(\dot{x}) + L_q(\dot{x})(\dot{x}_n - \dot{x}).$$

Integrating gives us

$$\int_0^T L(\dot{x}_n)\,dt \geqslant \int_0^T L(\dot{x})dt + \int_0^T L_q(\dot{x})(\dot{x}-\dot{x}_n)dt.$$

We are done if $\lim_{n\to\infty} \int L_q(\dot{x})(\dot{x}_n - \dot{x}) = 0$. We ahve replaced our nonlinear dependence on $\dot{x}_n - \dot{x}$ by a linear property. Since $\dot{x} \in L^2$, we can approxiamte $L_q(\dot{x})$ by smooth functions. Suppose $y_k \in \mathrm{C}^\infty$ with $y_k \to L(\dot{x})$ in L^2. It's enough to see that

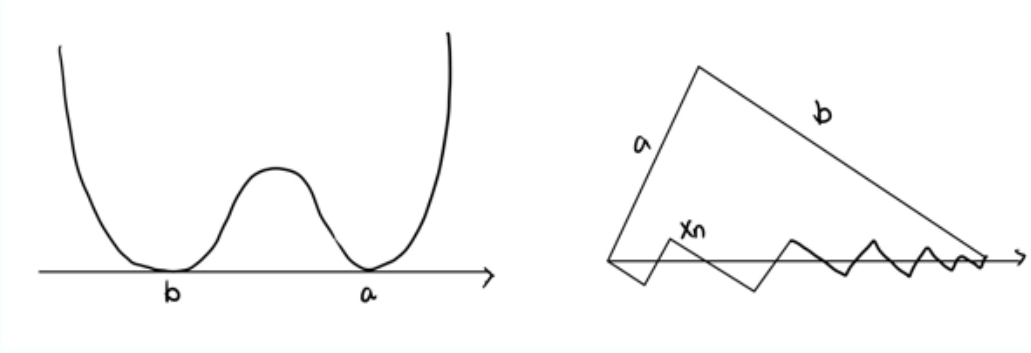
$$\lim_{n\to\infty} \int_0^T y_k\,(\dot{x}_n - \dot{x})\,dt = 0$$

In this context, we can integrate by parts. The integral equals

$$\int_0^T y_k\,(\dot{x}_n - x)\,dt = \int_0^T \dot{y}_k\,(x_n - x)\,dt + y_k\,(x_n - x)|_0^T \overset{n\to\infty}{\longrightarrow} 0$$

by uniform convergence of $x_n \to x$. □ □

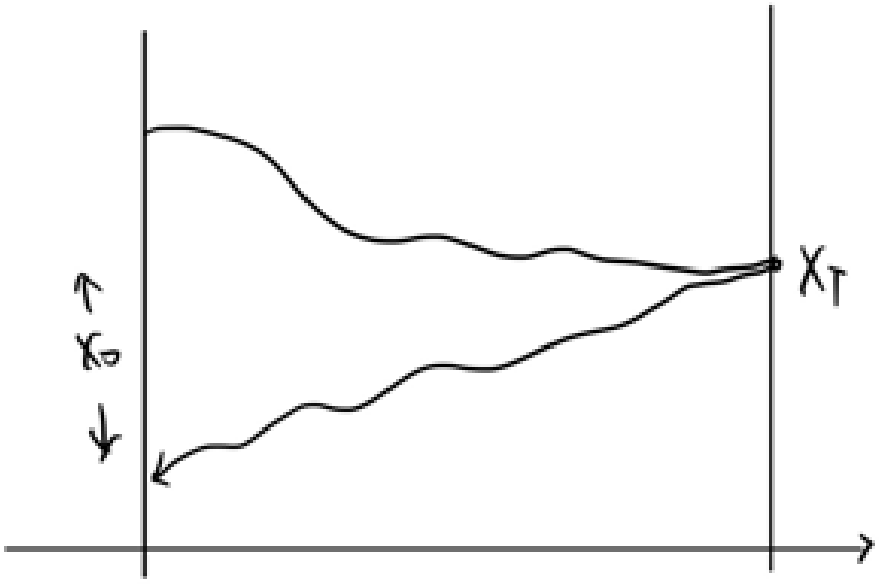
Example 1.55.
Recall our double well potential



In this case, if x_n is a wiggle approximating the 0 trajectory, we have $\mathcal{L}(\dot{x}_n) = 0$ by $\mathcal{L}(\dot{(x)}) = \mathcal{L}(0) > 0$.

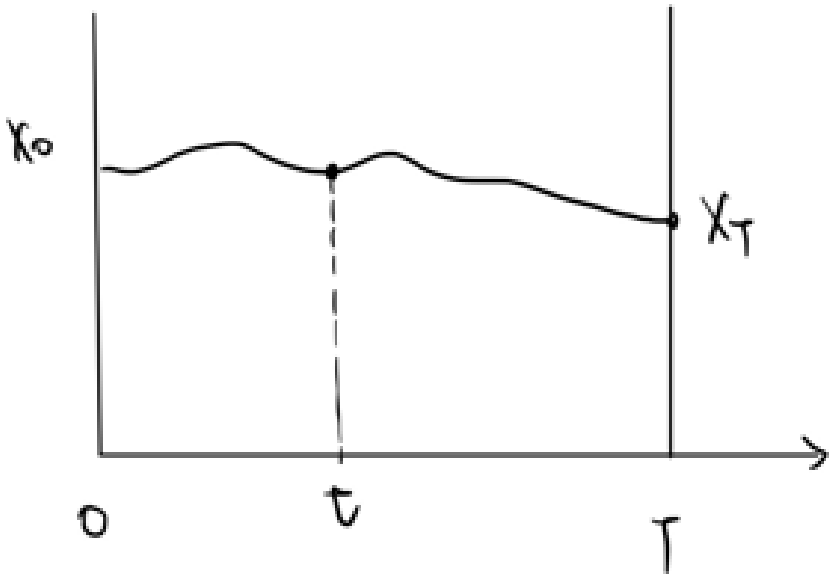
Remark 1.56.
The Hamilton-Jacobi equation can be solved for a short time using characteristics. In calculus of variations, the analogue turns out to be that minimizers are unique for a short time.

We want to think of two minimizers in calculus of variations as characteristics that interect.



1.10.3 E-L equation minimizers slove H-J equations

Proof.
Suppose x is a minimizer for the action functional. We can choose a intermediate point t, and first minimize relative to the time t.



$$\min_x \int_0^T L(x,\dot{x})dt + u_0(x_0) = \min_x \int_0^t L(x,\dot{x})ds + u_0(x_0) + \int_t^T L(x,\dot{x})ds$$

If $x|_{[0,T]}$ is a minimizer, then $x|_{[0,t]}$ is also a minimizer. So

$$u(x_T, x_0) = \min u(x_t, x_0) + \int_t^T L(x,\dot{x})ds$$

This is called the **Dynamic programming principle**. This principle tells us that for minimizers,

$$u\left(x_T, x_0\right) = u\left(x_t, x_0\right) + \int_t^T L(x, \dot{x}) ds$$

which we can differentiate with respect to t to get

$$\begin{aligned} \frac{d}{dt} u\left(x_t, x_0\right) &= L(x, \dot{x}) \\ &= p \cdot q - H(x, p) \\ &= p \cdot H_p - H \end{aligned}$$

We conclude that $u\left(t, x_t\right)$ from the calculus of variations is the same as the $u\left(t, x_t\right)$ from the Hamilton-Jacobi equation because they solve the same equation with the same initial data at time 0. □ □

Remark 1.57.
This is not an entirely correct proof. How do we know that there is an optimal trajectory starting at x_0? If the time is short enough, we can guarantee a minimizer starting at x_0, but this is exactly the issue of uniqueness of minimizers. This proof can be made rigorous for short times.

Remark 1.58.
More generally, this is related to control theory, where we try to find

$$u(x_0, T) = \min \int_0^T L(x, f) dt + u_0(x(0)), \quad \dot{x} = h(x, f).$$

Here we chan choose some weight of influence by changing f and we are trying to optimize some cost functional. The function $u(x_0, T)$ solves a H-J equation. We can think of our claculus of variations problem as the case where the ODE for x is given by $\dot{x} = f$.

Remark 1.59.
Calculus of variation allows us to obtain meaning solutions for H-J equations after characteristics to intersect. Instead of picking which characteristic to continue, we can just look for a minimizer for a calculus of variations problem in longer time.

1.11 The Hopf-Lax Solution to Hamilton-Jacobi Equations

Date: Sep 28, 2021

1.11.1 The Hamiltonian in classical mechanics

Last time, we were solving the Hamilton-Jacobi equation

$$\begin{cases} u_t + H(x, Du) = 0 \\ u(0) = u_0 \end{cases}$$

using the calculus of variations:

$$u(x,t) = \inf_{y(t)=x} \int_0^t L(y(s), \dot{y}(s))ds + u_0(y(0))$$

Theorem 1.60.
The function u solves the Hamilton-Jacobi equation as long as the solutions the solutions stay smooth.

In the proof, we had the convex duality

$$H(x,p) = \max_q p \cdot q - L(x,q)$$

for the Hamiltonian $H(x,p)$ and the Lagrangian $L(x,q)$.

Example 1.61.
Here is an example from classical mechanics. Consider the Lagrangian

$$L(x,q) = \frac{1}{2}mq^2 - \phi(x),$$

where $\frac{1}{2}mq^2$ is kinetic energy and $\phi(x)$ is potential energy. Then

$$H(x,p) = \sup_q p \cdot q - \frac{1}{2}mq^2 + \phi(x)$$

Complete the square to get

$$\begin{aligned} &= \sup_q \frac{1}{2m}p^2 - \frac{1}{2m}(p - mq)^2 + \phi(x) \\ &= \frac{1}{2m}p^2 + \phi(x) \end{aligned}$$

In the physical interpretation, the Hamiltonian $H(x,p)$ plays the role of the energy of the system.

1.11.2 The Hopf-Lax formula

Now we will consider a special case, where $L = L(q)$ does not depend on x (and consequently $H = H(p)$). Assume that L, H are strictly convex and coercive (i.e. $\lim_{q\to\infty} \frac{L(q)}{|q|} = \infty$). The Euler-Lagrange equation tells us that

$$\frac{d}{dt}L_q(y, \dot{y}) = 0.$$

So we get that $L_q(\dot{y})$ is constant. Since L_q is a local diffeomorphism, we get that $\dot{y}$ is constant. That is, the solutions to the Euler-Lagrange equation are linear.

We claim that fixing the endpoints $y(0), y(t)$, the minimum is attained for linear trajectories.

Theorem 1.62 (Hopf-Lax formula).
If $L = L(q)$ is convex, then

$$u(x,t) = \inf_y u_0(y) + tL\left(\frac{x-y}{t}\right).$$

Proof.
Since

$$int_0^t \dot{y}(s)\, ds = y(t) - y(0),$$

we can average to get

$$\frac{1}{t}\int_0^t \dot{y}(s)\, ds = \frac{y(t)-y(0)}{t},$$

where the right hand side is the average velocity for a straight path. Then

$$\int_0^t L(\dot{y}(s))ds = t \cdot \frac{1}{t}\int_0^t L(\dot{y}(s))ds \geqslant t \cdot L(\frac{y(t)-y(0)}{t}).$$

In the inequality, we apply Jensen's inequality. In other owrds, the cost of an arbitrary path is larger than the cost of the straight path. □ □

We are not done yet, we still have to minimize $u_0(y(0))$ over the choice of $y(0)$.

1.11.3 Properties of the Hopf-Lax solution

Assume L is convex and coercive. For simplicity, also assume that u_0 is bounded . Observe that if $t > 0$, then we can restrict $q = \frac{x-y}{t}$ to a compact set. So if u_0 is also continuous, then the infimum is attained.

Proposition 1.63.
If $u_0 \in$ Lip, then $u \in$ Lip.

Proof.
Suppose we have points x_1, x_2, and y such that

$$tL(\frac{x_2-y}{t}) + g(y) = u(x,t).$$

Then

$$\begin{aligned} u(x_1,t) - u(x_2,t) &= \min_z \left\{ tL\left(\frac{x_1-z}{t}\right) + u_0(z)\right\} - tL\left(\frac{x_2-y}{t}\right) - u_0(y) \\ &\leqslant u_0(x_1-x_2+y) - u_0(y) \leqslant \mathrm{Lip}(u_0)|x_1-x_2| \end{aligned}$$

Interchaging the roles of x_1, x_2, we find

$$|u(x_1,t) - u(x_2,t)| \leqslant \mathrm{Lip}(u_0)|x_1-x_2|.$$

□ □

What if we don't assume u_0 is Lipschitz? Can we still conclude that u is Lipschitz?

Proposition 1.64.
If u_0 is continuous, then $u(t)$ is Lipschitz.

Proof.
In this case, we compare x_1 and x_2 to the same y. We have

$$u(x_1) = \inf_y u_0(y) + tL\left(\frac{x_1 - y}{t}\right)$$
$$u(x_2) = \inf_y u_0(y) + tL\left(\frac{x_2 - y}{t}\right)$$

The difference

$$\left|L\left(\frac{x_1 - y}{t}\right) - L\left(\frac{x_2 - y}{t}\right)\right| \leqslant C \cdot \frac{|x_1 - x_2|}{t},$$

where the Lipschitz constant $C = C(t)$ in the set where $\frac{x_1 - y}{t}$ and $\frac{x_2 - y}{t}$ live. Where should we look? $\frac{y - x_1}{t}, \frac{y - x_2}{t}$ cannot be too large. Let $x = x_1 = x_2$, and compare the straight trajectory to an arbitrary trajecotry. The oblique trajectory loses if $u_0 + tL(0) \leqslant u_0(y) + tL(\frac{x-y}{t})$. This is when $2M/t \leqslant L(\frac{x-y}{t})$. So we can restrict to y such that $L(\frac{x-y}{t}) \leqslant 2M/t$. So $\frac{x-y}{t}$ is in a compact set depending on t. Then the conclusion is that

$$|u(x_1, t) - u(x_2, t)| \leqslant C(t) \cdot \frac{|x_1 - x_2|}{t}$$

where $C(t)$ is the Lipschitz constant for L in the region $L(q) \leqslant \frac{C}{t}$. This Lipschitz constant goes to ∞ as $t \to 0$. □ □

In terms of the Hamilton-Jacobi equation, there will be lots of velocities with different speeds. So there is only an average velocity that survives. We say that this PDE has a mild **regularizing effect**. [Will: Todo]Ask about the proof and the regularizing effect

1.11.4 Almost everywhere solvability of the H-J equation

Theorem 1.65.
If u is a Lipschitz function, then u is differentiable almost everywhere.

Corollary 1.66.
The solution u is differentiable almost everywhere.

Proposition 1.67.
Let (x, t) be a differentiable point for u. Then the H-J equation hold at (x, t).

Corollary 1.68.
The function u solves the H-J equation almost everywhere.

Let's prove the proposition first.

Proof.
We can think of the Hamilton-Jacobi equation as proving two separate inequalities. If our trajectory is optimal, then it is optimal if we only look at the trajectory at a shorter length of time. Look at the optimal trajectory, ending at y and with slope $\frac{x-y}{t}$. Then,

$$u(x, t) = u_0(y) + tL\left(\frac{x - y}{t}\right)$$

so

$$u\left(x - h\frac{x - y}{t}, t - h\right) = u_0(y) + (t - h)L\left(\frac{x - y}{t}\right)$$

The first equation tells us that y is the optimal trajectory for (x,t), and the second says that y is optimal for $\left(x \cdot h\frac{x-y}{t}, t-h\right)$. Let $q = \frac{x-y}{t}$. Then dividing by h gives

$$\frac{u(x,t) - u(x-hq, t-h)}{h} = hL(q)$$

Letting $h \to 0$ gives

$$\partial_x u \cdot q + \partial_t u = L(q)$$

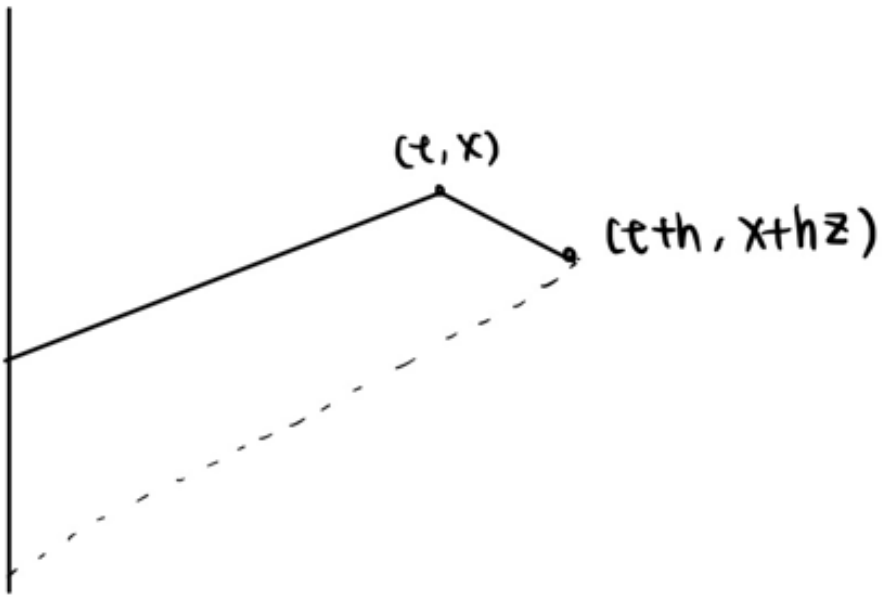
So for this special q we have chosen,

$$\partial_t u + \partial_x u \cdot q - L(q) = 0$$

We want to think of this in terms of the Legendre transform. Since $H(p) = \sup p \cdot q - L(q)$, the latter half of our equation, $\partial_x u \cdot q - L(q)$, is $\leqslant H\left(\partial_x u\right)$. So we get

$$\partial_t u + H\left(\partial_x u\right) \geqslant 0$$

Now we want to produce the other inequality. Notice that for the previous inequality, it was enough to work with a specific value of q, whereas for this direction, we will need to look at all values of q. Instead of looking at the past of (t,x), look at the future of (t,x). Our trajectory looks like



One trajectory from $(t+h, x+hz)$ is to go through x, but this may not be optimal. So

$$u(t+h, x+hz) \leqslant u(t,x) + \underbrace{hL(z)}_{\int_t^{t+h} L(z)ds}$$

As before, subtract the right hand side, divide by h and let $h \to 0$. Then we get

$$\frac{u(t+h, x+hz) - u(t,x)}{h} \leqslant L(t) \Longrightarrow \partial u + \partial_x u z \leqslant L(z).$$

So we have proven that for all z,

$$\partial_t u + \partial_x u \cdot z - L(z) \leqslant 0.$$

Taking the supremum over all z, we get

$$\partial_t u + H\left(\partial_x u\right) \leqslant 0.$$

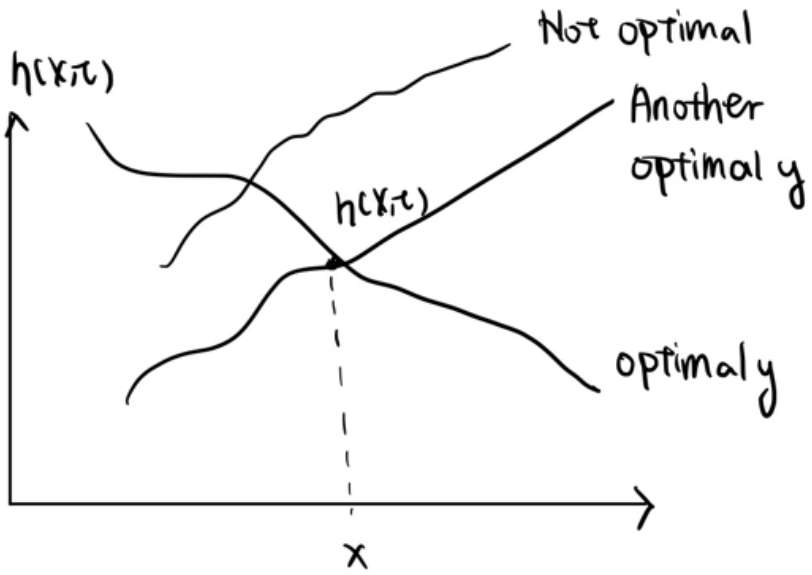
□ □

Now we will tell a story. The details are in Evans' book, but the overall story is more important. We want to ask a question: Does solving the Hamilton-Jacobi equation almost everywhere suffice to guarantee uniqueness for Hamilton-Jacobi? Equivalently, does this guarantee that u is the minimal value function? The answer is no.

Are there other interesting properties for the function u ? Look at the Hopf-Lax formula

$$u(x,t) = \inf_y u_0(y) + tL\left(\frac{x-y}{t}\right)$$

Observe that this is an infimum of functions which are smooth in x. We can compare what this looks like for different optimal/nonoptimal y:



Since we are taking a minimum, we can see that our curve could ave a corner pointing upwards, but a conner pointing downwards is not possible. This points to a concavity property of our solution.

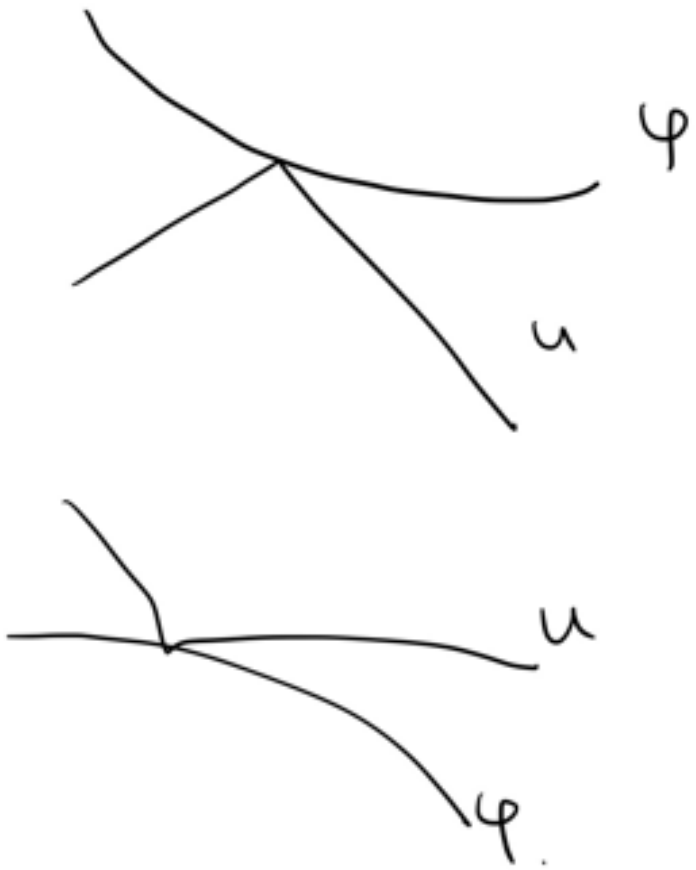
Proposition 1.69.
u is semiconcave.

Concave means that $u(t,x) \geqslant \frac{u(t,x+y)+u(t,x-y)}{2}$. **Semiconcave** means that

$$u(t,x) \geqslant \frac{u(t,x+y)+u(t,x-y)}{2} - c\cdot|x-y|^2.$$

Theorem 1.70.
The optimal value function u is the unique seminconcave solution to the H-J equation.

The proof is in Evans, but it is a little hard to follow. There is a better way to do things! Instead of plugging in u to check whether it satisfies the equation, if we have a corner, draw a tangent test function φ with $\varphi_t + H(\partial_x\phi) \geqslant 0$ or $\varphi_t + H(\partial_x\phi) \leqslant 0$.



These are called **viscosity solutions** for H-J equations.

1.12 Introduction to Distribution Theory

Date: Sep 30, 2021

1.12.1 Weak solutions to PDEs

For the next month or so, our goal will be to study linear, constant coefficient PDEs

$$P(\partial)u = f, \quad P(\partial) = \sum_{|\alpha| \leqslant m} c_\alpha \partial^\alpha.$$

We will first take a detour to study the theory of distributions. First, some motivation:

Example 1.71.
Recall the transport equation

$$\begin{cases} \left(\partial_t + \sum_j A_j \partial_j\right) u = 0 \\ u(0) = u_0 \end{cases}$$

with constant coefficients A_j. The characteristics are given by $\dot{x} = A$, which gives $x(t) = x(0) + tA$. This means that $\dot{u} = 0$ along these characteristics, so $u(x(t), t) = u(x(0), 0)$. In other words,

$$u(x, t) = u_0(x - tA)$$

Classically, if $u_0 \in C^1$, then $u \in C^1$. What if $u_0 \in C$? It doesn't make sense to say that the solution u is continuous because we need to take derivatives. If we interpret the equation as a directional derivative, $u_0 \in C$ gives a solution. This interpretation relies strongly on the specific problem. Can we treat this problem in general?

Suppose we have a smooth function $\varphi \in C_0^\infty$. We can write the equation as the condition

$$\int_{\mathbb{R}^{n+1}} (\partial_t + A_j \partial_j)\, u \varphi dx = 0$$

where a function is 0 if it integrates to be 0 against all $\varphi \in C_0^\infty$. Now integrate by parts to get

$$-\int_{\mathbb{R}^{n+1}} u\, (\partial_t + A_j \partial_j)\, \varphi dx = 0, \quad \forall \varphi \in C_0^\infty$$

which applies to all $u \in C$. Our continuous solution will be a solution to this integral equation.

Definition 1.72 (Weak solution).
u is a weak solution to a PDE if the corresponding integral equation holds for all $\varphi \in \mathrm{C}_0^\infty$.

Example 1.73.
Recall the Burgers' equation

$$u_t + uu_x = 0, \quad u(0) = u_0$$

The characteristics are given by $\dot{x} = u$ and $\dot{u} = 0$. The characteristics will intersect, and at the point of intersection of characteristics, the solution will start to develop a jump discontinuity, known as a shock.

In this problem, if we think of the equation as a directional derivative, the derivative along the characteristics are different when they intersect, so we cannot get a solution. However, we can similarly look for a weak solution by integrating by parts as before. When we do this, we want to think of uu_x as $\frac{1}{2}\partial_x(u^2)$.

1.12.2 Topologies on vector space

The key idea in the theory of distribution is that we can think of a function $u : \mathbb{R}^n \to \mathbb{R}$ as a linear map on all $\varphi \in C_0(\mathbb{R}^n)$ via

$$u(\varphi) := \int_{\mathbb{R}^n} u \cdot \varphi dx$$

Observe that if $u(\varphi) = 0$ for all φ, then $u = 0$. We will use the notation $\mathcal{D} = C_0^\infty$ to refer to the smooth functions with compact support. Obesrve that $\mathcal{D}$ is a linear space. What is the topology of $\mathcal{D}$? Recall that C is a normed space, with

$$\|u\|_C = \sup_{x \in \mathbb{R}^n} |u(x)|$$

Recall:

Definition 1.74 (Normed space).
A normed space is a vector space V, with a norm map $\|\cdot\| : V \to \mathbb{R}$ (or $\mathbb{C}$) satisfying

(a) $\|u\| \geqslant 0$, with equality iff $u = 0$.
(b) $\|\lambda u\| = |\lambda| \|u\|$ for all $\lambda \in \mathbb{R}$ (or $\mathbb{C}$).
(c) $\|u + v\| \leqslant \|u\| + \|v\|$.

We obtain a metric space structure is given by $d(u, v) = \|u - v\|$. Recall that completeness of a metric space means that every Cauchy sequence is convergent.

Definition 1.75 (Banach Space, Hilbert space).
A Banach space is a complete normed space. A Hilbert space is a vector space with a complete inner product $\langle u, v \rangle = u \cdot v$.

Example 1.76.
The L^2 sapce is given by

$$L^2(\mathbb{R}^n) = \left\{ u : \mathbb{R}^n \to \mathbb{R} \middle| \int |u|^2 dx < \infty \right\}$$

This space is a Hilbert space, given the inner product

$$u \cdot v = \int_{\mathbb{R}^n} uv \, dx$$

(with v replaced by $\bar{v}$ in the complex case).

Hilbert spaces are a special case of Banach spaces, but a single space can have different norm structures on it.

Example 1.77.
We can equip $\mathbb{R}^n$ with the norm $\|v\|^2 = \sum_j v_j^2$ which comes from the usual dot product (a Hilbert space structure). We can also equip $\mathbb{R}^n$ with the L^p norm $\|v\|^p = \sum_j |v_j|^p$ with $1 \leqslant p < \infty$, which gives a Banach space structure.

Example 1.78.
C^k is a Banach space with the norm

$$\|u\| = \sup_{|\alpha| \leqslant K} \sup_{x \in \mathbb{R}^n} |\partial^\alpha u(x)|$$

Returning to our objective, what norm can we give $C^\infty(\mathbb{R}^n)$? We can define

$$\|u\|_\alpha = p_\alpha(u) := \sup_{x\in\mathbb{R}^n} |\partial^\alpha u(x)|$$

The problem is that we have infinitely many of these. What would $u_n \to u$ mean in $C^\infty(\mathbb{R}^n)$? We want to say that $\partial^\alpha u_n \to \partial^\alpha u$ uniformly for all α.

The solution is to use all the $\|\cdot\|_\alpha$ as seminorms, which satisfy all the norm conditions except for $\|u\| = 0 \implies u = 0$.

Definition 1.79 (Locally convex space, Fréchet space).
Locally convex spaces are vector spaces equipped with a family of seminorms. A complete, locally convex space is called a Fréchet space.

In a locally convex space,

$$p_\alpha(u) = 0 \quad \forall\alpha \implies u = 0$$

Why is this called a "locally covnex space"? The idea is that each seminorm gives you neighborhoods of points, which may not be nested in each other for different seminorms. But these are all convex neighborhoods, and we can intersect these neighborhoods to get more convex neighborhoods around every point. The picture of our function spaces looks like

$$\text{Hilbert spaces} \subsetneq \text{Banach spaces} \subsetneq \text{Fréchet spaces}.$$

Example 1.80.
We will use the notation $\mathcal{E} = \{u \in \mathbb{R}^n \to \mathbb{R} \mid u \text{ is smooth }\}$. Here, we ask for nothing at ∞. What does $u_n \to u$ mean in $\mathcal{E}$? We can define this as $\partial^\alpha u_n \to \partial^\alpha$ uniformly on compact sets. For this space, we need to use the collection of seminorms

$$p_{\alpha,K}(u) = \sup_{x\in K} |\partial^\alpha u(x)|, \quad \alpha \in \mathbb{N}^d, K \text{ compact}.$$

We don't need to check all compact sets; it suffices to take nested balls with radius going to ∞. With this topology, $\mathcal{E}$ is a locally convex space.

For $\mathcal{D}$, we have an issue: if we have a sequence of functions of compact support, the support may grow to not be compact in the limit. To solve this, there is a notion called the inductive limit of locally convex spaces, essentially cooked up only to describe $\mathcal{D}$.

Definition 1.81 (Inductive limit).
We say u is the inductive limit of u_n in $\mathcal{D}$ if

- $\partial^\alpha u_n \to \partial^\alpha u$ uniformly.
- There is a compact set K such that $\operatorname{supp} u_n \subseteq K$.

Remark 1.82.
If $u \in \mathrm{C}(\mathbb{R}^n)$ and $\varphi \in \mathcal{D}$, the map $\varphi \to u(\varphi) = \int u\varphi\, dx$ is continuous.

Definition 1.83 (Space of distrubtions).
The space of distributions, denoted D' or $D*$ is the sapce of linear, continuous $f : \mathcal{D} \to R$.

This seems to separate us from our original goal. If we have a function, we can get a distribution, but if we have a distribution, we can't always get a function back; instead, we get generalized functions. The term "distribution" comes from the French school, whereas the term "generalized functions" comes from the Russian school.

1.12.3 Examples of distributions

Here are some examples of distributions.

Example 1.84.
The dirac mass at 0 is

$$\delta_0(\varphi) = \varphi(0)$$

Example 1.85.
Another distribution is

$$\delta_0'(\varphi) = -\varphi'(0)$$

The reason for the minus sign will become apparent later on. In general, we can define

$$\delta_x^{(\alpha)}(\varphi) = (-1)^{|\alpha|}\partial^\alpha \varphi(x)$$

The space $\mathcal{D}'$ of distributions is a linear space. [2] It has the topology of weak convergence: $f_n \to f$ in $\mathcal{D}'$ if

$$f_n(\varphi) \to f(\varphi) \quad \forall \varphi \in \mathcal{D}$$

Example 1.86.
Can we approximate δ_0 with functions? This may shed some light on what generalized functions look like. Let

$$u_n(x) = \begin{cases} n/2 & x \in [-1/n, 1/n] \\ 0 & \text{otherwise} \end{cases}$$

Here, $\int u_n = 1$ for all n. If we try to take the limit in the sense of distributions, we get (in 1 dimension):

$$\begin{aligned} u_n(\varphi) &= \int u_n \cdot \varphi \, dx \\ &= \frac{n}{2} \int_{-1/n}^{1/n} \varphi(x)\, dx \\ &= \frac{1}{2} \int_{-1}^{1} \varphi(y/n) dx \overset{n\to\infty}{\longrightarrow} \varphi(0), \end{aligned}$$

So $u_n(\varphi) \to \delta_0(\varphi)$. That is, $u_n \to \delta_0$.

Remark 1.87.
In $\mathbb{R}^n$, we could use

$$u_\varepsilon = \frac{1}{\varepsilon^n c_n} 1_{B(0,\varepsilon)}, \quad c_n = |B(0,1)|$$

In n dimensions, this has size $\sim 1/\varepsilon^n$.

Remark 1.88.
We could also use $\mathcal{D}$ functions. If $\varphi \in \mathcal{D}$, with $\int \varphi = 1$, then we can define the rescaled function (at scale ε)

$$\varphi_\varepsilon(x) = \frac{1}{\varepsilon}\varphi(x/\varepsilon)$$

By the same argument, $\varphi_\varepsilon \to \delta_0$ in $\mathcal{D}'$.

Next time, we will see how we can think of distributions as solutions to PDEs. This will require knowing things like how to differentiate distributions.

1.13 Distribution & Homogeneous Distribution

Date: Oct 5, 2021

1.13.1 Operations on distributions

Last time, we introduced distributions. We had the set $\mathcal{D} = C_0^\infty$ of test functions and the set $\mathcal{D}'$ of distributions, continuous linear maps $\mathcal{F} : \mathcal{D} \to \mathbb{R}$. If u is a function, we interpreted it as a distribution via

$$u(\phi) = \int u\phi dx$$

So we can think of distributions as generalized functions. We also saw distributions as a limit of functions, in this weak sense.

Now, we want to see distributions as solutions to PDEs, so we need to think about operations with distributions.

Differentiation

We want to define $u \mapsto \partial_j u$ for distributions. First suppose u is a function. Then $\partial_j u$ is a function with

$$\begin{aligned}\partial_j u(\phi) &= \int \partial_j u\phi dx \\ &= -\int u \cdot \partial_j \phi dx \\ &= -u\left(\partial_j \phi\right)\end{aligned}$$

We can take this as a definition.

Definition 1.89 (Weak derivative).
If $u \in \mathcal{D}'$, we define $\partial_j u$ by $\partial_j u(\varphi) = -u(\partial_j \varphi)$.

Remark 1.90.
If $u \in \mathrm{C}^1$, then u is the same classically and as a distribution.

Example 1.91.
Consider the Heaviside function

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

in 1 dimension. Then $\partial_x H = \delta_0$ away from 0 , in the classical sense. We can check that

$$\begin{aligned}\partial_x H(\phi) &= -H\left(\partial_x \phi\right) \\ &= -\int H(x)\partial_x \phi dx \\ &= -\int_0^\infty -\partial_x \phi(x) dx \\ &= -\phi|^\infty \\ &= \phi(0) \\ &= \delta_0(\phi)\end{aligned}$$

so $\partial_x H = \delta_0$ as a distribution. The idea is that when we have a jump discontinuity, differentiating

gives us a Dirac mass.

Example 1.92.
What is the derivative of the Dirac mass?

$$\begin{aligned}\partial_x \delta_0(\phi) &= -\delta_0(\partial_x \phi) \\ &= -\delta_x \phi(0) \\ &= \delta_0'(0)\end{aligned}$$

So the derivative of δ_0 is what we previously called δ_0'. Similarly, we can have $\partial^\alpha \delta_0 = \delta_0^{(\alpha)}$ for a multi-index α.

Multiplication by smooth functions

Suppose $\psi \in \mathcal{E}$ and u is a function. Then ψu is a function. What if $u \in \mathcal{D}'$? If u is a function, then

$$\begin{aligned}\psi u(\phi) &= \int \psi u \phi dx \\ &= \int u \underbrace{\psi\phi}_{\in \mathcal{D}} dx \\ &= u(\psi\phi)\end{aligned}$$

We can again take this as a definition.

Definition 1.93.
If $u \in \mathcal{D}'$ and $\varphi \in \mathcal{E}$, we define φu by $\phi u(\phi) = u(\varphi\phi)$.

The Leibniz rule for derivatives says

$$\partial(\psi u) = \partial\psi \cdot u + \psi \cdot \partial u$$

Using these definitions, this rule also holds for $u \in \mathcal{D}'$ and $\psi \in \mathcal{E}$. If we have the equation $P(x,\partial)u = f$ with $P(x,\partial) = \sum c_\alpha(x)\partial^\alpha$, then all these operations are well-defined for distributions, so we can think of distribution solutions to PDEs.

1.13.2 The support of a distribution

Recall that if u is a function, its support is the largest closed set "where u is nonzero." In particular,

$$x_0 \notin \operatorname{supp} u \iff u = 0 \text{ in } B(X_0, r) \text{ for some } r > 0$$

Definition 1.94 (Support of distribution).
If $u \in \mathcal{D}'$, its support is the closed set defined by

$$x_0 \notin \operatorname{supp} u(\phi) \iff u(\phi) = 0 \text{ for all } \phi \in \mathcal{D} \text{ with } \operatorname{supp}\phi \subseteq B(x_0, r)$$

Example 1.95.
The support of the Dirac mass is $\operatorname{supp}\delta_0 = \{0\}$. If $x_0 \neq 0$, there is a ball $B(x_0, r) \cap \{0\} = \varnothing$. Then if we let $\phi \in \mathcal{D}$ have $\operatorname{supp}\phi \subseteq B(x_0, r)$, then $\delta_0(\phi) = \phi(0) = 0$.

Let $\mathcal{E}'$ denote the **compactly supported distributions**.

Proposition 1.96.
If $f \in \mathcal{E}'$, then f "naturally" extends to a continuous linear function of $\mathcal{E}$.

Proof.
We know $f(\phi)$ when $\phi \in \mathcal{D}$. Because supp $f \in B(0,R), f(\phi) = 0$ if ϕ is supported outside $B(0,R)$. We can truncate ϕ outside B as follows: Replace ϕ by $\chi\phi$, where χ is a cutoff function with compact support, supp $\chi \subseteq B(0,2R)$, and $\chi = 1$ in $B(0,R)$. Then

$$\begin{aligned} f(\phi) &= f(\chi\phi) + f((1-\chi)\phi) \\ &= f(\chi\phi) \end{aligned}$$

So for $\phi \in \mathcal{E}$, define $f(\phi) := f(\chi\phi)$. □

We have the following picture:

$$\begin{array}{ccc} \mathcal{D} & \xrightarrow{dual} & \mathcal{D}' \\ \downarrow \subseteq & & \subseteq \uparrow \\ \mathcal{E} & \xrightarrow{dual} & \mathcal{E}' \end{array}$$

We will extend this picture later when we learn about the Fourier transform.

1.13.3 Homogeneous distributions

Example 1.97.
The polynomial $f(x) = x^n$ is a homogeneous polynomial. We can express this homogeneity by

$$f(\lambda x) = \lambda^n f(x)$$

where n is the homogeneity index.

Example 1.98.
The homogeneity index does not have to be an integer. If we have $f(x) = |x|^\alpha$, then

$$f(\lambda x) = \lambda^\alpha f(x)$$

for $\lambda > 0$. If α is not an integer, this is not smooth at 0. Is $|x|^\alpha$ a distribution? This is related to the question of whether $|x|^\alpha$ is integrable (away from infinity). In 1 dimension, $\int |x|^\alpha dx$ exists if $\alpha > -1$. In n dimensions, we can use polar coordinates:

$$\int |x|^\alpha dx = c_n \int r^\alpha r^{n-1} dr$$

where c_n is the volume of the unit ball in n-dimensions. Here, we need $\alpha + n - 1 > -1$, i.e. $\alpha > -n$. So $\frac{1}{|x|^n}$ is borderline.

Example 1.99.
The heaviside function is homogeneous of index 0 :

$$H(\lambda x) = \lambda^0 H(x)$$

for $\lambda > 0$.

Example 1.100.
In 2 dimensions (expressed in polar coordinates (r, θ)), the function

$$f(x) = r^\alpha g(\theta)$$

is homogeneous of index α.

For functions, the homogeneity condition $f(\lambda x) = \lambda^\alpha f(x)$ has a distributional interpretation:

$$\int f(\lambda x)\phi(x)dx = \lambda^\alpha \int f(x)\phi(x)dx$$

Applying a change of variables on the left,

$$\int f(y)\phi(y/\lambda)\frac{1}{\lambda^n}dy = \lambda^\alpha \int f(x)\phi(x)dx$$

Denoting $\phi_\lambda(x) = \lambda^{-n}\phi(x/\lambda)$, we get the relation

$$f(\phi_\lambda) = \lambda^\alpha f(\phi)$$

which is meaningful for distributions.

Definition 1.101 (Homogeneous distribution).
A distribution $f \in \mathcal{D}'$ is homogeneous of order α if

$$f(\phi_\lambda) = \lambda^\alpha f(\phi)$$

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

Example 1.102.
Can we think of the Dirac mass δ_0 as a homogeneous distribution?

$$\delta_0(\phi_\lambda) = \phi_\lambda(0) = \lambda^{-n}\phi(0) = \lambda^{-n}\delta_0(\phi)$$

so δ_0 has homogeneity $-n$.

In calculus, we have $\partial_x x^n = nx^{n-1}$. That is, we differentiate something which is homogeneous of order n and get something which is homogeneous of order $n-1$.

Proposition 1.103.
If $f \in \mathcal{D}'$ is homogeneous of order α, then $\partial_x f$ is homogeneous of order $\alpha - 1$.

Proof.
The chain rule works for functions, so it also works using the definition for distributions by passing the derivative to the test function. □

Example 1.104.
The Heaviside function is homogeneous of order 0 , and $\partial_x H = \delta_0$ is homogeneous of order -1. Similarly, $\partial_x \delta_0 = \delta_0'$ is homogeneous of order -1

In 1 dimension, we want to classify homogeneous distributions. Start with functions and $\alpha > -1$. We need to assign $f(-1)$ and $f(1)$, so this is a linear space of dimension 2. Here is a basis:

$$x_+^\alpha = \begin{cases} 0 & x < 0 \\ x^\alpha & x > 0 \end{cases} \quad x_-^\alpha = \begin{cases} |x|^\alpha & x < 0 \\ 0 & x > 0 \end{cases}$$

Then $|x|^\alpha = x_+^\alpha + x_-^\alpha$, and

$$\partial_x x_+^\alpha = \alpha x_+^{\alpha-1}, \quad \partial_x x_-^\alpha = -\alpha x_-^{\alpha-1}$$

Now look at when $\alpha \in (-2, -1)$. We can define

$$\partial_x x_+^{\alpha+1} := (\alpha+1) x_+^\alpha$$

If we repeat this, we can get homogeneous distributions to all noninteger negative αs.

What about $\alpha = -1$? We have δ_0. At order 0, we have 2 homogeneous distributions: H and the constant 1 function. But differentiating these gives δ_0 and 0, which do not have a 2 dimensional span. Other candidates are $\frac{1}{|x|}$ or $\frac{1}{x}$. We can look at the integrals

$$\int \frac{1}{|x|}\phi(x)dx \quad \int \frac{1}{x}\phi(x)dx$$

On the left, there may be no cancelation at 0, but we may be able to get some cancelation at 0 for the right integral. We may try to define

$$\int_{\mathbb{R}} \frac{1}{x}\phi(x)dx := \lim_{\varepsilon \to 0} \int_{R \setminus [-\varepsilon, \varepsilon]} \frac{1}{x}\phi(x)dx$$

Does this limit exist? We can look at

$$\int_{[-1,1]\setminus[-\varepsilon,\varepsilon]} \frac{1}{x}\phi(x)dx = \int_{-1}^{1} + \int_{\varepsilon}^{1}$$

Use the change of variables $y = -x$ on the left integral to get

$$= \int_\varepsilon^1 \frac{\phi(x) - \phi(-x)}{x} dx$$

$\phi(x) - \phi(-x)$ is $o(x)$, so this converges. Thus, we can define the principal value PV

Definition 1.105 (Principal value of $\frac{1}{x}$).
The principal value PV of $\frac{1}{x}$ is

$$\mathrm{PV}\frac{1}{x}(\phi) = \lim_{\varepsilon \to 0} \int_{\mathbb{R}\setminus -\varepsilon,\varepsilon]} \frac{\phi(x)}{x} dx$$

which is homogeneous of order -1.

1.14 Homogeneous Distribution of Order -1, Convolution, and Fundamental Solutions

Date: Oct 7, 2021

1.14.1 Special homogeneous distributions of order -1

The principle value of $1/x$ as a complex limit

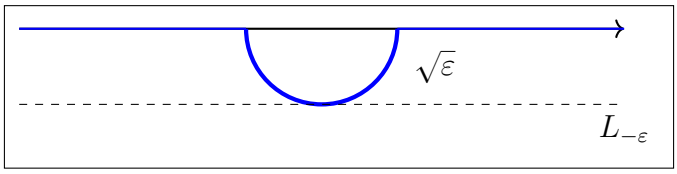
Last time, we were discussing homogeneous distributions. When classifying homogeneous distributions of order -1 in 1 dimension, we saw two interesting distributions:

$$\delta_0, \quad \mathrm{PV}\frac{1}{x}$$

If you like complex analysis, you can consider the function

$$f(z) = \frac{1}{z} = \frac{1}{x+iy}$$

Then $f(z) = \frac{1}{x-i\varepsilon}$ on the line $L_{-\varepsilon}$ below the real line:



What is $\lim_{\xi\to 0}\frac{1}{x-i\varepsilon}$? Apply this to a test function:

$$\begin{aligned}\frac{1}{x-i\varepsilon}(\varphi) = \int \frac{\varphi(x)}{x-i\varepsilon}dx \\ &\approx \int_{\mathbb{R}\setminus[\varepsilon,\varepsilon)} \frac{\varphi(x)}{x-i\varepsilon} + \int_{\frac{1}{2}C_\varepsilon} \frac{\varphi(z)}{z}dz \\ &\approx PV\frac{1}{x}(\varphi) + \varphi(0)\cdot\int_{1/2C_\varepsilon}\frac{1}{z}\,dz\end{aligned}$$

Write $\ln z = \ln|z| + i\arg z$. Then $z = \varepsilon e^{i\theta}$ for $\theta \in [\pi, 2\pi]$

$$\begin{aligned}&= \mathrm{PV}\,\frac{1}{x}(\varphi) + \varphi(0)\cdot\int_\pi^{2\pi}\frac{i\varepsilon e^{i\theta}}{\varepsilon e^{i\theta}}d\theta \\ &= \mathrm{PV}\,\frac{1}{x}(\varphi) + \varphi(0)\pi i\end{aligned}$$

So

$$\lim_{\varepsilon\to 0}\frac{1}{x-i\varepsilon} = \mathrm{PV}\frac{1}{x} + \pi i\delta_0$$

If we do the same approximation from the line L_ε above the real line, we get

$$\lim_{\varepsilon\to 0}\frac{1}{x+i\varepsilon} = \mathrm{PV}\frac{1}{x} - \pi i\delta_0$$

What is $\partial_x \mathrm{PV}\frac{1}{x}$? We can calculate that

$$-\lim_{\varepsilon\to 0}\frac{1}{(x-i\varepsilon)^2} = \left(\mathrm{PV}\frac{1}{x}\right)' + \pi i\delta_0'$$

and repeat this idea to find the derivatives of PV $\frac{1}{x}$.

$1/|x|$ as a distribution

What is $\frac{1}{|x|}$ as a distribution?

$$\lim_{\varepsilon\to 0}\int_{[-1,1]\setminus[-\varepsilon,\xi)}\frac{1}{|x|}\varphi(x)dx = \int\frac{1}{|x|}(\varphi(x)-\varphi(0))dx+\varphi(0)\int\frac{1}{|x|}dx$$
$$\to\int_{-1}^{1}\frac{1}{|x|}(\varphi(x)-\varphi(0))dx+2\varphi|\log\varepsilon|$$

But this does not converge as $\varepsilon\to 0$. So we can try to **renormalize**, calculating the integral when we subtract out the divergent term:

$$\frac{1}{|x|}(\varphi) := \lim_{\varepsilon\to 0}\int_{\mathbb{R}\setminus[-\varepsilon,\varepsilon}\frac{1}{|x|}(\varphi(x)-\varphi(0))dx-2\varphi(0)|\log\varepsilon|$$

However, this breaks the homogeneity.

1.14.2 Convolution

Definition 1.106 (Convolution).
Let $\varphi,\psi\in\mathcal{D}$. The convolution is the function

$$(\varphi*\psi)(x)=\int\varphi(y)\psi(x-y)dy$$

Observe that this is smooth in x. What about the support?

Proposition 1.107.

$$\operatorname{supp}\varphi*\psi\subseteq\operatorname{supp}\varphi+\operatorname{supp}\psi.$$

So we can think about convolution as a function

$$*:\mathcal{D}\times\mathcal{D}\to\mathcal{D}$$

Proposition 1.108 (Commutativity of convolution).

$$\varphi*\psi=\psi*\varphi.$$

Proposition 1.109 (Associativity of convolution).

$$\varphi*(\psi*\zeta)=(\varphi*\psi)*\zeta.$$

So $(\mathcal{D},+,*)$ is a commutative algebra. We have another commutative algebra structure on $\mathcal{D}$, $(\mathcal{D},+,\cdot)$. We will later see that these structures are not unrelated; they are mirror images of each other. With multiplication, we have the Leibniz rule:

$$\partial(\psi\varphi)=\partial\psi\cdot\varphi+\psi\cdot\partial\psi$$

We don't exactly have a Leibniz rule for convolution:

Proposition 1.110.

$$\varphi(\psi * \varphi) = \psi * \partial\varphi = \varphi * \partial\psi.$$

Proposition 1.111.
If $\varphi \in L^1$ and $\psi \in L^\infty$, then

$$\|\varphi * \psi\|_{L^\infty} \leqslant \|\varphi\|_{L^1}\|\psi\|_{L^\infty}$$

When you think of convolution, you want to think of two things: regularity and support. If $\varphi \in \mathcal{D}$ and $\psi \in \mathcal{E}$, then we lose information about the support, so $\varphi * \psi \in \mathcal{E}$. So $\mathcal{D} * \mathcal{E} \to \mathcal{E}$. On the other hand, if we take a derivative of the convolution, we just need to be able to take a derivative of one of the factors. Here is the takeaway:

- For support, we need the support of both factors.
- For regularity, we need the regularity of just one factor.

We can think of convolutions as distributions: If $\varphi \in \mathcal{E}$ and $\psi \in \mathcal{D}$

$$\varphi * \psi(x) = \varphi(\psi(x - \cdot))$$

This right hand side is well-defined even if $\varphi \in \mathcal{D}'$. So we see that

$$\mathcal{D}' * \mathcal{D} \to \mathcal{E}$$

Similarly, we have

$$\mathcal{E}' * \mathcal{D} \to \mathcal{D}$$

What about $\mathcal{E}' * \mathcal{E}'$? If $u, v, \varphi \in \mathcal{D}$, then

$$(u * v)(\varphi) = \iint u(y)v(x - y)dy\varphi(x)dx$$

Change variables using $z = x - y$ so $\varphi(x) = z + y$.

$$\begin{aligned} &= \iint u(y)v(z)\varphi(z + y)dydz \\ &= \int u(y) \underbrace{\int v(z)\varphi(z + y)dz}_{v(\varphi(y+))}\, dy \\ &= u(v(\varphi(y + \cdot))) \end{aligned}$$

This conclusion makes sense even if $u, v \in \mathcal{E}'$. We can make this precise if we can approximate elements of $\mathcal{E}'$ by elements in $\mathcal{E}$. So we get

$$\mathcal{E}' * \mathcal{E}' \to \mathcal{E}'$$

However, $\mathcal{D}' * \mathcal{D}'$ is undefined.

1.14.3 Fundamental solutions to PDEs

Now suppose we have the PDE

$$P(\partial)u = f$$

where P is linear with constant coefficients and f is a distribution. The simplest f we can consider is δ_0, which gives us the equation

$$P(\partial)K = \delta_0$$

The next simplest f we can consider is δ_{x_0}. So we get

$$P(\partial)K(\cdot - x_0) = \delta_{x_0}$$

by invariance with respect to translations. Can we write a general function as a superposition of δ functions? If we have a Riemann integral, we can approximate it by a sum of pieces which look like Dirac masses.

So can we make sense of something that looks like

$$f = \int f(x_0)\,\delta_{x_0} dx_0?$$

We can define this by applying f to a test function:

$$\varphi(\varphi) = \int f(x_0)\,\underbrace{\delta_{x_0}(\varphi)}_{=\varphi(x_0)}\,dx_0$$

So if we can deal with a Dirac masses, we can deal with a lienar combination of Dirac masses and hence any function as a superposition of Dirac masses. So the solution should looks like

$$u(x) = \int f(x_0)\,K(x - x_0)\,dx_0$$

This was some intuition (Or maybe confusion!), but here are some definitions.

Definition 1.112 (Fundamental solution).
K is a fundamental solution of $P(\partial)$ if

$$P(\partial)K = \delta_0.$$

Proposition 1.113.
The function $u = K * f$ solves the equation

$$P(\partial)u = f$$

Proof.

$$\begin{aligned} P(\partial)u &= P(\partial)(K * f) \\ &= P(\partial K) * f \\ &= \delta_0 * f \end{aligned}$$

We are done if $f * \delta_0 = f$. If $f \in \mathcal{D}$, then

$$f * \delta_0(x) = \delta_0(f(x - \cdot)) = f(x)$$

The same works for $f \in \mathcal{D}'$. □ □

In this proof, we saw that δ_0 is the identity with respect to $*$ For multiplication, 1 is the identity. The constant 1 function has support on all of $\mathbb{R}^n$, but it has regularity; conversely, δ_0 has 1 point as it support but no regularity. You can think of these as opposites.

Besides, with our notation, the fundamental theorem of calculus looks like this:

Theorem 1.114.
If $\partial_x u = f$ in $\mathbb{R}$, then

$$u = \int f(x)dx + C$$

If we specify that $u(-\infty) = 0$, then

$$u(x) = \int_{-\infty}^{x} f(y)dy$$

We want to interpret this as a convolution. First, let's compute the fundamental solution:

$$\partial_x K = \delta_0, \quad K(-\infty) = 0$$

This tells us that

$$K = H(x)$$

is the Heaviside function. By our proposition, $u = K * f$. We can write this as

$$u(x) = \int H(x-y)f(y)dy$$

For $H(x-y)$ to give 1 and not 0 , we need $x - y > 0$.

$$= \int_{-\infty}^{x} f(y)dy$$

Is the fundamental solution K unique? In general, if K is a constant solution, then $K + C$ is a fundamental solution for any constant C. If we ask for $K = 0$ at $-\infty$, we get $K = H$. But if we ask for $K = 0$ at $+\infty$, we get $K = H - 1$. If we ask for K to be odd, we get $K = H - 1/2$.

1.15 Fundamental Solutions for 2-D PDE & Laplacian

Date: Oct 12, 2021

1.15.1 Fundamental solutions in 1 and 2 dimensions

Last time, we discussed fundamental solutions for partial differential equations. Suppose we have a differential operator in 1 dimension

$$P(\partial)K = \delta_0$$

Solve the homogeneous equation and look for the fundamental solution

$$K(x) = \begin{cases} u_1^{\text{hom}}(x) & x < 0 \\ u_2^{\text{hom}}(x) & x > 0 \end{cases}$$

Plug this in into $P(\partial)K = \delta_0$ and get a linear system for the constants. As an exercise, try to solve the equation with the operator $P(\partial) = \partial^2 - 1$.

What about in 2 dimensions? In complex analysis, one way to specify whether a function is holomorphic is via the Cauchy-Riemann equations. If our coordinates are (x, y), then let $z = x + iy$.

Definition 1.115 (Holomorphic functions).
A function $f : \mathbb{R}^2 \to \mathbb{C}$ is **holomorphic** if

$$(\partial_x + i\partial_y)f = 0.$$

If we write $f = u + iv$, we can express this as equations for the real and imaginary parts:

$$\begin{cases} \partial_x u - \partial_y v = 0 \\ \partial_y u + \partial_x v = 0 \end{cases}$$

These are the **Cauchy-Riemann equations**. From the perspective of PDEs, this is just one equation.

Denote the operator

$$\bar{\partial} = \partial_x + i\partial_y$$

Sometimes people will use this notation to denote $1/2$ this quantity. Complex differentiation is given by the operator

$$\partial = \partial_x - i\partial_y$$

Our goal is to find the fundamental solution for $\bar{\partial}$. Looking at $\bar{\partial}K = \delta_0$, notice that δ_0 is homogeneous of order -2 and $\bar{\partial}$ reduces order of homogeneity by 1 . So we should look for a K which is homogeneous of order -1. Away from $z = 0, \bar{\partial}K = 0$, so K is holomorphic. So we should look for K of the form $K = \frac{c}{z}$, where c is a constant. This is locally integrable, unlike in 1 dimension. So we can define

$$K(\phi) = c\int_{\mathbb{R}^2} \frac{\phi(z)}{z} dxdy,$$

where we can use $dzd\bar{z}$ instead of $dxdy$. If K is a fundamental solution, $\bar{\partial}K = \delta_0$, so $\partial K(\phi) = \phi(0)$, which gives $K(-\bar{\partial}\phi) = \phi(0)$. Here,

$$\begin{aligned} K(-\bar{\partial}\phi) &= -c\iint_{\mathbb{R}^2} \frac{(\partial_x + i\partial_y)\,\phi(z)}{z} dxdy \\ &= \lim_{\varepsilon \to 0} -c \iint_{\mathbb{R}^2 \setminus B_\varepsilon} (\partial_x + i\partial_y)\,\phi \cdot \frac{1}{z} dxdy \end{aligned}$$

We want to use integration by parts. Using Divergence theorem,

$$= \lim_{\varepsilon \to 0} c \iint_{\mathbb{R}^2 \setminus B_\varepsilon} \phi \cdot \underbrace{(\partial_x + i\partial_y)\frac{1}{z}}_{=0} dxdy - c\int_{\partial B_\varepsilon} (\nu_x + i\nu_y)\,\phi \cdot \frac{1}{z} ds$$

where ν is the inner normal vector to the boundary of B_ε. In particular, $\nu = -\frac{(x,y)}{|z|}$.

$$\begin{aligned} &= \lim_{\varepsilon \to 0} c \int_{\partial B_\varepsilon} \frac{z}{|z|}\phi \cdot \frac{1}{z} ds \\ &= \lim_{\varepsilon \to 0} \frac{c}{\varepsilon} \int_{\partial B_\varepsilon} \phi(z) dz \\ &= 2\pi c \phi(0). \end{aligned}$$

We want $2\pi c = 1$, so we should pick $c = \frac{1}{2\pi}$. Thus, our fundamental solution is

$$K(z) = \frac{1}{2\pi z}.$$

Remark 1.116.
We can rewrite this line integral in a complex fashion, as

$$\int \frac{\phi(z)}{z}\, dz = 2\pi i \phi(0),$$

by the residue theorem. So we have recovered the residue theorem. In essence, the residue theorem is the analogue of the fundamental theorem of calculus for 2 dimensions.

1.15.2 Fundamental solution for the Laplacian

Our next exercise is to find the fundamental solution to $P(\partial) = -\Delta$, where

$$\Delta = \partial_1^2 + \cdots + \partial_n^2$$

Since δ_0 is homogeneous of order $-n$, and $P(\partial)$ will decrease the order of homogeneity by $2, K$ should be homogeneous of order $2-n$. To look for a candidate for a solution, we should look at the symmetries of Δ, in particular invariance with respect to rotations.

If $y = Ax$ is a linear change of variables, then $\frac{\partial}{\partial x_i} = A_{i,j}\frac{\partial}{\partial y_j}$. Then $\Delta = A_{i,j}A_{i,k}\frac{\partial}{\partial y_j}\frac{\partial}{\partial y_k}$ Here, we are using Einstein summation notation, in which the sum is implicit but unwritten. Do we get back Δ in y ? The answer is yes, if

$$A_{i,j}A_{i,k} = I_n \iff A^\top A = I$$

That is, we want A to be orthogonal. Recall that if A is orthogonal,

$$\begin{aligned} \|Ax\|^2 &= \langle Ax, Ax\rangle \\ &= \left\langle x, A^\top Ax\right\rangle \\ &= \langle x, x\rangle \\ &= \|x\|^2 \end{aligned}$$

So we can look for K which is invariant with respect to rigid rotations, i.e. K is a spherically symmetric distribution.

Remark 1.117.
We must be careful with this line of reasoning. We are just hoping that there exists some fundamental solution with this property. Not all fundamental solutions will have this property. For example, if we add x_1 to K, we will still have a fundamental solution, but it will not be radial.

We will guess

$$K = c_n \cdot \frac{1}{|x|^{n-2}}$$

where we will set the case $n = 2$ dimensions aside for now. Observe that

$$\begin{aligned} -\Delta K = \delta_0 &\iff -\Delta K(\phi) = \phi \\ &\iff K(-\Delta\phi) = \phi(0) \\ &\iff \int_{\mathbb{R}^n} -\Delta\phi \frac{1}{|x|^{n-2}} dx = \phi(0) \end{aligned}$$

As before, write this integral as

$$\lim_{\varepsilon \to 0} \int_{\mathbb{R}^n \setminus B_\varepsilon} -\Delta\phi \cdot \frac{1}{|x|^{n-2}}$$

We want to integrate by parts. Here is Green's theorem in this setting:

Theorem 1.118 (Green's theorem for the Laplacian).

$$\int_\Omega \Delta u \cdot v - u \cdot \Delta v dx = \int_{\partial\Omega} \frac{\partial u}{\partial \nu} v - u \frac{\partial v}{\partial \nu} d\sigma$$

Note that ν is the normal derivative.

Proof.

$$\begin{aligned} \int_\Omega \Delta u \cdot v dx &= \int_\Omega \partial_j \partial_j u \cdot v \\ &= -\int_\Omega \partial_j u \cdot \partial_j v dx + \int_{\partial\Omega} \nu_j \partial_j u \cdot v d\sigma \end{aligned}$$

where σ is surface measure on $\partial\Omega$. Observe (for the future) that $\partial_j u \cdot \partial_j v = \nabla u \cdot \nabla v$ in the first term and $\nu_j \partial_j u = \nu \cdot \nabla u := \frac{\partial u}{\partial \nu}$ is the normal derivative in the second term.

$$= \int_\Omega u \cdot \underbrace{\partial_j \partial_j}_{\Delta} v + \int_{\partial\Omega} \frac{\partial u}{\partial \nu} v - u \frac{\partial v}{\partial \nu} d\sigma$$

□

Returning to our computation, we want

$$\phi(0) = \lim_{\varepsilon \to 0} \int_{\mathbb{R}^n \setminus B_\varepsilon} \phi \left(-\Delta \frac{1}{|x|^{n-2}} \right) dx - \int_{\partial B_\varepsilon} \frac{\partial \phi}{\partial \nu} \cdot \frac{1}{|x|^{n-2}} - \phi \frac{\partial}{\partial \nu} \frac{1}{|x|^{n-2}} d\sigma$$

The first integral goes away because $-\Delta \frac{1}{|x|^{n-2}} = 0$. We can see this via a formula for the Laplacian on radial functions: $\Delta F(r) = \left(\partial_r^2 + \frac{n-1}{r} \partial_r \right) F(r)$. This is the chain rule, switching to polar coordinates in n dimensions. The second integral is

$$\int_{\partial B_\varepsilon} \frac{\partial \phi}{\partial \nu} \cdot \frac{1}{|x|^{n-2}} dA = O(\varepsilon) \to 0$$

as $\frac{\partial \phi}{\partial \nu}$ is bounded, $\frac{1}{|x|^{n-2}} = \varepsilon^{2-n}$, and dA has order ε^{n-1}.

The third integral is

$$\int_{\partial B_\varepsilon} \phi \cdot \frac{\partial}{\partial \nu} \frac{1}{|x|^{n-2}} d\nu = \int_{\partial B_\varepsilon} \phi \cdot (n-2) \frac{1}{|x|^{n-1}} d\sigma$$
$$\approx \phi(0) \cdot \frac{n-2}{\varepsilon^{n-1}} \varepsilon^{n-1} a_n,$$

where a_n is the area of the unit sphere.

$$= (n-2) a_n \phi(0)$$

So we need

$$c = c_n = \frac{1}{(n-2)a_n}$$

Theorem 1.119.
If $n \geqslant 3$, then the fundamental solution for $-\Delta$ is

$$K(x) = \frac{1}{(n-2)a_n} \cdot \frac{1}{|x|^{n-2}},$$

where a_n is the area of the unit sphere.

Returning to the 2 dimensional case, we want $K = K(r)$, and outside $K = 0$, we want

$$(\partial_r^2 + \frac{1}{r}\partial_r) K = 0.$$

We can write this as

$$\left(\partial_r + \frac{1}{r}\right) \underbrace{(\partial_r K)}_{L} = 0$$

This tells us that

$$\frac{L'}{L} = -\frac{1}{r}$$

so

$$\log L = -\log r + c$$

which we can write as

$$L = c \cdot \frac{1}{r}$$

Substituting back in for K, we have $\partial_r K = \frac{c}{r}$, which tells us that

$$K = c \ln r + d$$

where d is a constant that we can choose to fit our problem.

What is the constant c? Instead of a computation, we'll do some carefully selected handwaving. Note that

$$\frac{\partial}{\partial \nu} \log r = -\frac{1}{r}$$

so there is no $n-2$. We get the last line of the higher-dimensional computation, but without the $n-2$:

$$c = \frac{1}{a_2} = \frac{1}{2\pi}$$

So

$$K(x) = \frac{1}{2\pi} \ln r$$

where we can add a constant if we wish.

Remark 1.120.
If we think of the Laplacian in 2 dimensions as $\Delta = \partial\bar{\partial}$, then the fundamental solutions follow

$$K_{-\Delta} = K_{\partial} * K_{\bar{\partial}} = \frac{1}{z} * \frac{1}{\bar{z}}$$

We get a divergent integral, but with a proper renormalization, we can make sense of this.

1.16 Introduction to the Fourier Transform

Date: Oct 19, 2021

1.16.1 Motivation: diagonalization for differential operators

We would like to have a better way to think about fundamental solutions to PDEs. Here is an analogy for the Fourier transform. Suppose we have a symmetric matrix in $\mathbb{R}^n$. Then A is diagonalizable, with orthonormal eigenvectors $u_1, \dots, u_n$. If you want to better represent your matrix, you can change coordinates to this basis, or you can express an arbitrary vector with $u = c_1 u_1 + \cdots c_n u_n$, where $c_j = u \cdot u_j$. If you have two (or a family of) commuting matrices, you can find an orthonormal basis of eigenvectors for both (or all) matrices simultaneously.

If we have PDEs with constant coefficients, then the operators $P(\partial), Q(\partial), \dots$ are all commuting operators. Can we find a common eigenbasis of functions? Here are some candidates for eigenfunctions $e^{ix\cdot\xi}$, where the i is there to make sure that these don't blow up at ∞. Then

$$P(\partial)e^{ix\cdot\xi} = P(i\xi)e^{ix\cdot\xi}$$

so these exponentials naively serve as eigenfunctions for these operators with eigenvalues $P(i\xi)$. Here, we don't always have real eigenvalues, but we have complex eigenvalues.

Here are some issues:

- Are these functions orthogonal?
  Consider the Hilbert space $L^2(\mathbb{R}^n) = \{u : \mathbb{R}^n \to \mathbb{R} | \int_{\mathbb{R}^n} |u|^2 dx < \infty\}$. If we consider the $L^2(\mathbb{R}^n)$ inner product, $u \cdot v = \int_{RR^n} u(x)v(x)\,dx$ (with v replaced by $\bar{v}$ for complex functions), are these orthonormal? In fact, $e^{ix\cdot\xi} \notin L^2$, so we cannot properly analyze

$$\int_{\mathbb{R}^n} e^{ix\cdot\xi_1} e^{-ix\cdot\xi_2}\,dx.$$

- For our diagonalization, we have uncountable many eigenvectors. $L^2(\mathbb{R}^n)$ is a separable Hilbert space with a countable orthonormal basis. So we have too many functions.

However, we can think of $e^{ix\cdot\xi}$ as generalized eigenfunctions. We can still ask the question: Given $f \in L^2(\mathbb{R}^n)$, can we write it as a superposition as $e^{ix\cdot\xi}$? That is, can we write

$$f(x) = \int e^{ix\cdot\xi} c(\xi) d\xi?$$

If we disregard the above issues, can we still recover an identity like $c_j = u \cdot u_j$ as before? We may want to try

$$c(\xi) = \int f(x) e^{-ix\xi} dx$$

But since we have trouble normalizing the eigenfunctions, should there be a normalization constant in front? If we can achieve such a representation, then we get a lot out of it:

$$P(\partial)f = \int e^{ix\cdot\xi} c(\xi) P(i\xi) d\xi$$

So the map $f \mapsto P(\partial)f$ just acts diagonally on this basis: $c(\xi) \mapsto P(i\xi) \cdot c(\xi)$.

1.16.2 Properties of the Fourier transform

We will use the notation $D_j = \frac{1}{i}\partial_j$, so that $D_j e^{ix\cdot\xi} = \xi_j e^{ix\cdot\xi}$. So we will think of $P(D)$ instead of $P(\partial)$. In this notation, $P(D)e^{ix\cdot\xi} = P(\xi)e^{ix\cdot\xi}$, and we call $P(\xi)$ the **symbol** of P.

Example 1.121.
If $P(x,D)=\sum_\alpha c_\alpha(x)D^\alpha$, then the symbol is $P(x,\xi)=\sum_\alpha c_\alpha(x)\xi^\alpha$.

Definition 1.122 (Fourier transform).
The Fourier transform of a function f is

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

Our goal is to show that

$$f(x)=\frac{1}{(2\pi)^{n/2}}\int_{\mathbb{R}^n}e^{ix\cdot\xi}\widehat{f}(\xi)d\xi$$

For what f is $\widehat{f}$ well-defined? The integral is absolutely convergent if $f\in L^1$, i.e. $\int|f|<\infty$. We will not use L^1 functions much in our context. If we have $f\in L^1$, then

$$|\widehat{f}(\xi)|\leqslant\frac{1}{(2\pi)^{n/2}}\|f\|_{L^1}$$

which we can write as

$$\|\widehat{f}\|_{L^\infty}\leqslant\frac{1}{(2\pi)^{n/2}}\|f\|_{L^1}$$

The problem is that we want to be able to undo the Fourier transform, and for L^∞ functions, the Fourier transform is not well-defined.

What about the Fourier transform on test functions? If $f\in\mathcal{D}$, then $\widehat{f}\in\mathcal{E}$, so there is no compact support. But if we have $f\in\mathcal{E}$, then $\widehat{f}$ does not exist, since the integral may not converge. It seems that $\mathcal{D}$ is too small, and $\mathcal{E}$ is too large. What should be our intermediate space where $\mathcal{F}$ acts? We will use the Schwartz space $\mathcal{S}$ For $u\in\mathcal{S}$, we want the derivatives to not only be bounded but have decay at infinity.

Definition 1.123 (Schwartz space).
The Schwartz space is the space of $\mathrm{C}^\infty(\mathbb{R}^n)$ functions which are rapidly decreasing, in the sense that

$$\left|x^\alpha\partial^\beta u\right|\leqslant c_{\alpha,\beta}$$

for all $\alpha,\beta\in\mathbb{N}^n$.

The Schwartz space $\mathcal{S}$ is a locally convex space with seminorms

$$p_{\alpha,\beta}(u)=\left\|x^\alpha\partial^\beta u\right\|_{L^\infty}.$$

Theorem 1.124.
The Fourier transfomr is $\mathcal{F}:\mathcal{S}\to\mathcal{S}$, and the inverse $\mathcal{F}^{-1}:\mathcal{S}\to\mathcal{S}$.

We have not proven that $\left(\mathcal{F}^{-1}f\right)(\xi)=\frac{1}{(2\pi)^{n/2}}\int_{\mathbb{R}^n}e^{ix\cdot\xi}\widehat{f}dx$ gives the inverse, but we will call it the inverse for now. How do we prove this theorem?

Observe that in the expression $x^\alpha\partial^\beta$, the order of x^α and ∂^β does not matter. How do ∂,x interact with the Fourier transform?

Proposition 1.125.
For $f\in\mathcal{S}$, $\partial_j\hat{f}=-i\widehat{x_jf}$.

Proof.

$$\begin{aligned}\partial_j \widehat{f}(\xi) &= \frac{1}{(2\pi)^{n/2}} \int e^{-ix\cdot\xi} f(x)\,(-ix_j)\,dx \\ &= -i\widehat{x_j f}\end{aligned}$$

□ □

Proposition 1.126.
For $f \in \mathcal{S}$, $\xi\hat{f} = -i\widehat{\partial_x f}$.

Proof.

$$\xi_j \widehat{f}(\xi) = \frac{1}{(2\pi)^{n/2}} \int e^{-ix\cdot\xi} f(x)\xi_j dx$$

Use integration by parts.

$$\begin{aligned}&= \frac{1}{(2\pi)^{n/2}} \int i\frac{\partial}{\partial x_j}\left(e^{-ix\cdot\xi}\right) f(x)dx \\ &= \frac{1}{(2\pi)^{n/2}} \int -i\left(e^{-ix\cdot\xi}\right) f(x)dx\end{aligned}$$

□ □

Therefore, multiplication by x on the physical side is differentiation on the Fourier side, and multiplication by ξ on the Fourier side is differentiation on the physical side.

Now we can give proof of Thm 15.4.

Proof.
If $f \in \mathcal{S}$, then (using $\beta = 0$ and $|\alpha| \leqslant N$ for $N > n$)

$$|f(x)| \leqslant \frac{c_N}{(1+|x|)^N} \in L^1$$

So $\|\widehat{f}\|_{L^\infty} \leqslant c\|f\|_{L^1}$ Together, our propositions give us

$$\xi^\alpha \partial_\xi^\beta \widehat{f} = (-i)^{|\alpha|+|\beta|}\widehat{\partial_x^\alpha x^\beta f}$$

Here, we have

$$\left\|f^\alpha \partial_\xi^\beta \widehat{f}\right\|_{L^\infty} \leqslant \left\|\partial_x^\alpha x^\beta f\right\|_{L^1}$$

If $f \in \mathcal{S}$, then $\partial_x^\alpha x^\beta f \in \mathcal{S} \subseteq L^1$. So the right hand side is finite, controlled by finitely many of our Schwartz seminorms. □ □

Example 1.127 (Fourier transform of a Gaussian).
Suppose $f(x) = e^{-x^2/2}$. What is $\hat{f}$?

$$\begin{aligned}\widehat{f}(\xi) &= \frac{1}{(2\pi)^{n/2}} e^{-x^2/2} e^{-ix\xi} dx \\ &= \frac{1}{(2\pi)^{n/2}} e^{-\xi^2/2} \int e^{-(x+i\xi)^2/2} dx\end{aligned}$$

How do we deal with this integral? If we write $z = x = i\xi$, we are doing a complex integral on the

curve Γ_ξ. So we get

$$\begin{aligned}\widehat{f}(\xi) &= \frac{1}{(2\pi)^{n/2}} \int_{\Gamma_\xi} e^{-z^2/2} dz \\ &= e^{-\xi^2/2} \frac{1}{(2\pi)^{n/2}} \int_{\Gamma_0} e^{-z^2/2} dz \\ &= e^{-\xi^2/2} \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-x^2/2} dx\end{aligned}$$

We can recall $\int e^{-x^2} dx = \sqrt{\pi}$, so a change of variables gives

$$= e^{-\xi^2/2}$$

So we have seen that

$$\mathcal{F}\left(e^{-x^2/2}\right) = e^{-\xi^2/2}$$

In general, what is $\mathcal{F}\left(e^{-\lambda x^2/2}\right)$? Here is how the Fourier transform behaves under scaling:

Proposition 1.128.
For $f \in \mathcal{S}$,

$$\widehat{f(\mu \cdot)} = \frac{1}{\mu^n} \widehat{f}(\cdot/\mu)$$

Proof.

$$\mathcal{F}f(\mu x) = \int e^{-ix\cdot\xi} f(\mu x) dx$$

Make the change of variables $y = \mu x$.

$$\begin{aligned}&= \frac{1}{\mu^n} \int e^{-iy\cdot\xi/\mu} f(y) dy \\ &= \frac{1}{\mu^n} \hat{f}(\xi/\mu)\end{aligned}$$

□ □

Remark 1.129.
You might call $f(ux)$ an L^∞ scaling, whereas $\frac{1}{u^n}\hat{f}(\xi/\mu)$ is an L^1 scaling.

Example 1.130.
Setting $\mu = \sqrt{\lambda}$,

$$\mathcal{F}\left(e^{-\lambda x^2/2}\right) = \frac{1}{\lambda^{n/2}} e^{-\xi^2/(2\lambda)}$$

We will work toward the following Fourier inversion theorem:

Theorem 1.131.
$\mathcal{F}^{-1}\mathcal{F} = \mathcal{F}\mathcal{F}^{-1} = I$ in $\mathcal{S}$.

Remark 1.132.
You can think of $\mathcal{F}\mathcal{F}^{-1}$ as the complex conjugate of $\mathcal{F}^{-1}\mathcal{F}$.

1.17 Fourier Inversion, Plancherel's Theorem and Temper Distributions

1.17.1 Fourier inversion

Last time, we introduced the Fourier transform

$$\mathcal{F}u(\xi) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-ix\cdot\xi} u(x)dx$$

We had an "inverse"

$$\mathcal{F}^{-1}v(x) = \frac{1}{(2\pi)^{1/2}} \int_{\mathbb{R}^n} e^{ix\cdot\xi} v(\xi)d\xi$$

Both $\mathcal{F}$ and $\mathcal{F}'$ are functions from $\mathcal{S} \to \mathcal{S}$, where $\mathcal{S} = \left\{\varphi : \left|x^\alpha \partial^\beta \varphi\right| \leqslant c_{\alpha,\beta}\right\}$ is the Schwartz space.

Theorem 1.133.
$\mathcal{F}^{-1}\mathcal{F} = I$ on $\mathcal{S}$.

Proof.
Let's first try a brute-force approach and see what happens.

$$\begin{aligned}
\mathcal{F}^{-1}\mathcal{F}u &= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{ix\cdot\xi} \widehat{u}(\xi)d\xi \\
&= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{ix\cdot\xi} \int_{\mathbb{R}^n} e^{-i\xi\cdot y} u(y)dyd\xi \\
&\overset{?}{=} \frac{1}{(2\pi)^n} \iint e^{i(x-y)\cdot\xi} \mu(y)d\xi dy
\end{aligned}$$

We know $\widehat{u}$ has rapid decay, so the first integral is well-defined. But it is not clear how we can integrate here. The $d\xi$ integral should evaluate to be $\delta_{x=y}$ in some way. Here is what we actually do:

$$= \lim_{\varepsilon\to 0} \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{i(x-y)\cdot\xi} e^{-\frac{\varepsilon}{2}\xi^2}\, d\xi \int_{\mathbb{R}^n} u(y)dy$$

Now we can legitimately apply Fubini's theorem.

$$\begin{aligned}
&= \lim_{\varepsilon\to 0} \frac{1}{(2\pi)^n} \iint u(y) e^{i(x-y)\cdot\xi} e^{-\frac{\varepsilon}{2}\xi^2} d\xi dy \\
&= \lim_{\varepsilon\to 0} \int u(y) \frac{1}{(2\pi)^n} e^{-\frac{(x-y)^2}{2\varepsilon}} \varepsilon^{-n/2} dy \\
&= \lim_{\varepsilon\to 0} u * \varphi_\varepsilon \\
&= u
\end{aligned}$$

where

$$\varphi_\varepsilon(y) = \frac{1}{(2\pi)^n} e^{-\frac{y^2}{2\varepsilon}} \frac{1}{\varepsilon^{n/2}} \xrightarrow{\varepsilon\to 0} \delta_0$$

□ □

1.17.2 Isometry properties of $\mathcal{F}$ on L^2

Now let's shift our attention to L^2, with inner product $\langle u, v\rangle = \int u\bar{v}\, dx$.

Proposition 1.134.
The Fourier transform is unitary on L^2. That is

$$\mathcal{F}^* = \mathcal{F}^{-1}, \quad \left(\mathcal{F}^{-1}\right)^* = \mathcal{F}$$

Proof.

$$\begin{aligned}
\langle \mathcal{F}, uv \rangle &= \iint e^{-ix\xi} u(x) dx \bar{v}(\xi) d\xi \\
&= \iint e^{-ix\cdot\xi} \bar{v}(\xi) d\xi u(x) dx \\
&= \iint \overline{e^{ix\xi} v(\xi)} d\xi u(x) dx \\
&= \left\langle u, \mathcal{F}^{-1} v \right\rangle
\end{aligned}$$

□ □

Theorem 1.135.
$\mathcal{F} : \mathcal{S} \to \mathcal{S}$ is an L^2-isometry.

Proof.
If we set $u = v$, we get

$$\|u\|_{L^2}^2 = \int |u|^2 dx = \|\mathcal{F}u\|_{L^2}^2$$

□ □

We can use this to extend $\mathcal{F}$ to $L^2(\mathbb{R}^n)$ by density. If $u \in L^2$, find $u_n \in \mathcal{S}$ such that $u_n \to u$ in L^2. Then u_n is Cauchy in L^2, so $\mathcal{F}u_n$ is Cauchy in L^2. So $\lim_{n\to\infty} \mathcal{F}u_n =: \mathcal{F}u$.

Remark 1.136.
The Hahn-Banach theorem says that we can extend operators that are densely defined, but in general, there is no guarantee of uniqueness.

However, it is not immediately clear that we can do this approximation of elements of L^2 by elements in $\mathcal{S}$.

Proposition 1.137.
If $u \in L^2$, then there exists $u_n \in \mathcal{D}$ such that $u_n \to u$ in L^2. Note that this means $\mathcal{D}$ is dense in L^2.

Proof.
Step 1: Approximate u by compactly supported functions $u = \lim_{n\to\infty} u_n := u1_{\{|x|\leqslant n\}}$.

Step 2: Regularize $u = \lim_{\varepsilon\to 0} u * \varphi_\varepsilon$. Here, $\varphi \in \mathcal{D}$ with $\int \varphi = 2$, and $\varphi_\varepsilon = \varepsilon^{-n}\varphi(x/\varepsilon)$ so $\varphi_\varepsilon \to \delta_0$ as $\varepsilon \to 0$. So $u * \varphi_\varepsilon \to u$ in $\mathcal{D}'$ if $u \in \mathcal{D}'$ and in L^2 if $u \in L^2$.

□ □

So we get the following theorem:

Theorem 1.138 (Plancherel).
$\mathcal{F} : L^2 \to L^2$ is an isometry.

1.17.3 Temperate distributions

Can we extend $\mathcal{F}$ to any larger spaces? First, we will talk about the Fourier transform as a map $\mathcal{F} : \mathcal{S}' \to \mathcal{S}'$

Definition 1.139 (Temperate distributions).
$\mathcal{S}'$, the space of temperate distributions, is the space of distributions which extend to continuous linear functionals on $\mathcal{S}$.

Given $u \in \mathcal{D}'$, $u \in \mathcal{S}'$ if there is a constant c such that for $R\varphi \in \mathcal{S}$,

$$|u(\varphi)| \leqslant c \sum_{\text{finite}} p_{\alpha,\beta}(\varphi), \quad p_{\alpha,\beta}(\varphi) = \sup \left|x^\alpha \partial^\beta \varphi\right|.$$

Here is how we extend Fourier transform and inverse transforms to $\mathcal{S}'$: For $u, v \in \mathcal{S}$,

$$\langle \mathcal{F}u, v\rangle = \left\langle u, \mathcal{F}^{-1}v\right\rangle$$

so we have $\mathcal{F}u(\bar{v}) = u\left(\overline{\mathcal{F}^{-1}v}\right)$. Replacing v by $\bar{v}$ give $\mathcal{F}u(v) = u(\mathcal{F}v)$, where $u \in \mathcal{S}'$ and $\mathcal{F}v \in \mathcal{S}$. So we can define

$$\mathcal{F}u = u(\mathcal{F}v)$$

for $u \in \mathcal{S}', v \in \mathcal{S}$. $\mathcal{S} \subseteq \mathcal{E}$, so $\mathcal{E}' \subseteq \mathcal{S}'$. If $u \in \mathcal{E}'$ (is compactly supported), then

$$\mathcal{F}u(\xi) = u\left(\frac{1}{(2\pi)^{n/2}} e^{-x\xi}\right)$$

Note $u\left(\frac{1}{(2\pi)^{n/2}} e^{-x\xi}\right)$ is a function of x. So we see that $\mathcal{F} : \mathcal{E}' \to \mathcal{E}$. The moral here is that "$\mathcal{F}$ interchanges decay and regularity."

1.17.4 Examples of temperate distributions

When is a function a temperate distribution? If $u \in \mathcal{S}'$ and $\varphi \in \mathcal{S}$,

$$u(\varphi) := \int u(x)\varphi(x)dx$$

where $\varphi(x)$ is rapidly decreasing. So if $|u(x)| \leqslant c\left(1 + |x|^N\right)$, then the integral is convergent.

Example 1.140.
All rational functions are temperate distributions. You should not get the idea that these are all the temperate distributions.

Example 1.141.
Consider

$$u(x) = e^x \cos e^x$$

Think of $u = \frac{\partial}{\partial x} \sin e^x = \partial_x f$. Then

$$u(\varphi) = -f\left(\partial_x \varphi\right)$$

where $\partial_x \varphi \in \mathcal{S}$ if $\varphi \in \mathcal{S}$. So a temperate distribution may not have much decay if it has enough oscillation, and there is a delicate balance between the two. Here, if we have $x, \partial : \mathcal{S} \to \mathcal{S}$, we have extended $x, \partial : \mathcal{S}' \to \mathcal{S}'$.

1.17.5 The Fourier transform of δ_0 and H

What is $\widehat{\delta_0}$?

$$\widehat{\delta_0}(\xi) = \delta_0\left(\frac{1}{(2\pi)^{n/2}} e^{ix\cdot\xi}\right) = \frac{1}{(2\pi)^{n/2}}$$

Remark 1.142.
People will often change the normalization constant in the Fourier transform to get $\widehat{\delta}_0 = 1$. So people will also replace $e^{ix\cdot\xi}$ with $e^{2\pi ix\cdot\xi}$. This is useful if you want to deal with Fourier series or if you want to make a distinction between the $\mathbb{R}^n$ of the input and the $\mathbb{R}^n$ of the ourput. These are actually the same space because $\mathbb{R}^n$ is the cotangent space fo $\mathbb{R}^n$. For more general spaces, the Fourier transform will not have the same input and output domain. We will not need to worry about this for our PDEs.

In 1 dimension, we have $\partial_x H = \delta_0$. Then

$$\mathcal{F}(\partial_x H) = \mathcal{F}(\delta_0)$$

which tells us that $-i\xi\mathcal{F}(H) = \frac{1}{(2\pi)^{n/2}}$. So we get that

$$\widehat{H} = \frac{i}{(2\pi)^{n/2}} \cdot \frac{1}{\xi}$$

Take u compactly ussported in $[0, \infty)$. Then

$$\widehat{u}(\xi) = \int e^{-ix\cdot\xi} u(x) dx$$

Switch to complex numbers $\xi + i\zeta$. This integral ecomes

$$\int e^{-ix\xi + x\zeta} u(x) dx$$

If $\zeta < 0$, we have exponential decay for $x > 0$. So $\widehat{u}(\xi)$ extends to a holomorphic function in $\{\operatorname{Im} z \leqslant 0\}$. In this picture, we can think of

$$\widehat{H} = \frac{i}{(2\pi)^{n/2}} \cdot \frac{1}{\xi - i0}$$

We can also look at

$$\widehat{H - 1} = \frac{i}{(2\pi)^{n/2}} \cdot \frac{1}{\xi + i0}$$

So if we take the average, we get

$$H - \frac{1}{2} = \frac{i}{(2\pi)^{n/2}} \mathrm{PV} \frac{1}{\xi}$$

1.18 Using the Fourier Transform to Find Fundamental Solutions

1.18.1 The Paley-Wiener theorem and the Fourier transform of even and odd functions

We have been looking at the Fourier transform

$$\widehat{u}(\xi) = \frac{1}{(2\pi)^{n/2}} \int e^{-ix\cdot\xi} u(x)dx$$

We initially defined $\mathcal{F} : \mathcal{S} \to \mathcal{S}$, but we can also define it $L^2 \to L^2$ (with the isometry property) and $\mathcal{S}' \to \mathcal{S}'$. We have also seen that $\mathcal{F} : L^1 \to L^\infty$.

Last time, we also saw that

$$\widehat{H} = \frac{i}{(2\pi)^{n/2}(\xi - i0)}$$

If $u \in \mathcal{S}'$ with supp $u \subseteq [0, \infty)$, then $\widehat{u}$ has a holomorphic extension to $\{\operatorname{Im} z \leqslant 0\}$. If u is a measure, then $\widehat{u}$ is bounded in $\{\operatorname{Im} z \leqslant 0\}$. This leads us to the following property. First, let's generalize this statement.

Suppose supp $u \subseteq [a, \infty)$. Then

$$\widehat{u}(\xi + i\zeta) = \int e^{ix\xi + x\zeta} u(x)dx$$

so

$$|\widehat{u}(\xi + i\zeta)| \leqslant c \cdot e^{a\zeta}$$

The best we can hope for is a bound of the form $e^{a\zeta}|\xi|^N$.

Theorem 1.143 (Paley-Wiener).
$u \in \mathcal{S}'$ has supp $u \subset [a, \infty)$ if and only if $\hat{u}$ has a holomorphic extension to the lower half-plane such that

$$|\hat{u}(z)| \leqslant e^{-a \operatorname{Im} z}|z|^N.$$

Remark 1.144.
There is a Paley-Wiener theorem in higher dimensions. If supp $u \subseteq K$ for some compact K, then $\widehat{u}(\xi)$ is defined for $\xi \in \mathbb{C}^n$. Instead of getting the support of u as K in the other direction, we get the convex hull of K.

We can think of the $e^{-ix\cdot\xi}$ in the Fourier transform as $\cos(-x \cdot \xi) + i \sin(-x \cdot \xi)$.

- If u is real and even, hence $\hat{u}$ is real and even.
- If u is real and odd, then $\hat{u}$ is imaginary and odd.
- If u is imaginary and even, then $\hat{u}$ is imaginary and even.
- If u is imaginary and odd, then $\hat{u}$ is real and odd.

1.18.2 Using the Fourier transform to find fundamental solutions

Suppose we have a constant partial differential operator $P(\partial)$, and we want to compute a fundamental solution $P(\partial)K = \delta_0$. Let $D = -i\partial$ and rewrite $P(\partial)$ as $P(D)$. Taking the Fourier transform gives

$$P(\xi)\hat{K} = \frac{1}{(2\pi)^{n/2}}1.$$

This tells us that

$$\hat{K} = \frac{1}{(2\pi)^{n/2}} \frac{1}{P(\xi)}.$$

So we can invert the Fourier transform to get K :

$$K = \frac{1}{(2\pi)^{n/2}} \mathcal{F}^{-1}\left(\frac{1}{P(\xi)}\right)$$

Here are some issues:

- $P(\xi)$ may have zeros.
- If P has zeroes, hen $1/P$ is not uniquely determined as a distribution.
- This procedure only gives fundamental solutions which are temperate distributions.

The easy case is when $P(\xi) \neq 0$ for any $\xi \in \mathbb{R}^n$. Then $\frac{1}{p} \in \mathcal{S}'$, so this computation is justified.

Example 1.145.
Suppose $P = -\partial_x^2 + 1 = D_x^2 + 1$. Then $P(\xi) = (1 + \xi^2)$. So we compute

$$K(x) = \mathcal{F}^{-1}\left(\frac{1}{1+\xi^2}\right)$$

This $K(x)$ is real and even. We are looking at

$$\int_{\mathbb{R}} \frac{1}{\xi^2 + 1} e^{ix\xi} d\xi$$

This integral has a pole at i and a pole at $-i$. However, we can expend this using partial fractions:

$$\frac{1}{1+\xi^2} = \frac{i}{2}\frac{1}{\xi + i} - \frac{i}{2}\frac{1}{\xi - i}$$

where the first term is holomorphic if $\operatorname{Im}\zeta > 0$ and the second is holomorphic if $\operatorname{Im}\zeta < 0$. So the Paley-Wiener theorem tells us that the first one will have an inverse Fourier transform supported in $(-\infty, 0]$, and the second one will have an inverse Fourier transform supported in $[0, \infty)$. If $x < 0$, we can use complex analysis to say

$$\int_{\mathbb{R}} \frac{1}{\xi + i} e^{ix\xi} d\xi = \text{ Residue at } i = e^x$$

A similar computation for $x > 0$ suggests that we should get

$$\int_{\mathbb{R}} \frac{1}{1+\xi^2} e^{ix\xi} d\xi = ce^{-|x|}$$

In general, if K is a fundamental solution, then so will be $K + K_0$, where K_0 solves the homogeneous equation $P(\partial)K_0 = 0$. In this case, our general solution is $K = ce^{|x|} + c_1 e^x + c_2 e^{-x}$. We did not get these latter two terms before because they are not temperate distributions.

Example 1.146.
If $P = -\Delta + 1$, then $P(\xi) = \xi^2 + 1$ in $\mathbb{R}^n$. Then

$$K = \mathcal{F}^{-1}\left(\frac{1}{1+\xi^2}\right)$$

gives the unique temperate fundamental solution. Note that $e^{ix\cdot\xi}$ is a solution iff $1 + \xi^2 = 0$ In 3

dimensions, this is $K(x) = e^{-|x|}\frac{1}{|x|}$.

Example 1.147.
Let $P = -\Delta$, so $P(\xi) = \xi^2$. Then $K = \frac{1}{\xi^2}$ is locally integrable in $\mathbb{R}^n$ if $n \geqslant 3$. So if $n \geqslant 3$, we get that $K \in \mathcal{S}'$ is a homogeneous temperate distribution. Since $\frac{1}{\xi^2}$ is homogeneous of order $-2, K = \mathcal{F}^{-1}\left(\frac{1}{\xi^2}\right)$ will be homogeneous of order $2 - n$.

Proposition 1.148.
If u is homogeneous of order s, then $\hat{u}$ is homogeneous of order $-n - s$.

The example to keep in mind to make sure your numbers are right is $\widehat{\delta} = \frac{1}{(2\pi)^{n/2}}$. The Dirac mass is homogeneous of order $-n$, whereas this constant function is homogeneous of order 0.

Example 1.149.
If $P = -\Delta$ with $n = 2$, perform the same computation as before, but interpret $\frac{1}{\xi^2}$ as a distribution:

$$\frac{1}{|\xi|^2}(\varphi) = \lim_{\varepsilon \to 0} \int_{\mathbb{R}^2 \setminus B(0,\varepsilon)} \frac{\varphi(\xi)}{|\xi|^2} d\xi - \varphi(0) \ln \varepsilon$$

so we pay a price of log, which makes us lose the homogeneity property.

Example 1.150.
Example 1.5. Suppose $P(\xi) = A\xi \cdot \xi$, where A is a positive deifnite matrix. This is a second order, elliptic, constant coefficient PDE with $P = a^{i,j}\partial_i \partial_j$. We can transform $A \to \mathrm{Id}$ by a linear fransformation. Let $x = By$, so $x \cdot \xi = By \cdot \xi = y \cdot B^\top \xi$. If we carry out the computation, we end up with

$$K = \frac{1}{(A^{-1}x \cdot x)^{(n-2)/2}}$$

Hormander's book extensively discusses how the Fourier transform behaves under linear changes of coordinates.

1.18.3 Fundamental solution of the heat equation

Recall the heat equation

$$(\partial_t - \Delta)\, u = f$$

We think of u as the temperature of an infinite solid and f as describing the heat sources. This is also called the diffusion equation, since we can, for example, interpret $u(t, x)$ as a local concentration of salt in the water of an ocean. In probability theory, the heat equation has connections to Brownian motion, where we let a particle move randomly at every time, independently of the movement at other times.

Our Fourier variables will be ξ (corresponding to x) and τ (corresponding to t). We can write our operator as

$$\partial_t - \Delta = iD_t + D_x^2,$$

so

$$P(\xi, \tau) = i\tau + \xi^2$$

which vanishes only at $\tau = 0, \xi = 0$. Is $\frac{1}{i\tau + \xi^2} \in L^1_{\mathrm{loc}}$? Yes! The $1/\tau$ increases the local integrability of this expression, so we will not need to make a distinction between the cases $n = 2$ and $n \geqslant 3$. We want to calculate

$$\mathcal{F}^{-1}\left(\frac{1}{i\tau + \xi^2}\right)$$

First integrate in τ : We have a pole at $\tau = i\xi^2$. This pole is in the upper half plane, so $\mathcal{F}_\tau^{-1}\left(\frac{1}{i\tau+\xi^2}\right)$ is supported where $t > 0$. This says that the evolution of heat is well-defined in the future, rather than in the past. We conclude that

$$\mathcal{F}_\tau^{-1}\left(\frac{1}{i\tau+\xi^2}\right) = ce^{-t\xi^2}1_{\{t\geqslant 0\}}$$

for some constant c. Then we can calculate

$$\mathcal{F}^{-1}\left(\frac{1}{i\tau+\xi^2}\right) = \frac{1}{(4\pi t)^{n/2}}e^{-\frac{x^2}{4t}}1_{\{t\geqslant 0\}}$$

Here is another approach. We can try to solve

$$\begin{cases} (\partial_t - \Delta)\, u = 0 \\ u(0) = \delta_0 \end{cases}$$

Take the Fourier transform in x to get

$$\begin{cases} (\partial_t + \xi^2)\, \widehat{u} = 0 \\ \widehat{u}(0) = \frac{1}{(2\pi)^{n/2}} \end{cases}$$

This gives

$$\widehat{u} = \frac{1}{(2\pi)^{n/2}}e^{-t\xi^2}$$

So we get the same result. For $t > 0$, we can consider

$$\begin{cases} (\partial_t - \Delta)\, u = 0 \\ u(0) = u_0 \end{cases}$$

Extend u to

$$\widetilde{u} = \begin{cases} u & t > 0 \\ 0 & t < 0 \end{cases}$$

Then

$$(\partial_t - \Delta)\, \widetilde{u} = u_0(x)\delta_{t=0}$$

Here, $u_0 = \delta_{x=0}$, so $u_0\delta_{t=0} = \delta_{(0,0)}$.

1.19 Schrödinger Equation, the Uncertainty Principle, and Oscillatory Integrals

1.19.1 Fundamental solution of the Schrödinger equation

Recall the heat equation

$$(\partial_t - \Delta)\, u = f \quad \text{in } \mathbb{R}_t \times \mathbb{R}^n_x$$

This has fundamental solution

$$K(t,x) = \frac{1}{(4\pi t)^{n/2}} e^{-x^2/(4t)} \mathbb{1}_{\{t \geqslant 0\}}$$

This is the unique temperate distribution for the heat equation. We also have the Schrödinger equation

$$(i\partial_t + \Delta)\, u = f \quad \text{in } \mathbb{R} \times \mathbb{R}^n$$

Unlike the heat equation, this equation fundamentally has complex-valued solutions. This is the fundamental PDE in quantum mechanics, where $u(t)$ is interpreted as the state of a particle at time t in a probabilistic sense as follows: $\|u\|_{L^2} = 1$, and $|u|^2$ is viewed as a probability distribution. In particular,

$$\mathbb{P}(p \in E) = \int_E |u|^2 dx$$

where p can be the position of a particle. In this picture, the Fourier transform also plays a role. Here, $|\widehat{u}|^2$ is the probability density of the velocity of the particle. Plancherel's theorem tells us that $\|\widehat{u}\|_{L^2} = 1$, as well.

Let $P(\tau, \xi) = \tau - \xi^2$. Then the fundamental solution to the Schrödinger equation should be $K = \mathcal{F}^{-1}\left(\frac{1}{\tau - \xi^2}\right)$. The issue is that $\tau - \xi^2$ has an entire parabola worth of zeroes. How do we think of $\frac{1}{\tau - \xi^2}$ as a distribution? If we just view this as a distribution in the variable τ, this is like the distribution $\frac{1}{x}$, which gives a few different ways to think of it:

$$\frac{1}{\tau - \xi^2 - i0}, \quad \frac{1}{\tau - \xi^2 + i0}, \quad \mathrm{PV}\frac{1}{\tau - \xi^2}$$

Note that these first two solutions indicate that the Schrödinger equation, unlike the heat equation, can be run backwards in time. How do we pick one of these options? We might want to look for a solution that looks like it's moving forward in time: $\operatorname{supp} K \subseteq \{t \geqslant 0\}$. This implies that $\widehat{K}$ should have a holomorphic extension in the lower half-plane. Then our forward fundamental solution is

$$K(t,x) = \mathcal{F}^{-1}\left(\frac{1}{\tau - \xi^2 - i0}\right)$$

First, we will take the Fourier transform with respect to τ. That $\xi = 0$, this gives $H(t)$. Recall that $\mathcal{F}^{-1}\delta_0 = 1$, and $\mathcal{F}^{-1}\delta_{\xi_0} = e^{ix\xi_0}$. This is a general rule for the Fourier transform of the translation of a distirbution, so when $\xi \neq 0$, we get $K(t,\xi) = H(t)e^{-i\xi^2 t}$.

Alternatively, take only a spatial Fourier transform of the Schrödinger equation

$$\begin{cases} (i\partial_t + \Delta)\, u = 0 \\ u(0) = u_0 = \delta_0(u) \end{cases}$$

to get

$$\begin{cases} (i\partial_t + \xi^2)\, \widehat{u}(\xi) = 0 \\ \widehat{u}(0) = 1 \end{cases}$$

This gives $\widehat{u}(\xi) = e^{it\xi^2}$, so $u = \mathcal{F}^{-1}\left(e^{-it\xi^2}\right)$. Recall that $\mathcal{F}(e^{-\xi^2/2}) = e^{-x^2/2}$. and more generally that $\mathcal{F}(e^{-\lambda\xi^2/2}) = \frac{1}{\lambda^{n/2}} e^{-x^2/(2\lambda)}$ for $\lambda \in \mathbb{R}^+$.

Extend this to complex λ. For what complex λ is $\frac{1}{\lambda^{n/2}} e^{-x^2/(2\lambda)}$ a temperate distribution? This is the right half plane $\{\lambda : \operatorname{Re}\lambda \geqslant 0\}$. For $\operatorname{Re}\lambda > 0$, the function $e^{-\lambda\xi^2/2}$ is analytic with values in $\mathcal{S}$. This

tells us that its Fourier transform is analytic for $\operatorname{Re}\lambda > 0$ and we can uniquely extend it to an analytic function on $\{\operatorname{Re}\lambda > 0\}$. What about when $\operatorname{Re}\lambda = 0$? As $\lambda = it + \varepsilon \to it, e^{-(it+\varepsilon)\xi/2} \to e^{-it\xi^2/2}$ in $\mathcal{S}$, i.e. in the topology of temperate distributions. So the Fourier transforms converge in the same sense. Thus, we get fundamental solution

$$K(t,x) = \frac{1}{(4\pi it)^{n/2}} e^{ix^2/(4t)} 1_{\{t \geqslant 0\}}$$

Remark 1.151.
Note that $\widehat{u}(t,\xi) = e^{it\xi}\widehat{u}_0(\xi)$, which means that

$$|\widehat{u}(t,\xi)| = |u_0(\xi)| \implies \|\widehat{u}(t)\|_{L^2} = \|u_0\|_{L^2}.$$

So $|\widehat{u}|$ remains a probability distribution for all time $t \geqslant 0$.

1.19.2 The uncertainty principle

Can we closely predict both position and velocity? Can we have supp $u \subseteq I$ and supp $\widehat{u} \subseteq J$ for compactly supported intervals I, J? The answer is no. If supp u is compact, then $\widehat{u}$ is analytic. So u must be 0.

Let's try to localize our particle at $x = 0, \xi = 0$. Let

$$(\delta x)^2 = \int |u|^2(x) \cdot x^2 dx$$

be the mean square deviation from 0 . We can do the same for velocity to get

$$(\delta \xi)^2 = \int |\widehat{u}|^2(\xi) \cdot \xi^2 d\xi$$

Is there a function $u \in L^2$ with $\|u\|_{L^2} = 1$ such that δx and $\delta\xi$ are simultaneously small? This is not possible. Observe that

$$\delta x = \|x \cdot u\|_{L^2}$$

while Plancherel's theorem tells us that

$$\delta\xi = \|\xi \cdot \widehat{u}\|_{L^2} = \|\partial_x u\|_{L^2}$$

We can compute the inner product

$$\operatorname{Re}\int xu \cdot \overline{\partial_x u} dx = \int x \cdot \frac{1}{2} \underbrace{\partial_x |u|^2}_{u\partial_x\bar{u} + \bar{u}\partial_x u} dx$$

Now integrate by parts to get

$$\begin{aligned} &= -\int_n \frac{n}{2}|u|^2 dx \\ &= -\frac{n}{2}\|u\|_{L^2}^2 \end{aligned}$$

So we conclude that

$$\begin{aligned} \|u\|_{L^2}^2 &= -2n \operatorname{Re}\langle xu, \partial_x u\rangle_{L^2} \\ &\leqslant 2n\|xu\|_{L^2}\|\partial_x u\|_{L^2} \end{aligned}$$

So we get the following:

Theorem 1.152 (Uncertainty principle).

$$\delta x \cdot \delta \xi \geqslant \frac{1}{2n}$$

Remark 1.153.
This says that we cannot know the position of an electron without sacrificing information about its velocity. In physics, people write the Schrödinger equation as $i\partial_t u + c\Delta u = f$ where c is a constant involving $\hbar$, Planck's constant. This gives the following physically normalized version of the uncertainty principle:

$$\delta x \cdot \delta \xi \geqslant \frac{\hbar}{2n}$$

1.19.3 Oscillatory fintegrals and the KdV equation

We have seen the integral $\int e^{it\xi^2} e^{ix\cdot\xi}$. Can we compute the more general integral $\int e^{i\lambda\varphi(\xi)} d\xi$ where φ is a **phase function**? How does this integral behave as $\lambda \to \infty$? Let us make the following observation in 1 dimension.

Proposition 1.154.
If $\varphi' \neq 0$, then for any N,

$$\int e^{i\lambda\varphi(\xi)} a(\xi) d\xi = o\left(\lambda^{-N}\right)$$

This is called an oscillatory integral.

Proof.
Suppose $\varphi' \neq 0$. Then localize to a compact set with a function a and integrate by parts:

$$\begin{aligned}\int e^{i\lambda\varphi(\xi)} a(\xi) d\xi &= \int \varphi' e^{i\lambda\varphi} \cdot \frac{a}{\varphi'} d\xi \\ &= \frac{i}{\lambda} \int e^{i\lambda\varphi(\xi)} \partial_\xi \left(\frac{a}{\varphi'}\right) d\xi\end{aligned}$$

so we have gained a factor of $1/\lambda$. Now repeat this.

□ □

The conclusion is that the main contribution comes from the critical points of φ. The study of oscillatory integrals via their critical points is called the method of **stationary phase**. From the perspective of PDEs, we want to use oscillatory integrals to compute asymptotic expansions of fundamental solutions which are not explicit.

Example 1.155 (Korteweg-de Vries equation).
The KdV equation is

$$\left(\partial_t + \partial_x^3\right) u = 0$$

It describes unidirectional waves in a canal.

If you want to make this a linear equation, we can consider the case where this equals $6uu_x$. Let's compute a fundamental solution. We want to compute the inverse Fourier transform of $\frac{1}{\tau - \xi^3}$. For a forward fundamental solution, we want

$$K = \mathcal{F}^{-1}\left(\frac{1}{\tau - \xi^3 - i0}\right)$$

We have

$$K(t, \xi) = e^{it\xi^3}$$

If we take the Fourier transform in time, we get $K(t,\xi) = e^{it\xi^3}$. So now we want to take the integral

$$\int e^{i\left(t\xi^3+x\xi\right)} d\xi$$

The solution will not be an algebraic function; instead, it will be something we label as a "special function," the Airy function. In particular, $\mathcal{F}^{-1}\left(e^{i\xi^3}\right) = \mathrm{Ai}(x)$.

Let's try to compute the asymptotic behavior. The phase is $\varphi(\xi) = t\xi^3 + x\xi$. The critical points are when

$$3t\xi^2 + x = 0 \Longrightarrow \xi^2 = -\frac{x}{3t}$$

This has roots only when $x < 0$, which is why this equation only gives waves in 1 direction. We get two critical points:

$$\xi^1 = \sqrt{-\frac{x}{3t}}, \quad \xi^2 = -\sqrt{-\frac{x}{3t}}$$

At each critical point, replace the cubic polynomial with a quadratic polynomial which is the Taylor series of the polynomial, and take the Fourier transform like with our analysis of the Schrödinger equation.

1.20 The KdV Equation and the Wave Equation

1.20.1 Fundamental solution of the KdV equation

Last time, we were discussing the KdV equation

$$\left(\partial_t + \partial_x^3\right) u = f$$

We saw that the fundamental solution was given by

$$\widehat{K}(t,\xi) = e^{it\xi^3}$$

Taking the inverse Fourier transform in x gives

$$K(t,x) = \int e^{i\left(t\xi^3 - x\xi\right)} d\xi$$

This problem admits a type of scaling. If we want

$$\left(\partial_t + \partial_x^3\right) u = 0$$

then we can make a change of variables $u(x,t) \mapsto u\left(\lambda x, \lambda^3 t\right)$. If we want to get rid of the time variable, we can set $\xi = t^{-1/3}\xi$, so the integral becomes

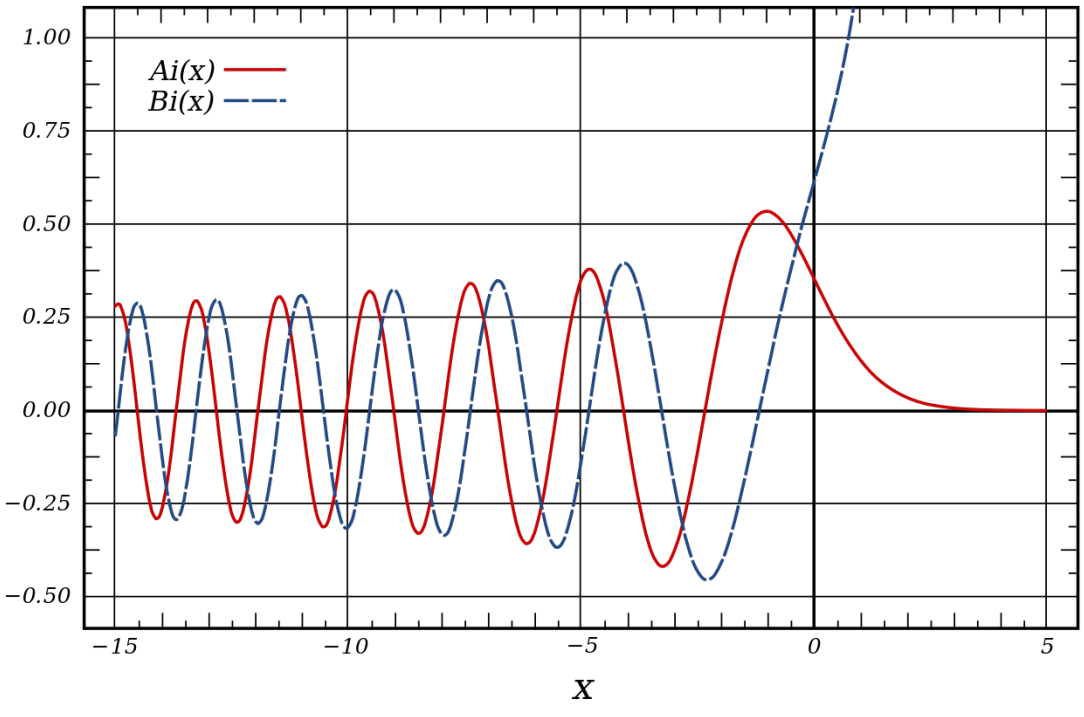
$$\begin{aligned} K(t,x) &= t^{-1/3} \int e^{i\eta^3 + x/t^{1/3}\eta} d\eta \\ &= t^{-1/3} K\left(1, x/t^{1/3}\right) \\ &= t^{-1/3} \operatorname{Ai}\left(x/t^{1/3}\right) \end{aligned}$$

where

$$\operatorname{Ai}(x) = \mathcal{F}^{-1}\left(e^{i\xi^3}\right) = \int e^{i\left(t\xi^3 + x\xi\right)} d\xi$$

In this integral, we have the phase function $\varphi(\xi) = \xi^3 + x\xi$. The critical points, with $\varphi_\xi = 0$, are $\xi_{1,2} = \pm\sqrt{-x/3}$ with $x < 0$.

Let's draw a picture of the Airy function; this is real-valued because the equation is real, so the real and imaginary parts of any solution should also be solutions. At $+\infty$, we have no stationary points, so we expect rapid decay. This decay is $O\left(e^{-x^{3/2}}\right)$, which one can prove by changing the contour in the integral (to some other integral over a contour in the complex plane).



Choose $\xi_1 = \sqrt{-x/3}$, look at the contribution around ξ_1, and take 2Re.

$$\varphi(\xi) = \varphi(\xi_1) + \frac{1}{2}\varphi''(\xi_1)(\xi - \xi_1)^2 + \underbrace{O\left((\xi - \xi_1)^3\right)}_{\text{discard}}$$

We are multiplying two functions, a Gaussian and a function with oscillation.

Recall that $\mathcal{F}^{-1}\left(e^{i\lambda\xi^2/2}\right) = -\frac{1}{(i\lambda)^{n/2}}e^{ix^2/(2\lambda)}$. Now obsesrve that the Fouerier transform lets us figure out the integral of a function: $\widehat{u}(0) = \frac{1}{(2\pi)^{n/2}} = \int u(x)dx$. So can calculate this integral:

$$\int e^{i\left(\varphi(\xi_1)+\frac{1}{2}\varphi'(\xi_1)(\xi-\xi_1)^2\right)}d\xi = e^{i\varphi(\xi_1)}\frac{1}{(i\varphi''(\xi_1))^{1/2}}$$

Now write

$$\varphi(\xi_1) = \xi_1\left(\xi_1^2 + x\right) = \frac{2}{3}x\sqrt{-x/2} = c(-x)^{3/2}$$
$$\varphi''(\xi) = 6\xi = c(-x)^{1/2}$$

In total, we get something of the form

$$e^{ic(-x)^{3/2}}(-x)^{-1/4}$$

This left term oscillates faster and taster, while the right term has a decay. So we can improve our picture of the Airy function.

The homework says that $\mathrm{Ai}''(x) = x\mathrm{Ai}(x)$ (up to some constants/signs). There are two solutions to this equation; why are we only getting the Airy function? This is because using the Fourier transform only solves for temperate solutions. The other solution will look like the Airy function for negative x but has exponential (specifically $e^{+x^{3/2}}$) growth as $x \to \infty$. The Airy function has nice properties, and it actually extends to a holomorphic function.

Professor Tataru really likes the Airy function. He used to put it on exams, until one time when he put it on a calculus exam. That didn't go so well.

1.20.2 Analysis of the wave equation

Definition 1.156 (d'Allembertian).
The d'Allembertian is the partial differential operator

$$\Box = \partial_t^2 - \Delta_x$$

Definition 1.157 (Wave equation).
The wave equation is the equation

$$\begin{cases} \Box u = f \\ u(t=0) = u_0 \\ \partial_t u(t=0) = u_1 \end{cases}$$

This is an evolution equation which is 2nd order in t.

What does the wave equation model? In 1 dimension, this modes an elastic string. In 2 dimensions, it models an elastic drum, and in 3 dimensions, it models an elastic solid. The wave equation also models, sound, light, and electromagnetism.

Our goal is to find the fundamental solution. The symbol for the equation is $P(\tau, \xi) = -\tau^2 + \xi^2$. Then we get

$$K(t,x) = \mathcal{F}^{-1}\left(\frac{1}{-\tau^2 + \xi^2}\right)$$

The zero set of P, (the characteristic set) contains the points where $\tau^2 = \xi^2$. In 1 dimension, this looks like an X, but in n dimensions, this looks like 2 cones.

Like we have seen before, this is not uniquely defined as a distribution. We want to pick a forward fundamental solution, so we will look at this as a function of τ and think of this as a function which is holomorphic in the lower half plane:

$$K(t,x) = \mathcal{F}^{-1}\left(\frac{1}{-(\tau - i0)^2 + \xi^2}\right)$$

We will take the Fourier transform first in τ and then in ξ. First, expand the fraction into partial fractions:

$$\frac{1}{-\tau^2 + \xi^2} = \frac{A}{\tau - |\xi|} + \frac{B}{\tau + |\xi|}$$
$$A = -\frac{1}{2|\xi|}, \quad B = \frac{1}{2|\xi|}$$

So

$$\mathcal{F}^{-1}\left(\frac{1}{-(\tau - i0)^2 + \xi^2}\right) = -H(t)\frac{e^{it|\xi|} - e^{-it|\xi|}}{2|\xi|}$$
$$= -i\frac{\sin(t|\xi|)}{|\xi|}$$

This is hard to compute the Fourier transform directly in n dimensions, but the 1-dimensional computation is easier: We are looking at

$$-i\frac{\sin(t\xi)}{\xi} = \frac{1}{2}\frac{e^{it\xi} - e^{-it\xi}}{\xi - i0}$$

Taking the inverse Fourier transform, we note that multiplying phase factors just translate the Fourier transform. We get

$$\frac{1}{2}(H(x+t) - H(x-t)) = \frac{1}{2}1_{[-t,t]}.$$

Theorem 1.158.
The fundamental solution to the wave equation in 1 dimension is

$$K(t,x) = \begin{cases} 1/2 & x > 0, -t \leqslant x \leqslant t \\ 0 & \text{otherwise} \end{cases}$$

How should we approach this for $n \geqslant 2$? The distribution $\frac{1}{-(\tau - i0)^2 + \xi^2}$ is homogeneous of order -2, so K will be homogeneous of order $(-n-1) - (-2) = -n + 1$. We could try to replace x by $r = |x|$, making an ansatz that the solution is radial, but this is not very nice because we still have a PDE in 2-dimensions. This is easier to solve in 3 dimensions, so we can add a dimension and then pretend it doesn't exist after we solve the equation; this is probably how it was done in the early 1900 s.

Instead, let's look at the symmetries of $\square$. This is translation-invariant and invariant under rigid rotations in x. The latter suggests that we could look for more general linear transformations which $\square$ is invariant under. We will change our notation from (t,x) to $(x_0, \dots, x_n)$, where $t = x_0$ and $x = (x_1, \dots, x_n)$. Here's how to make sure we won't get confused about which variables are time. A notational convention which goes back to Einstein says we write x_α for $\alpha = 0, \dots, n$ and x_j for $j = 1, \dots, n$. Now apply a change of variables to get $y = Ax$. Then

$$\frac{\partial}{\partial x_k} = \frac{\partial y_j}{\partial x_k}\frac{\partial}{\partial y_j} = a_{j,k}\frac{\partial}{\partial y_j}$$

1.21 Fundamental Solutions to the Wave Equation

1.21.1 Lorentz invariance of fundamental solution

Last time, we were solving the wave equation where

$$\begin{aligned}\Box &= \partial_t^2 - \Delta_x \\ &= m^{\alpha,\beta}\partial_\alpha\partial_\beta\end{aligned}$$

in coordinates $t = x_0$ and $\partial_t = \partial_0$. The matrix M is given by

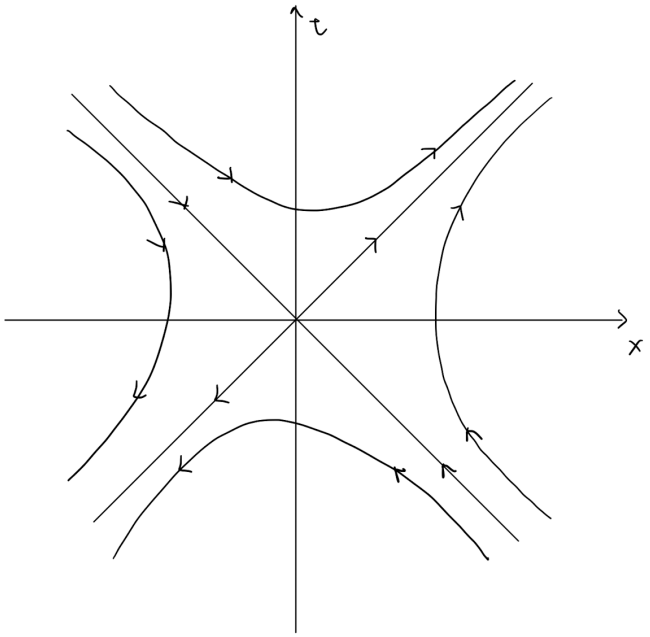
$$M = \begin{bmatrix} -1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix}$$

Last time, we determined that a fundamental solution is homogeneous of order $1 - n$ and must move forward in time. We looked at symmetries of the equation when we make a linear change of coordinates $x = Ay$. We saw that such a linear change of coordinates leaves $\Box$ unchanged if and only if

$$A^\top M A = M$$

This is a group, called the **Lorentz group**; if M were the identity matrix, this would be the group or orthogonal matrices. What are the generators for this group?

1. Rigid rotations: $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, where O is an $n \times n$ orthogonal matrix. These were the symmetries corresponding to the Laplacian.
2. Look at $1 + 1$ dimensions and leave the rest unchanged: Since we can apply rotations to the last n dimensions, we only need to mix the time dimension and the first space dimension. Observe that
$$\begin{bmatrix} a & c \\ b & d \end{bmatrix}\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$
If -1 were 1 , we would get rotations:
$$A = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} = \text{ rotation by angle } \theta$$
This keeps $t^2 + x^2$ unchanged; this is like rotating around a point in a circle by angle θ. With the -1, we get
$$A = \begin{bmatrix} \cosh\varphi & \sinh\varphi \\ \sinh\varphi & \cosh\varphi \end{bmatrix} = \text{ hyperbolic rotation by angle } \varphi$$
Such matrices keep $t^2 - x^2$ unchanged. Rather than circles, here's what the level sets look like:



Here, we dilate the $t = x$ direction and shrink the $t = -x$ direction. This suggests that we make a change of variables $u = t + x$ and $v = -x$. Then $\partial_t^2 - \partial_x^2 = 4\partial_u\partial_v$. Then the transformation $u \mapsto \lambda u, v \mapsto \lambda^{-1}v$ preserves the operator in this null frame.

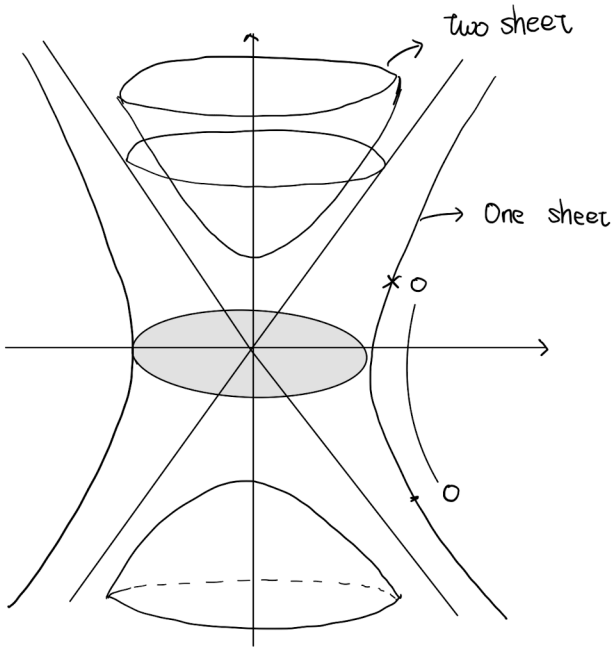
Theorem 1.159.
The Lorentz group is generated by rigid spatial rotations and 1-d hyperbolic rotations.

Remark 1.160.
Hyperbolic rotations are sometimes referred to as Lorentz boosts. These hyperbolic rotations are what happens in special relativity when you switch between observers in different reference frames.

We say that the solution to the wave equation is **Lorentz invariant**.

1.21.2 Calculation of fundamental solutions

We now know that the fundamental solution of the wave equation should be a "function" of $t^2 - x^2$. Here is what the picture should look like in higher dimensions.



The level sets should be forward and backward cones and hyperboloids. We get 1-sheeted and 2-sheeted hyperboloids. On the 1-sheeted hyperboloids, the forward in time points are connected to the backwards in time points, which must give 0 for our forward time solution. So these must be 0 . Thus, $K = K\left(t^2 - x^2\right) = K(y)$ must be supported in the forward cone $\{t^2 - x^2 \geqslant 0\}$ We want a homogeneous distribution of y which is $\frac{1-n}{2}$ homogeneous (since we are now working with the squares of t, x) and supported in $y \geqslant 0$.

- In 1 dimension, we want a homogeneous distribution of order 0 , supported where $y > 0$. So $K(y) = cH(y)$, and we saw earlier that this constant is $c = 1/2$.
- In 2-dimensions, we want a homogeneous distribution of order $-1/2$, supported where $y > 0$. So we get

  $$K(y) = \begin{cases} c_2 \frac{1}{\sqrt{y}} & y > 0 \\ 0 & y \leqslant 0 \end{cases}$$

  So we get

  $$K(t, x) = c_2 \frac{1}{\sqrt{\left(t^2 - x^2\right)_+}} 1_{t \geqslant 0}$$

- In 3 -dimensions, we cannot get a function which is homogeneous of order -1. The two distributions that span the space of homogeneous distributions of order -1 are δ_0 and PV $\frac{1}{y}$. The latter is supported everywhere, so we take $K(y) = \delta_{y=0}$

  $$K(t, x) = c_3 \delta_{t^2 - x^2 = 0} 1_{t \geqslant 0}$$

- In 4 dimensions, we need homogeneity of order $-3/2$. However, $\frac{1}{y_+^{3/2}} \notin L^1_{\text{loc}}$. Define

  $$\frac{1}{y_+^{3/2}} := -2\partial_y \frac{1}{y_+^{1/2}}$$

  This is a distribution, not a function. We can repeat this differentiation procedure to get a solution for all even dimensions.
- In 5 dimensions, we can get a solution which is homogeneous of order -2 by differentiating $\delta_{y=0}$. We can keep differentiating to get solutions in all odd dimensions.

1.21.3 Determination of constants for fundamental solutions

Here is a formal computation: If $\Box u = f$, let's see how $\int u dx$ behaves as a function of time.

$$\frac{d}{dt}\int u dx = \int u_t dx$$

$$\frac{d^2}{dt^2}\int u dx = \int u_{tt} dx = \int \Delta_x u + f dx = \int f dx,$$

since we can get rid of the Laplacian using integration by parts. If $f = \delta_0$ and $u = K$, then $u = 0$ for $t < 0$, so

$$I(t) = \int u dx = 0 \quad \text{for all } t < 0.$$

Additionally, we get

$$I''(t) = \delta_{t=0}.$$

This tells us that

$$I(t) = t 1_{t \geqslant 0}.$$

so

$$\int K(t,x) dx = t.$$

- In 2 dimensions, we have

  $$K(t,x) = \frac{c_2}{\sqrt{(t^2 - x^2)_+}}$$

  so the equation

  $$t = c_2 \int \frac{1}{\sqrt{(t^2 - x^2)_+}} dx$$

  holds for all t. if we set $t = 1$, then we get

  $$1 = c_2 \int_{B(0,1)} \frac{1}{\sqrt{1 - r^2}} r dr d\theta = c_2 2\pi \left[-\sqrt{1 - r^2}\right]_0^1$$

  which tells us that

  $$c_2 = \frac{1}{2\pi}.$$

- In 3 dimensions, we want to find c_3. What is $\delta_{t^2 - x^2}$?

  $$\delta_0 = \frac{1}{2\pi i}\left(\frac{1}{y - i0} - \frac{1}{y + i0}\right)$$

  so we can write

  $$\delta_{t^2 - x^2} = \frac{1}{2\pi i}\left(\frac{1}{t^2 - x^2 + i0} - \frac{1}{t^2 - x^2 - i0}\right)$$

  Note that

  $$\frac{1}{t^2 - x^2} = \frac{1}{t - |x|}\frac{1}{t + |x|},$$

  where the left term vanishes on the cone, and $t + |x|$ is $2t$ on the cone. so we can write

  $$\delta_{t^2 - x^2 = 0} = \underbrace{\delta_{t=|x|}}_{\text{surface measure on } |x| = t} \cdot \frac{1}{2t}$$

  If we have a surface $\Sigma = \{\phi = 0\}$, this is like normalizing to make $|\nabla \phi| = 1$ The computation becomes

  $$K(t,x) = c_3 \frac{1}{t}\delta_{|x|=t}$$

  $$t = \int \frac{c_3}{t}\delta_{|x|=t} dx = \frac{c_3}{t}\underbrace{\text{Area}(\{|x| = t\})}_{= 4\pi t^2}.$$

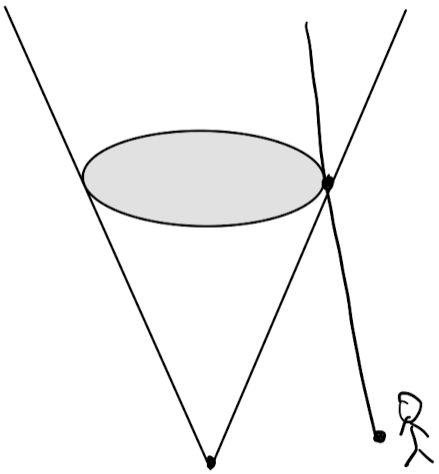
  so we get

  $$c_3 = \frac{1}{4\pi}$$

1.21.4 Physical interpretation of solutions to the wave equation

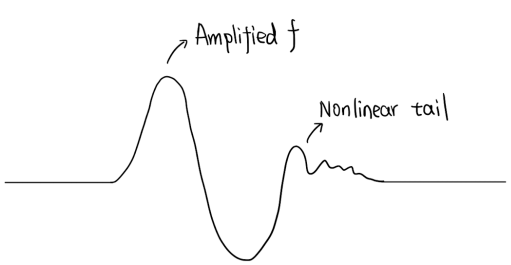
Here are two key properties of the wave equation:

1. All forward solutions are supported on the forward cone. This is referred as the finite speed of propagation. This says that waves move with speed $\leqslant 1$. If we normalize the equation with physical constants to get $c^2\partial_t^2 - \Delta_x$, where c is the speed of light, then this says that no waves move faster than the speed of light. An observer at position x only observes the wave at the time at which the cone hits the observer's timeline:



- Consider 3 dimensions, where the fundamental solution K is supported exactly on the cone. Here, waves hit the observer just once, and we don't see them again. This is called the Huygens principle.

Remark 1.161.
The equations of physics are nonlinear; this linear PDE is just the best linear approximation. The finite speed of propagation remains, but Huygen's principle does not hold in general. When scientists observed gravitational waves recently, they observed both a Dirac mass and a nonlinear tail.



1.21.5 Next steps: Fourier series

Our next goal is to learn about the connection between the Fourier transform and Fourier series. The Fourier transform $\widehat{u}$ of $u : \mathbb{R}^n \to \mathbb{C}$ is given by

$$u(x) = \int \widehat{u}(\xi) e^{ix\cdot\xi} d\xi$$

In calculus, you may have encountered Fourier series:

Definition 1.162 (Fourier series).
If $u : [0, 2\pi] \to \mathbb{C}$, then the Fourier series for u is given by $u(x) = \sum_n c_n e^{inx} = \sum_n c_n(\cos(nx) + i\sin(nx))$

Not all PDEs can be solved; we will see more about this next time.

1.22 Fourier Transforms of Periodic Functions and Local Solvability of Partial Differential Operators

1.22.1 Fourier transforms of periodic functions

A function f is periodic if

$$f(x) = f(x+a)$$

for some a and for all x.

Definition 1.163 (Periodic).
$f \in \mathcal{D}'$ is periodic of period a if

$$f(\phi) = f(\phi(\cdot + a))$$

Suppose f is periodic, what can we say about $\widehat{f}$? Recall that for functions,

$$\widehat{f}(\cdot + a) = e^{ia\cdot\xi}\widehat{f}$$

Using the periodic condition, write this as the multiplication

$$\widehat{f} \cdot \left(1 - e^{ia\xi}\right) = 0$$

Note that $1 - e^{ia\xi} \neq 0$ unless $\xi = \frac{2\pi n}{a}$. Then $\operatorname{supp} \widehat{f} \subseteq \frac{2\pi n}{a}\mathbb{Z}$. As an analogy look at the condition $xf = 0 \implies f = c\delta_0$; here, we have zeros at many points. So we conclude that

$$\widehat{f} = \sum_n c_n \delta_{\frac{2\pi n}{a}}$$

Theorem 1.164.
The coefficients c_n are the Fourier coefficients for f in the interval $[0, a]$, and

$$f(x) = \sum_n c_n e^{\frac{2\pi i}{a} nx}.$$

Here, we have ignored the factors of 2π.

Remark 1.165.
We can multiply f by $e^{-\frac{2\pi i}{a}m}$ and integrate from 0 to a to get

$$c_n = \int f(x) e^{-\frac{2\pi i}{a} mx}.$$

Example 1.166.
The simplest periodic distribution is

$$f_a = \sum_n \delta_{na}.$$

Then

$$\widehat{f_a} = \sum_n c_n \delta_{\frac{2\pi}{a} n}$$

If we write

$$f_a \left(1 - e^{\frac{2\pi i x}{a}}\right) = 0$$

then we get

$$\widehat{f_a} = \widehat{f_a}\left(\cdot + \frac{2\pi}{a}\right).$$

Thus, all the c_n s are the same. So

$$\widehat{f_a} = c_a \sum_n \delta_{\frac{2\pi n}{a}} = c_a f_{\frac{2\pi}{a}}$$

What is c_a? Apply this to a Schwarz function: $\widehat{f}(\phi) = f(\widehat{\phi})$ by definition, so

$$c_a \sum_{n\in\mathbb{Z}} \phi\left(\frac{2\pi n}{a}\right) = \sum_{m\in\mathbb{Z}} \widehat{\phi}(ma)$$

This is called the **Poisson summation formula**. Now what happens if we replace ϕ by $\phi e^{ix}\cdot\xi_0$? Then $\widehat{\phi}(\xi)$ becomes $\widehat{\phi}\,(\xi-\xi_0)$. The Poisson summation formula gives

$$c_a \sum_n \phi\left(\frac{2\pi n}{a}\right) e^{i\frac{2\pi n}{a}\xi_0} = \sum \widehat{\phi}\,(ma-\xi_0)$$

The dependence of ξ_0 on the left hand side is simple. Integrate to get

$$\underbrace{\int_0^a \sum_n \phi\left(\frac{2\pi n}{a}\right) e^{i\frac{2\pi n}{a}\xi_0} d\xi_0}_{=ac_a\phi(0)} = \int_0^a \sum_m \widehat{\phi}\,(ma-\xi_0)\,d\xi_0$$

$$\begin{aligned} &= \int \widehat{\phi}(\xi) d\xi \\ &= \widehat{\phi}(1) \\ &= \phi\left(\frac{1}{\sqrt{2\pi}}\delta_0\right) \\ &= \frac{1}{\sqrt{2\pi}}\phi(0) \end{aligned}$$

Accounting for the constants we ignored before, we get

$$c_a = \frac{1}{2\pi a}.$$

Remark 1.167.
We can use the Poisson summation formula to compute all sorts of series. Recall that $\mathcal{F}\left(\frac{1}{1+x}\right) = ce^{-|\xi|}$ (perhaps omitting constants). Choose $a = 2\pi$. The Poisson summation formula tells us that

$$\sum_{m\in\mathbb{Z}} \frac{1}{n^2+1} = \sum_m e^{-2\pi|m|} = \frac{2}{1-e^{2\pi}} - 1$$

where we have ignored the constants.

1.22.2 Local solvability of partial differential operator

Let $P(D)$ be our partial differential operator with constant coefficients.

Definition 1.168 (Solvable).
$P(D)$ is solvable if for each f, the equation $P(D)u = f$ admits at least one solution.

If $f \in \mathcal{D}'$, then $u \in \mathcal{D}'$. If $f \in \mathcal{S}$, then $u \in \mathcal{S}$. In general, the regularity of f and u will be related, so when we say $P(D)$ is solvable, we specify a class of functions f.

Definition 1.169 (Locally solvable I)**.**
$P(D)$ is locally solvable if for each $f \in \mathcal{E}'$, there exists a solution $u \in \mathcal{D}'$ in a neighborhood of the support of f.

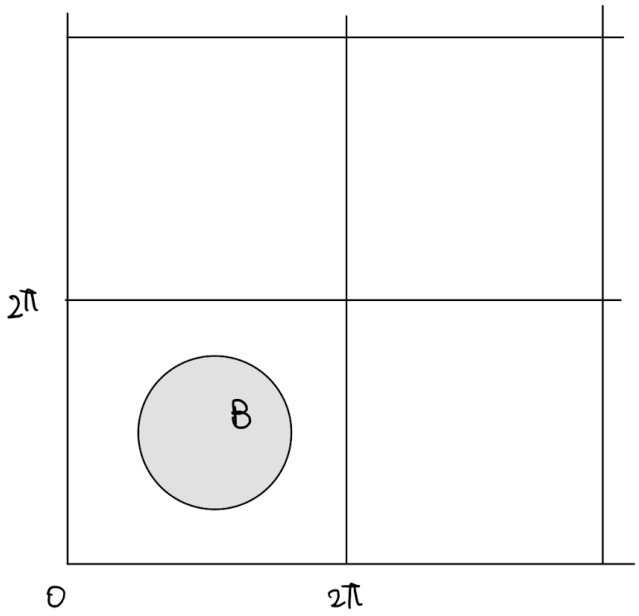
If $u \in \mathcal{E}'$, then $P(\xi)\widehat{u}(\xi) = \widehat{f}(\xi)$ for $\xi \in \mathbb{C}^n$. Here is a narrower version, which we may regard as the "real definition" of local solvability:

Definition 1.170 (Locally solvable II)**.**
$P(D)$ is locally solvable if for each x_0, there is an $\varepsilon > 0$ such that if $\operatorname{supp} f \subseteq B(x_0, \varepsilon)$, then a solution exists.

For today, we will deal with the first, more relaxed definition.

Theorem 1.171.
Every constant coefficient partial differential operator is locally solvable (in the relaxed sense).

Proof.
Suppose f is supported in $B \subseteq [0, 2\pi]^n$. Take $\widetilde{f}$ to be the periodic extension of f, and look for a periodic solution $\widetilde{u}$ to $P(D)\widehat{u} = \widetilde{f}$.



What does this periodization do? Originally, $P(D)u = f$ gives $P(\xi)\widehat{u} = \widehat{f}$, so $\widehat{u} = \frac{1}{P(\xi)}\widehat{f}$ However, this has issues because $P(\xi)$ can have issues. In the periodic case, we know

$$\widehat{\widetilde{f}}(\xi) = \sum_{m \in \mathbb{Z}^n} f_m \delta_m$$
$$\widehat{\widetilde{u}}(\xi) = \sum_{m \in \mathbb{Z}} u_m \delta_m$$

The advantage is that we only $P(m) \neq 0$ on lattice points $m \in \mathbb{Z}^n$. However, the Fourier transform is defined for temperate distributions, so we need about on $\frac{f_m}{P(m)}$. More precisely, we need a bound

$$|P(m)| \geqslant (1 + |m|)^{-N}$$

What if P has zeroes on the lattice points? Make the change of notation $f \mapsto fe^{ix\cdot\xi} = g$ so $u \mapsto ue^{ix\xi} = v$. We can ask this question for the phase-shifted variables. To study our equation, we need to expand

$$P(D)u = P(D)\left(ve^{-ix\cdot\xi}\right)$$

To use the Leibniz rule, note that,

$$\begin{aligned} D_j\left(ve^{-ix\xi}\right) &= D_j ve^{-ix\xi} + vD_j e^{-ix;\xi} \\ &= e^{-ix\cdot\xi}\left(D_j v - v\xi_j\right) \\ &= e^{-x\cdot\xi}(D_j - \xi_j)v, \end{aligned}$$

We can write this as $e^{ix\cdot\xi}D_j e^{-ix\xi} = D_j - \xi_j$, which we may think of as a conjugation. Referring to our equation, we get

$$\begin{aligned} P(D)u &= P(D)(ve^{-ix\cdot\xi}) \\ &= e^{-ix\cdot\xi}p(D-\xi)v \\ &= f \end{aligned}$$

which tells us that we have replaced $P(D)u = f$ with

$$P(D-\xi)v = g$$

So we only need to solve the new periodic problem is to define

$$v_m = \frac{g_m}{P(m-\xi)}, \quad m \in \mathbb{Z}$$

Now we only need to find some $\xi \in [0,1]^n$ such that

$$|P(m-\xi)| \geqslant (1+|m|)^{-N} \quad \forall m$$

The following lemma tells us we can find such a ξ.

Lemma 1.172.
If δ is small enough, then

$$\int \frac{1}{(P(\eta))^\delta}\frac{1}{(1+|\eta|)^N}d\eta < \infty.$$

Proof.
In 1 dimension, use partial fractions. Then reduce any number of dimensions to the 1-dimensional case. □ □

How does this help us? Write $\eta = m + \xi$ with $m \in \mathbb{Z}^n$ and $\xi \in [0,1]^n$. Then

$$\int_\xi \sum_m \frac{1}{P(m-\xi)^\delta}\frac{1}{(1+|m|)^N}d\xi < \infty$$

So for almost every ξ,

$$\sum_m \frac{1}{|P(m-\xi)|^\delta}\frac{1}{(1+|m|)^N} = M < \infty$$

This tells us that

$$|P(m-\xi)| \geqslant M^{-1/d}(1+|m|)^{-N/\delta}$$

which is exactly the relation we want to have. □ □

1.23 Properties of Harmonic Functions

1.23.1 Elliptic regularity

Recall that if we have the Laplace equation

$$-\Delta u = f \quad \text{in } \mathbb{R}^n$$

then we have the fundamental solution

$$K(x) = \begin{cases} \frac{c_n}{|x|^{2-n}} & n \geqslant 3 \\ \frac{1}{2\pi} \ln |x| & n = 2, \end{cases}$$

and we can get a solution $u = K * f$. However, there are a number of questions we have not answered, such as uniqueness of solutions.

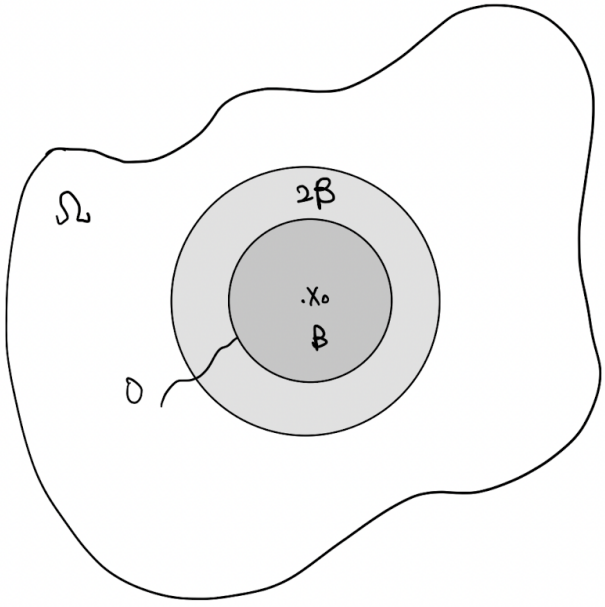
Definition 1.173 (Harmonic).
A function such that $-\Delta u = 0$ is called harmonic.

Theorem 1.174 (Elliptic regularity).
Harmonic functions are smooth.

That is, if we have a local solution $u \in \mathcal{D}'$, we want to show that $u \in C^\infty$. Why should harmonic functions be smooth? This is because the fundamental solution K is smooth away from 0 . Let's see how the reasoning goes.

Proof.
Let Ω be the domain where u lives. Choose a point $x_0 \in \Omega$, and we want to show that u is smooth around x_0. Draw a ball B around x_0 and a larger ball $2B$ around B. To use the fundamental solution, chop off u by using a cutoff function

$$\chi(x) = \begin{cases} 1 & x \in B \\ \text{smooth} & x \in 2B \backslash B \\ 0 & x \in 2B^c \end{cases}$$



If we let $v = \chi u$, then

$$-\Delta v = \underbrace{-\chi \Delta u}_{=0} - u\Delta\chi - 2\nabla u \cdot \nabla\chi$$

This gives us the new problem

$$-\Delta = f, \quad f \in \mathcal{D}', \quad \operatorname{supp} f \subseteq 2B\backslash B$$

Then

$$\begin{aligned} v(x) &= (K * f)(x) \\ &= \int K(x-y) f(y) dy \end{aligned}$$

Suppose we want a local solution in, say, $B/2$, where B has radius R. If $x \in B/2$ and $y \in 2B\backslash B$, then $|x-y| \geqslant r/2$. Now $K(z)$ is smooth where $|z| \geqslant r/2$, which means this convolution is smooth for $x \in B/2$. □ □

Remark 1.175.

- We didn't use much about the Laplace equation itself here. We only used the fact that K is smooth away from 0.
- This is not all there is to elliptic regularity. K is analytic away from 0, which tells us that u is analytic.
- More generally, we may want to make statements about what kind of regularity u has if f has a certain degree of regularity. This is what elliptic regularity really is, and this is only the tip of the iceberg.

1.23.2 The maximum principle

Definition 1.176 (subharmonic, superharmonic).
A function u such that $-\Delta u \leqslant 0$ is called subharmonic. A fuction u such that $-\Delta u \geqslant 0$ is called superharmonic.

We will prove results for harmonic functions and claim that they hold for sub and superharmonic functions, as well.

Suppose $-\Delta u = 0$ in Ω. Where is the max / min of u ? The first step to answering this question is to look at the **mean value property**.

Theorem 1.177 (Mean value property).
Suppose $-\Delta u = 0$ in $B(x_0, a)$. Then

$$u(x_0) = \frac{1}{|B|}\int_B u(x)dx = \frac{1}{|\partial B|}\int_{\partial B} u(x)d\sigma,$$

where σ is surface measure on the sphere ∂B.

Remark 1.178.
If we assume u is subharmonic, i.e. $-\Delta u \leqslant 0$, then we get $\leqslant$ instead of equalities. The reverse inequality holds for superharmonic functions.

Lemma 1.179 (Green's theorem).
Suppose $u : \Omega \to \mathbb{R}$. Then

$$\int_\Omega \partial_j u dx = \int_{\partial\Omega} u \cdot \nu_j d\sigma$$

where ν_j is the outward pointing normal to $\partial\Omega$. Equivalently,

$$\int \underbrace{\partial_j u_j}_{\operatorname{div} u} dx = \int_{\partial\Omega} u \cdot \nu d\sigma$$

Here's how we can use this: Integrating by parts twice in the following integral keeps the sign the same and introducing 2 boundary terms:

$$\int_\Omega \Delta u \cdot v dx - \int_\Omega u \cdot \Delta v dx = \int_{\partial\Omega} \underbrace{\partial_j u \nu_j}_{\frac{\partial u}{\partial \nu}} \cdot v - u \cdot \underbrace{\nu_j \partial_j v}_{\frac{\partial v}{\partial \nu}} d\sigma,$$

where these ν are normal derivatives. Now let's prove the mean value property:

Proof.
Suppose $B = B(0, r)$, and apply Green's theorem with a well-chosen v. Looking at our equation, it would be nice if we could make $v = 0$ on the boundary. So we can try

$$v = K(|x|) - K(r).$$

We get

$$u(0) = c \int_{\partial B} u d\sigma.$$

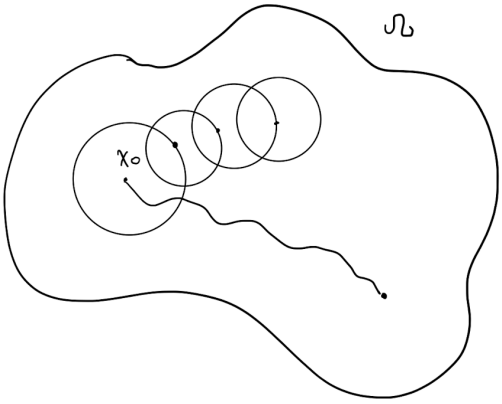
This holds for all harmonic functions. If we set $u = 1$, then we get $c = \frac{1}{|\frac{1}{|B|}, \text{ so } u=} \frac{1}{\partial B|} \int_{\partial B} u$. □ □

Corollary 1.180.
If $u(x_0) = \max u$ for $x_0 \in B$, then u is constant in B.

Remark 1.181.
If u is subharmonic, the same holds. But if u is superharmonic, then we need to replace the maximum with the minimum in this property.

Theorem 1.182 (Strong maximal principle).
Suppose $u \in C^2(\Omega) \cap C(\bar{\Omega})$ is harmonic. Then $\max_{\bar{\Omega}} u = \max_{\partial\Omega} u$ Moreover, if $\max u$ is attained inside Ω, then u is constant.

Proof.
If $\max u$ is only attained on $\partial\Omega$, then we ar edone. What if $\max u$ is attained at $x_0 \in \Omega$? Here is a proof by picture. Put a ball around x_0. By the corollary, u is constant in B. Then the other points in this ball are maximum points, and we can get to any other point via a sequence of balls.



If you want to write down a proof, you can use path-connectedness, or you can use an argument like this: Let $A = \{x \in \Omega : u(x) = u(x_0)\}$. Since u is continuous, A is closed. But the corollary says that if $x_0 \in A$, then $B(X_0, r) \subseteq A$. So A is open. Thus, $A \subseteq \Omega$ is open and closed, and if Ω is connected, we get $A = \Omega$. □ □

Remark 1.183.
If u is subharmonic, the same holds. But if u is superharmonic, then we need to replace the maximum with the minimum in this property. Besides, the hypotheses here are much stronger than they need to be.

Corollary 1.184 (Comparison principle).
Let u be subharmonic, i.e., $-\Delta u \leqslant 0$, and let v be subharmonic, i.e., $-\Delta v \geqslant 0$. If $u \leqslant v$ on $\partial\Omega$, then $u \leqslant v$ in $\bar{\partial}\Omega$.

This comparison principle is the correct statement for nonlinear elliptic stuff and also for the Hamilton-Jacobi equations. There is a simpler proof of the maximum principle without the use of the fundamental solution where we drop the strong part.

Proof.
Suppose first that $-\Delta u < 0$. Let x_0 be a maximum point inside Ω. Then $\nabla u(x_0) = 0$, and $Hu(x_0) < 0$, where $H = \frac{\partial^2 u}{\partial x_i \partial x_j}$ is the Hessian matrix. Observe that

$$\Delta u = \sum_j \partial_j \partial_j u = \operatorname{tr} Hu \leqslant 0$$

Then $\Delta u(x_0) \leqslant 0$, so $-\Delta u(x_0) \geqslant 0$. But this contradicts our assumption that $-\Delta u < 0$ Now if $-\Delta u \leqslant 0$, then we penalize u by replacing u by $u_\varepsilon = u + \varepsilon x^2$. Then

$$-\Delta u_\varepsilon = -\Delta u - 2u\varepsilon < 0$$

This tells us that

$$\max_{\bar{\Omega}} u_\varepsilon = \max_{\partial_\Omega} u_\varepsilon$$

If we let $\varepsilon \to 0$, both sides converge uniformly to $\max_{\bar{\Omega}} u$ and $\max_{\partial\Omega} u$, respectively. □

1.23.3 Liouville's thoerem

We have been looking at harmonic functions in a domain Ω. What if we are looking at harmonic functions in all of $\mathbb{R}^n$? If you allow exponential growth, then the sky is the limit as to what you can get. But what if we only want polynomial growth. Further yet, what if u is bounded?

Theorem 1.185 (Liouville).
Let u be harmonic in $\mathbb{R}^n$. If u is bounded, then u is constant.

Proof.
If u is harmonic, so are its derivatives. Then

$$\begin{aligned}\partial_j u(x_0) &\overset{\text{MVP}}{=} \int_\Omega \partial_j u(x) dx \\ &= \frac{1}{|B_R|}\int_{\partial B_R} u \cdot \nu_j d\sigma(x)\end{aligned}$$

If $|u| \leqslant M$, we can estimate this by

$$\begin{aligned}|\partial_j u(x_0)| &\leqslant \underbrace{\frac{1}{B_R}}_{R^n} M \underbrace{|\partial B_R|}_{R^{n-1}} \\ &\lesssim \frac{M}{R} \\ &\xrightarrow{R\to\infty} 0\end{aligned}$$

So $\nabla u(x_0) = 0$, which means that u is constant. □ □

Remark 1.186.
If u is temperate, then $\hat{u}\|\xi\|^2 = 0$, so $\hat{u}$ is supported at 0. Then $\hat{u} = \sum_\alpha c_\alpha \delta_0^{(\alpha)}$, which implies that u is a polynomial. Thus, all temperate harmonic functions are polynomials. This also serves as a proof of Liouville's theorem, since the only polynomials are constant.

1.23.4 Boundary value problem

Let $\Omega \subset \mathbb{R}^n$, and suppose that

$$\begin{cases} -\Delta u = f & \text{in } \Omega \\ u = g & \text{on } \partial\Omega \end{cases}.$$

This give us uniqueness: Suppose u_1, u_2 are solutions. If $u_1 - u_2 = v$, then v is harmonic. The maximum and minimum principles give

$$\begin{aligned}\max_\Omega v &\leqslant \max_{\partial\Omega} v = 0 \\ \min_\Omega v &\geqslant \min_{\partial\Omega} v = 0\end{aligned}$$

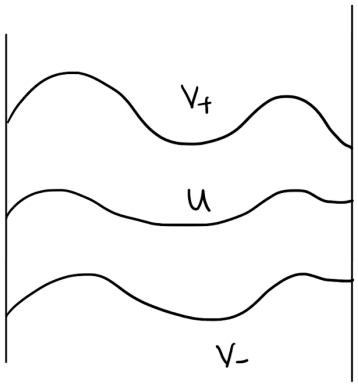
So $v = 0$. There is also a proof of existence using hte maximum principle. Consider a subsolution v^-satisfying

$$\begin{cases} -\Delta v^- \leqslant f \\ v \leqslant g \end{cases}$$

and a supersolution satisfying

$$\begin{cases} -\Delta v^+ \geqslant f \\ v \geqslant g \end{cases}$$

The maximum principle $v^+ \geqslant v^-$. Taking the maximum over all supersolutions and subsolutions gives the largest subsolution and the smallest supersolution.



This is called Perron's method. We can also find a fundamental solution in Ω, called a Green function.

1.24 Boundary Value Problem for the Laplace Equation

1.24.1 The Dirichlet and Neumann Problems

Last time, we were looking at the Laplace equation

$$-\Delta u = f \quad \text{in } \mathbb{R}^n$$

We saw a few ways to look at this:

- via the fundamental solution. This led to elliptic regularity.
- via the maximum principle. This gave us a way to prove uniqueness of solutions.
- via energy estimates. This is what we will discuss today.

When we look at the Laplace equation, we need some boundary behavior? The **Dirichlet problem** is to solve the Laplace equation with the following boundary condition.

$$\begin{cases} -\Delta u = f & \text{in } \Omega \subseteq \mathbb{R}^n \\ u = g & \text{on } \partial\Omega \end{cases}$$

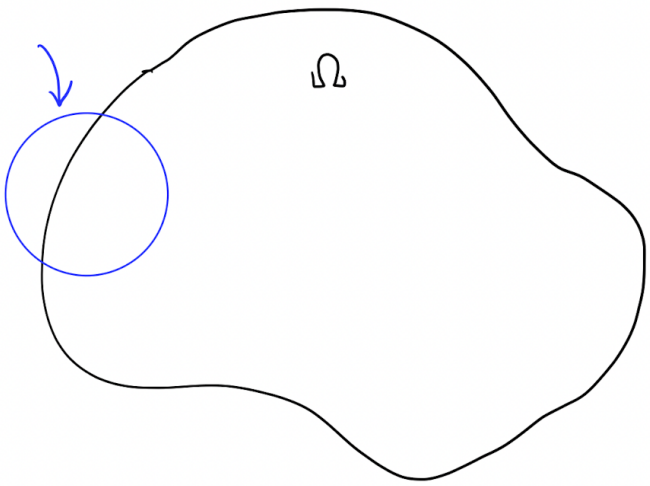
Alternatively, we can look at the **Neumann problem** with a boundary condition on the normal derivative of the solution.

$$\begin{cases} -\Delta u = f & \text{in } \Omega \subseteq \mathbb{R}^n \\ \frac{\partial u}{\partial \nu} = g & \text{on } \partial\Omega \end{cases}$$

Can we impose both the Dirichlet and Neumann boundary conditions? The answer is not always. The equation

$$\begin{cases} -\Delta u = f & \text{in } \Omega \subseteq \mathbb{R}^n \\ u = g_1 & \text{on } \partial\Omega \\ \frac{\partial u}{\partial \nu} = g_2 & \text{on } \partial\Omega \end{cases}$$

is an overdetermined problem. It makes sense to consider this locally.



This local problem will in general have uniqueness but not necessarily existence. This leads to a type of problem called a **unique continuation problem.**

1.24.2 Uniqueness concerns for the Dirichlet and Neumann problem

Proposition 1.187 (Uniqueness for the Dirichlet problem).
The solution to the Dirichlet problem

$$\begin{cases} -\Delta u = f & \text{in } \Omega \subseteq \mathbb{R}^n \\ u = g & \text{on } \partial\Omega \end{cases}$$

is unique.

Proof.
Suppose u_1, u_2 are solutions. Then $v = u_1 - u_2$ solves

$$\begin{cases} -\Delta v = 0 & \text{in } \Omega \subseteq \mathbb{R}^n \\ v = 0 & \text{on } \partial\Omega \end{cases}$$

We want to show that $v = 0$. We have

$$\begin{aligned} 0 &= \int -\Delta v \cdot v dx \\ &= \int -\partial_j \partial_j v \cdot v dx \end{aligned}$$

Green's theorem gives

$$= \int \partial_j v \dot{\partial}_j v - \int_{\partial\Omega} \frac{\partial v}{\partial \nu} \cdot \underbrace{v}_{=0} d\sigma$$

So we get

$$0 = \int_\Omega |\nabla v|^2 dx$$

which tells us that $\nabla v = 0$ in Ω. Thus, v is constant, and the boundary condition tells us that $v = 0$. □ □

Remark 1.188.
What about uniqueness of the Neumann problem?

$$\begin{cases} -\Delta u = f & \text{in } \Omega \subseteq \mathbb{R}^n \\ \frac{\partial u}{\partial \nu} = g & \text{on } \partial\Omega \end{cases}$$

By the same computation, we still get that

$$\int_\Omega |\nabla u|^2 = 0$$

which tells us that u is constant. The boundary condition is satisfied by any constant, however. So solutions are unique modulo constants.

1.24.3 Existence using energy type estimates

If $f : \mathbb{R} \to \mathbb{R}$, then a minimum point x_0 for f must have $f'(x_0) = 0$. We can do the reverse. If we have an equation for a function, we can write it as the derivative of another function and interpret our equation as finding the minimizers (or critical points) for this function. Looking at functions $u : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$, associate the functional

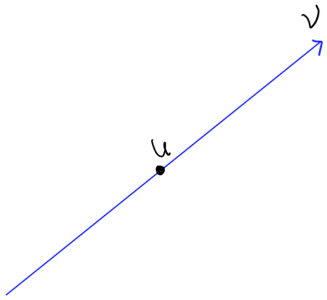
$$\mathcal{L}(u) = \int_\Omega |\nabla u|^2 - f \cdot u dx$$

We will call this the **Lagrangian** of the problem. Our goal is to minimize $\mathcal{L}(u)$ over a good class of functions u; we can assume some nice regularity and our boundary condition. Let $\mathcal{A} = \{u : \Omega \to \mathbb{R} \mid u \in C^2, u = 0 \text{ on } \partial\Omega\}$. So we want

$$\min_{u \in \mathcal{A}} \mathcal{L}(u)$$

Does a minimum exist? We will not answer this today, but observe that $\mathcal{L}$ is strictly convex because it is the sum of a positive quadratic form and a linear term. If we have a minimum, then by strict convexity, the minimum will be unique.

We may also ask: What equation does a minimum satisfy? Suppose u is a minimum. Take $v \in \mathcal{D}(\Omega)$, and set $u_h = u + hv$



Look at $\mathcal{L}(u_h)$ as a function of h. This has a minimum at $h = 0$, which tells us that

$$\frac{d}{dh}\mathcal{L}(u_h) = 0 \quad \text{at } h = 0$$

Write

$$\mathcal{L}(u_h) = \int |\nabla(u + hv)|^2 - f \cdot (u + hv)dx$$

so

$$\frac{\partial}{\partial h}\mathcal{L}(u_h)\Big|_{h=0} = \int \nabla u \cdot \nabla v - f \cdot v dx$$

Hence,

$$0 = \int \nabla u \cdot \nabla v - f \cdot v dx$$

for all $v \in \mathcal{D}(\Omega)$. Integration by parts gives us

$$= \int_\Omega -\Delta u \cdot v - fv dx$$
$$= \int_\Omega v(-\Delta u - f)dx$$

So we get

$$-\Delta = f \quad \text{in } \Omega$$

And we can append our favorite boundary condition.

Remark 1.189.
The regularity condition $u \in C^2$ is not the correct condition to use. Really, we want to use Soblev spaces, which we have not discussed yet.

1.24.4 Green's functions for domains with boundary

Circle back to the fundamental solution and try to use it in a domain with boundary. We will look at how this doesn't work and how it can be fixed. In $\mathbb{R}^n$, we have the formal computation

$$\int -\Delta u \cdot K(x - x_0)\, dx = \int u \cdot -\underbrace{\Delta K}_{\delta_{x_0}}(x - x_0)\, dx$$

If $-\Delta u = f$, then

$$\int f \cdot K(x - x_0)\, dx = u(x_0)$$

What about a domain with boundary?

$$\int_\Omega -\Delta u \cdot K(x - x_0)\, dx = \int u \cdot -\Delta K(x - x_0)\, dx + \int_{\partial\Omega} -\frac{\partial u}{\partial \nu} \cdot K(x - x_0) + u \cdot \frac{\partial}{\partial \nu} K(x - x_0)\, d\sigma$$

If u solves the Dirichlet problem

$$\begin{cases} -\Delta u = f & \text{in } \Omega \subseteq \mathbb{R}^n \\ u = g & \text{on } \partial\Omega \end{cases}$$

then

$$u(x_0) = \int_\Omega f(x) K(x - x_0)\, dx + \int_{\partial\Omega} \frac{\partial u}{\partial \nu} K(x - x_0) - g \cdot \frac{\partial K}{\partial \nu}(x - x_0)\, d\sigma,$$

but we do not know what $\frac{\partial u}{\partial \nu}$ is. We do not have this information, and recall that if we do, then we have an overdetermined problem.

How do we fix this? Perhaps the fundamental solution is not the object we want to be looking at. Replace $K(x - x_0)$ by $G(x, x_0)$ to get

$$\int_\Omega -\Delta u \cdot G(x, x_0)\, dx = \int u \cdot -G(x, x_0)\, dx + \int_{\partial\Omega} -\frac{\partial u}{\partial \nu} \cdot G(x, x_0) + u \cdot \frac{\partial}{\partial \nu} G(x, x_0)\, d\sigma$$

We would like the properties

$$\begin{cases} -\Delta_x G(x, x_0) = \delta_{x_0}, & x \in \partial\Omega \\ G(x, x_0) = 0 \end{cases}$$

If we have these properties, then

$$u(x_0) = \int_\Omega f \cdot G(x, x_0)\, dx + \int_{\partial\Omega} g \frac{\partial}{\partial \nu} G(x, x_0)\, d\sigma.$$

If we had such a function G, then we could solve the Dirichlet problem. Call this function G the **Green function** in Ω.If we are solving the Neumann problem, we may get a different Green function.

Remark 1.190.
$G = G(x, x_0)$, not $G(x - x_0)$ because translation invariance is broken by our domain. If you translate the domain with boundary, you will not get the same domain.

To find such a function G, we would try

$$G(x, x_0) = K(x - x_0) + \psi(x, x_0)$$

and look for an equation for ψ. We need

$$\begin{cases} -\Delta_x \psi(x, x_0) = 0 & \text{in } \Omega \\ \psi(x, x_0) = -K(x - x_0) & x \in \partial\Omega \end{cases}$$

$K(x - x_0)$ is smooth for $x \in \partial\Omega$ and $x_0 \in \Omega$, so by elliptic regularity, ψ should be smooth. ψ can be found by solving a Dirichlet problem.

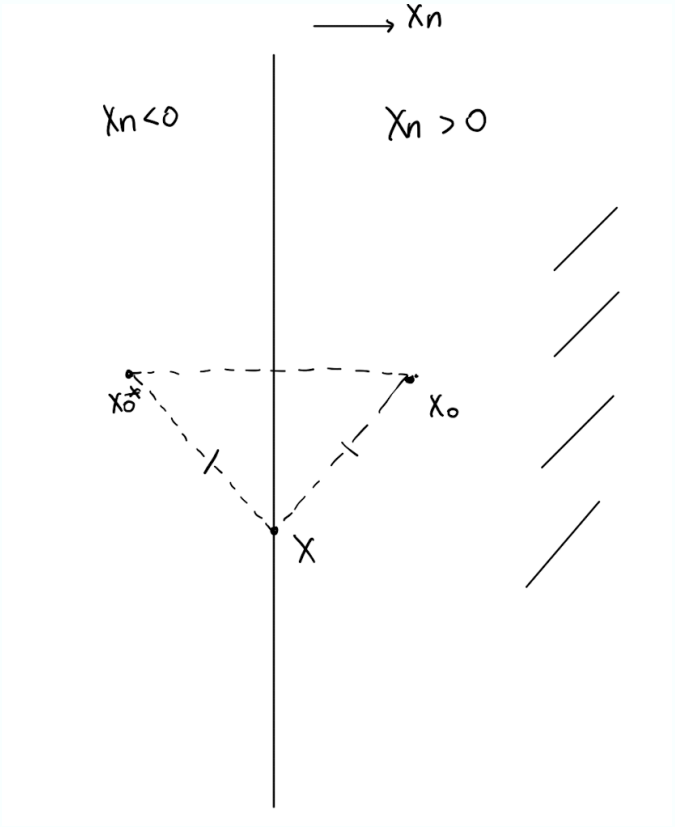
Remark 1.191.

- You may think this is leading us in a circle, but this is not the case: Here, we are solving a Dirichlet problem with a very special boundary data, $K(x - x_0)$. This may make the Green's function easier to find than solving the original equation otherwise.
- The Green function is symmetric:

$$G(x, y) = G(y, x)$$

Let's calculate some Green's functions.

Example 1.192 (Half-plane).
Let $\Omega = \{x_n > 0\}$. The idea is to calculate G using symmetries. Here is the picture we should have in mind:



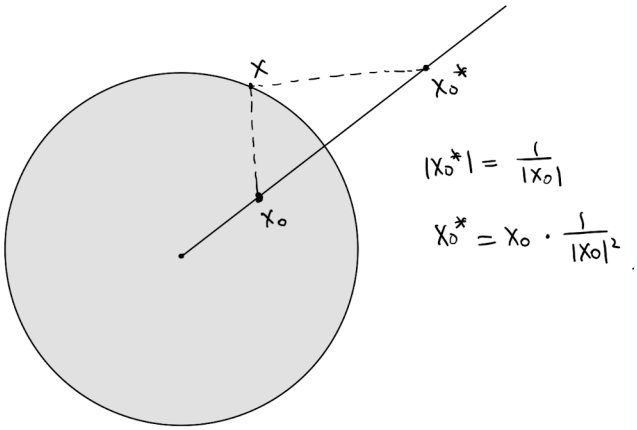
We have

$$G(x, x_0) = K(x - x_0) + \psi(x, x_0)$$

If we were to use x_0^*, then $K(x - x_0^*)$ is harmonic in Ω. Now also observe that for $x \in \partial\Omega$, $|x - x_0| = |x - x_0^*|$ for $x \in \partial\Omega$. Thus, the radial symmetry of K gives $K(x - x_0) = K(x - x_0^*)$. This implies that we can choose

$$G(x, x_0) = K(x - x_0) - K(x - x_0^*)$$

Example 1.193 (Unit ball).
Let $\Omega = B(0, 1)$. Here, we can try repeating the same argument but with inversion about the boundary of the circle:



If we have point $x \in \partial\Omega$, then

$$|x - x_0| = |x - x_0^*| \cdot |x_0|$$

So if we are in $\mathbb{R}^n$ for $n \geqslant 3$,

$$G(x, x_0) = K(x - x_0) - |x_0|^{2-n} K(x - x_0^*)$$

The proportionality constant comes from the fact that the first term is like $|x - x_0|^{2-n}$, while the second term is like $|x - x_0^*|^{2-n}$.

These kinds of computations are available only in very specific domains, so the existence of Green's functions is more of a qualitative question than a computational one.

1.25 Boundary Value Problem for the Heat Equation

1.25.1 Properties of the heat equation

Consider the heat equation in $\mathbb{R}^+ \times \mathbb{R}^n$.

$$\begin{cases} (\partial_t - \Delta)\, u = f \\ u(0) = u_0 \end{cases}$$

We have already seen how to derive a solution via the fundamental solution:

$$u = f *_{x,t} K(t) + u_0 *_x K(t), \quad K(t) = \frac{1}{(4\pi t)^{n/2}} e^{-x^2/(4t)} \mathbb{1}_{\{t \geqslant 0\}}$$

This is the unique solution going forward in time which is a temperate distribution.

Here are some key properties for the homogeneous equation given by this fundamental solution: Consider the heat equation in $\mathbb{R}^+ \times \mathbb{R}^n$.

$$\begin{cases} (\partial_t - \Delta)\, u = 0 \\ u(0) = u_0 \end{cases}$$

- Infinite speed of propagation: Even if u_0 has compact support, the solution u immediately spreads to all of $\mathbb{R}^n$.
- Instant regularization:
$$u(t) = K(t) * u_0$$
where $K(t)$ is smooth for $t > 0$. So u is smooth for $t > 0$.
- The fundamental solution has Gaussian decay at ∞ : This means that any initial data u_0 with $|u_0| \leqslant e^{cx^2}$ will generate a local in time solution.

1.25.2 The mean value property and the maximum principle

Now let's look at the heat equation in a domain $\Omega \subseteq \mathbb{R}^n$.

$$\begin{cases} (\partial_t - \Delta)\, u = f & \text{in } \Omega \times \mathbb{R}^+ \\ u(t=0) = u_0 & \text{in } \Omega \\ u(t,x) = g & \text{on } \partial\Omega \times \mathbb{R}^+ \end{cases}$$

The third equation is a **Dirichlet boundary condition**. We could replace it with a **Neumann boundary condition**

$$\frac{\partial u}{\partial \nu}(t,x) = g \quad \text{on } \partial\Omega \times \mathbb{R}^+$$

As with the Laplace equation, we use either one boundary condition or the other but not both. Here are several ways to approach this:

- Via a maximum principle.
- Via energy estimates.
- Using Green's functions.

We first discuss the maximum principle. First, is there a mean value property for the heat equation? We would like to write something like

$$u\,(t_0, x_0) = \frac{1}{|D|} \int_D u(t,x)\, dt dx$$

for some D. For the Laplace equation, we used a ball for D, but this should not be the case for the heat equation; unlike for the Laplace equation, balls are not level sets of the fundamental solution. We may also ask if we need any weights for the maximum principle.

Step 1: Green's theorem for the heat equation: Let u, v be such that v has compact support. Then

$$\iint (\partial_t - \Delta)\, u \cdot v dx dt = \iint (-\partial_t - \Delta)\, v \cdot u dx dt$$

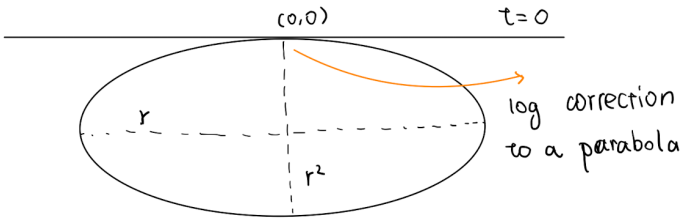
If we want to get $u(0,0)$ out of the right hand side, then we would need $(-\partial_t - \Delta)\, v = \delta_{(0,0)}$. Here, $-\partial_t - \Delta$ is the adjoint heat operator, which is a "backward heat operator" and gives a backward heat equation with a fundamental solution

$$K^{\text{back}}(x,t) = -\frac{1}{(4\pi|t|)^{n/2}} e^{x^2/(-4t)} 1_{\{t \leqslant 0\}}$$

Define the parabolic balls

$$D_r(0,0) = \left\{ \left| K^{\text{back}}(x,t) \right| \geqslant r^{-n} \right\}$$

What do these sets look like? If $x = 0$, then $K \asymp t^{-n/2}$, and $t^{-n/2} \geqslant r^{-n}$ iff $|t| \leqslant r^2$. To figure out the sideways boundaries of these regions, take $t \approx -\frac{1}{2} r^2$. Now change x so that $e^{x^2/(4t)} \gtrsim 1$. Then $|x| \leqslant \sqrt{-t} \approx r$. This looks like an ellipse, but near $(0,0)$, there is a logarithmic coorection to a parabola.



Our goal is to show that

$$u(0,0) = \int_{D_r(0,0)} \omega(t,x) u(t,x) dx$$

for some suitable positive weight ω (we want positive so we can think of this as an average). Look at our Green's theorem in $D_r(0,0)$, which gives boundary terms:

$$\iint\limits_{D_r(0,0)} (\partial_t - \Delta)\, u \cdot v dx dt = \iint\limits_{D_r(0,0)} (-\partial_t - \Delta)\, v \cdot u dx dt + \int_{\partial D_r(0,0)} \nu_t \cdot uv - \frac{\partial u}{\partial \nu} \cdot v + u \cdot \frac{\partial v}{\partial \nu} d\sigma$$

For $v = K^{\text{back}}(t,x)$, this does not work because we get boundary terms. Instead, we can try

$$v = K^{\text{back}}(t,x) + r^{-n}$$

which makes $v = 0$ on $\partial D_r(0,0)$. This makes the first two boundary terms equal 0 , but we would also like to make sure that $\frac{\partial v}{\partial \nu} = 0$ on $\partial D_r(0,0)$. This is the same as saying that $\nabla v = 0$ on ∂D_r. The way we can alter our fundamental solution to take advantage of this is

$$v = K^{\text{back}}(t,x) + r^{-n} + c \ln\left(-K^{\text{back}} \cdot r^n\right)$$

where c is chosen so that $\nabla v = 0$ on $\partial D_r(0,0)$. This choice gives us

$$\begin{aligned} \nabla v &= \nabla K + c \frac{\nabla K}{K} \\ &= \nabla K \left(1 + \frac{c}{K}\right) \end{aligned}$$

and since $K = -r^n$ on the boundary, we can pick $c = r^n$. If $(\partial_t - \Delta)\, u = 0$, then we get

$$\iint D_t(-\partial - \Delta)v \cdot u dx dt = 0$$

We can calculate

$$\begin{aligned}
(-\partial_t - \Delta)\, v &= \delta_{(0,0)} + c\,(-\partial_t - \Delta) \ln\left(-r^n K^{\text{back}}\right) \\
&= \delta_{(0,0)} - c\frac{\partial_t K^{\text{back}}}{K^{\text{back}}} - c\nabla \cdot \frac{\nabla K^{\text{back}}}{K^{\text{back}}} \\
&= \delta_{(0,0)} - c\underbrace{\frac{(\partial_t - \Delta)\, K^{\text{back}}}{K^{\text{back}}}}_{=0} + c\frac{\left(\nabla K^{\text{back}}\right)^2}{\left(K^{\text{back}}\right)^2} \\
&= \delta_{(0,0)} + c\left(\nabla \ln K^{\text{back}}\right)^2
\end{aligned}$$

where this is a spatial gradient.

$$= \delta_{(0,0)} - r^{-n}\frac{x^2}{4t^2}$$

We get:

Theorem 1.194 (Mean value property).
If $(\partial_t - \Delta)\, u = 0$ in $\Omega \times [0,T]$,

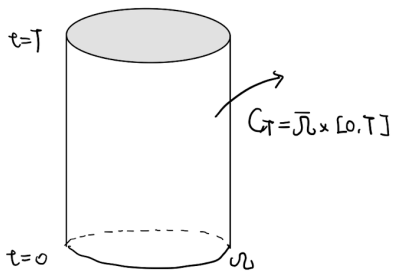
$$u(0,0) = r^{-n}\int_{D_r(0,0)} \frac{x^2}{4t^2} u(t,x) dx dt$$

Remark 1.195.
How do we know this is an average? This holds for all solutions to the heat equation, so plug in a constant. This gives

$$r^{-n}\int_{D_r(0,0)} \frac{x^2}{4t} dx dt = 1$$

So this is indeed a weighted average.

For our maximum principle, what is the boundary of our region $C_T = \bar{\Omega} \times [0,T]$?



If you consider causality, the $t = T$ boundary is determined by the rest, so it should not be considered. Write $\partial C_T = \bar{\Omega} \times \{0\} \cup \partial\Omega \times [0,T]$. The first part is the bottom, and the second part is the lateral boundary. Together, they make up the parabolic boundary of C_T.

Theorem 1.196 (Strong maximum principle).
If $(\partial_t - \Delta)\, u = 0$ in $\Omega \times [0,T]$, then

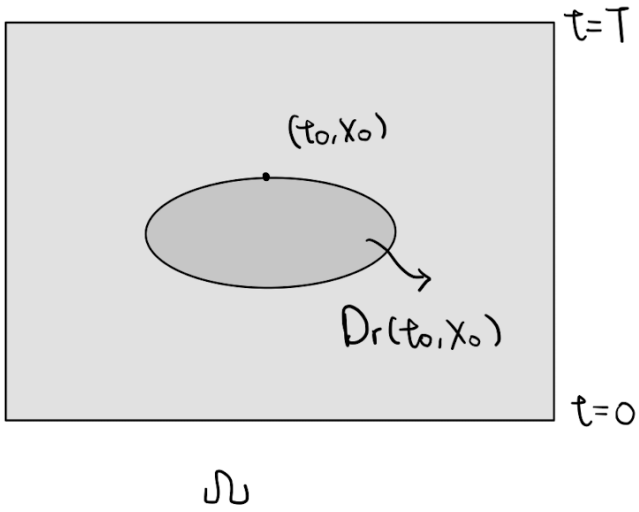
$$\max_{C_T} u = \max_{\partial C_T} u.$$

Further if $u\,(t_0, x_0) = \max u$ for some (t_0, x_0) inside, then u is constant for $t \leqslant t_0$.

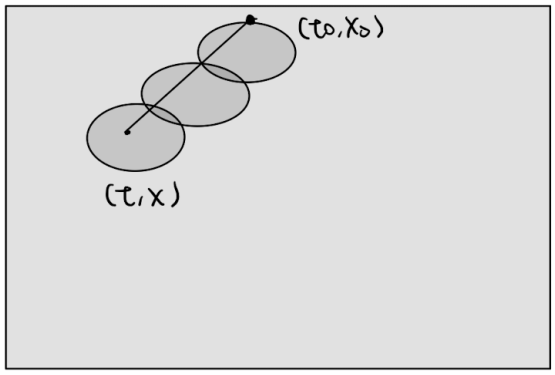
Proof.
Take (t_0, x_0) to be a maximum inside. Then the mean value property gives

$$\begin{aligned}\max u &= u(t_0, x_0) \\ &= r^{-n} \int \frac{(x - x_0)^2}{(t - t - 0)^2} u(t, x) dx dt \\ &\leqslant r^{-n} \int \frac{(x - x_0)^2}{(t - t - 0)^2} \max u dx dt \\ &= \max u\end{aligned}$$

Equality must hold, so $u = \max u$ in $D_r(t_0, x_0)$.



How do we get the whole region $t \leqslant t_0$? Here is a picture:



□ □

Remark 1.197.
Just like with the Laplace equation, we can talk about subsolutions

$$(\partial_t - \Delta) u \leqslant 0$$

and supersolutions

$$(\partial_t - \Delta) u \geqslant 0$$

Using the mean value property with inequalities gives a maximum principle for subsolutions and a minimum principle for super solutions.

Theorem 1.198 (Comparison principle).
Let u^-be a subsolution and u^+be a supersolution for the heat equation. If $u^- \leqslant u^+$on ∂C_T, then $u^- \leqslant u^+$in C_T.

Corollary 1.199.
The solution to the bounded heat equation is unique.

1.25.3 Energy estimates

Consider the homogeneous Dirichlet problem

$$\begin{cases} (\partial_t - \Delta)\, u = 0 & \text{in } \Omega \times [0, T) \\ u(0) = u_0 \\ u(\partial\Omega) = 0 \end{cases}$$

and let

$$E(u(t)) = \int |u(t,x)|^2 dx$$

Then we can compute

$$\begin{aligned} \frac{\partial}{\partial t} E(u(t)) &= 2 \int u \cdot u_t dx \\ &= 2 \int u \cdot \Delta u dx \\ &= -2 \int |\nabla u|^2 \, dx \\ &\leqslant 0, \end{aligned}$$

which tells us that E is nonincreasing in time $E(t) \leqslant E(0)$. So if $u_0 = 0$, then $E(t) = 0$, which gives $u(t) = 0$. We can also look at the relation

$$\|u(0)\|_{L^2}^2 = \|u(T)\|_{L^2}^2 + \int_0^t |\nabla u|_{L^2}^2 dx$$

If we start with $u(0) \in L^2$, we get $u(T) \in L^2$ for a.e. T. We can think of this as a **parabolic regularizing effect**.

1.26 Initial Value Problems and Energy Estimates for the Wave Equation

1.26.1 Initial value problems for the wave equation

Today, we will be looking at the wave equation

$$\begin{aligned} \Box u &= f \quad \text{in } \mathbb{M}^{n+1} = \mathbb{R} \times \mathbb{R}^n \\ u(0) &= u_0 \\ \partial_t u(0) &= u_1 \end{aligned}$$

where

$$\Box = \partial_t^2 - \Delta_x = -m^{\alpha\beta}\partial_\alpha \partial_\beta, \quad m = \begin{bmatrix} -1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix}.$$

We have seen that the fundamental solution (forward in time) is

$$K(t,x) = \begin{cases} \frac{1}{2} 1_{\{t>|x|\}} & n = 1 \\ c_n \left(t^2 - x^2\right)_+^{(1-n)/2} & n \geqslant 2 \text{ even} \\ c_n \delta_{t^2-x^2}^{\left(\frac{n-1}{2}\right)} & n \geqslant 2 \text{ odd} \end{cases}$$

The solution for the inhomogeneous problem is $u = K * f$ (as if the Cauchy data equals 0 at $-\infty$). The solution for the homogeneous problem ($f = 0, u_0, u_1 \neq 0$) is a bit more tricky. Let

$$\tilde{u} = \begin{cases} u & t > 0 \\ 0 & t < 0 \end{cases}$$

Let's find an equation for $\tilde{u}$.

$$\Delta \widetilde{u} = \begin{cases} \Delta u & y \geqslant 0 \\ 0 & t < 0, \end{cases}$$

$$\partial_t \widetilde{u} = \begin{cases} \partial_t u & y > 0 \\ 0 & t < 0, \end{cases} + \delta_{t=0} \cdot u_0.$$

The second time derivative is then

$$\partial_t \widetilde{u} = \begin{cases} \partial_t^2 u & y > 0 \\ 0 & t < 0, \end{cases} + \delta_{t=0} \cdot u_1 + \delta'_{t=0} \cdot u_0.$$

This gives us

$$\Box \widetilde{u} = \delta_{t=0} u_1 + \delta'_{t=0} u_0$$

which implies that

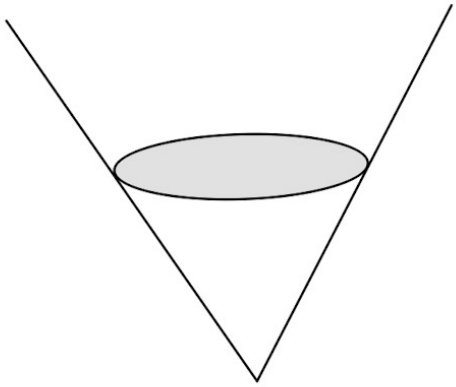
$$\widetilde{u} = K *_{t,x} \left(\delta_{t=0} u_1 + \delta'_{t=0} u_0\right).$$

Taking the convolution first in t gives

$$\widetilde{u}(t) = K(t) *_x u_1 + \partial_t K(t) *_x u_0.$$

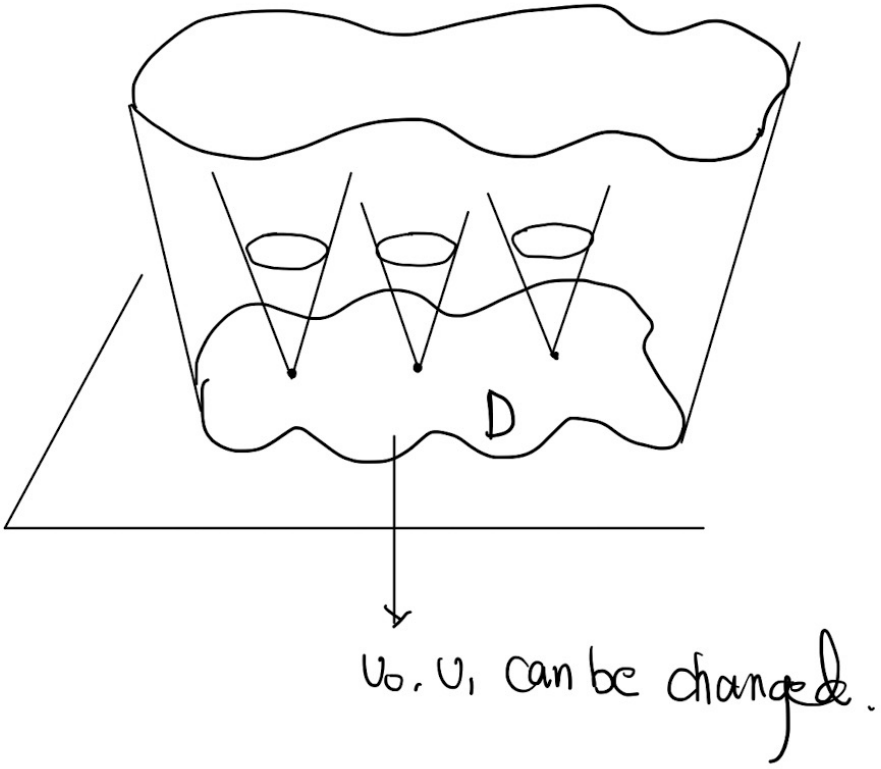
Here are some properties of the wave equation:

- Finite speed of propagation: The solution only exists inside the positive cone.

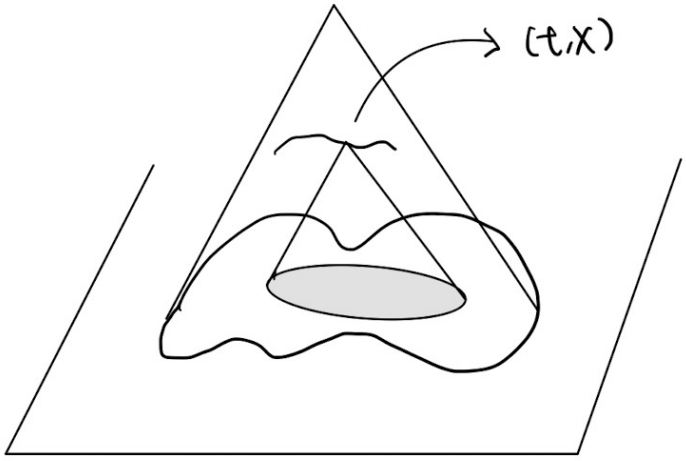


- Huygens principle: When $n \geqslant 3$ is odd, the fundamental solution is supported exactly on the cone.

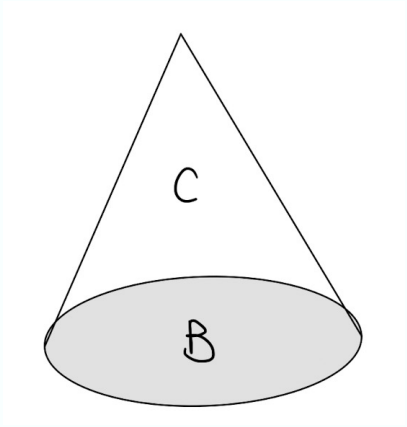
Suppose now that we have some region with initial data (u_0, u_1) which can be changed. Where does the solution change? At each point, we have an upward cone, and we take the union of these cones.



The **domain of influence** is $\Omega = \cup_{x \in D}\{(0, x) + K\}$, where K is the forward cone. Now suppose we only have initial data (u_0, u_1) in the domain D. Where can we find the solution? If we look at a point (t, x), then $u(t, x)$ depends on u_0, u_1 in $B(x, |t|)$. The value $u(x, t)$ is uniquely determined if $B(x, |t|) \subseteq D$. The union, $\{(t, x) : B(x, |t|) \subseteq D\}$ is called the **domain of uniqueness** for D.



Example 1.200.
When the base domain B is a ball, then the domain of uniqueness C is exactly a cone:



1.26.2 Energy estimates for the wave equation

Here's how energy estimates work for the wave equation. When we say energy, we want to think a quantity which is conserved. Suppose we have a vibrating string.

We can We can think of the energy as potential energy P, expressed in terms of "how extended is the string." This can be measured by some average of the slope of the string:

$$P = \int |\partial_x u|^2 \, dx.$$

The second part of the energy should be the kinertic energy, which measures the velocity of the string:

$$\int |\partial_t u|^2 \, dx$$

If we were physicists, we would have constants in front, but we are mathematicians, so we will set some constants equal to 1. We can write the total energy as

$$E(u(t)) = \frac{1}{2} \int |\partial_t u|^2 + |\nabla_x u|^2 \, dx$$

Theorem 1.201.
If $\Box u = 0$, then $E(u(t))$ is constant.

Proof.

$$\begin{aligned}\frac{d}{dt}E &= \int \partial_t u \partial_t^2 u + \sum_{j=1}^{n} \partial_j u \cdot \partial_t \partial_j u dx \\ &= \int \partial_t u \sum_{j=1}^{n} \partial_j \partial_j u + \sum_{j=1}^{n} \partial_j u \cdot \partial_t \partial_j u dx \\ &= 0\end{aligned}$$

By Green's theorem, assuming that $u = 0$ at ∞. □

Why should we improve on this? We have seen that "conservation laws" imply features of our problem. If we have

$$\partial_t \underbrace{u}_{\text{density}} + \partial_x \underbrace{F(u)}_{\text{flux}} = 0$$

which tells us that

$$\partial_t \int u \, dx = 0$$

For the wave equation, we have the energy density

$$e(t,x) = \frac{1}{2}\left|\partial_t u\right|^2 + \left|\nabla_x u\right|^2,$$

so that

$$E = \int e$$

Note that this doesn't go the other way around; there may be many densities that integrate to the same total energy E. We can look at

$$\begin{aligned}\partial_t e(t,x) &= \partial_t u \cdot \partial_t^2 u + \partial_j u \cdot \partial_t \partial_j u \\ &= \partial_t u \cdot \Box u + \partial_t u \cdot \partial_j^2 u + \partial_j u \partial_t \partial_j u \\ &= \partial_j (\underbrace{\partial_t u \cdot \partial_j u}_{\text{energy flux}}) + \partial_t u \cdot \Box u\end{aligned}$$

This leads us to another proof of the energy estimates for the wave equation:

Proof.
Start with $\Box u = 0$, and get $\Box u \cdot \partial u = 0$. Then integrate by parts. □ □

Let's see what happens when we take

$$\begin{aligned}\Box u \cdot \partial_k u &= \left(\partial_t^2 u - \partial_j \partial_j u\right) \cdot \partial_k u \\ &= \partial_t \left(\partial_t u \cdot \partial_k u\right) - \partial_t u \cdot \partial_t \partial_k u - \partial_j \left(\partial_j u \cdot \partial_k u\right) + \partial_j u \cdot \partial_j \partial_k u\end{aligned}$$

We can think of the first and third terms as derivatives.

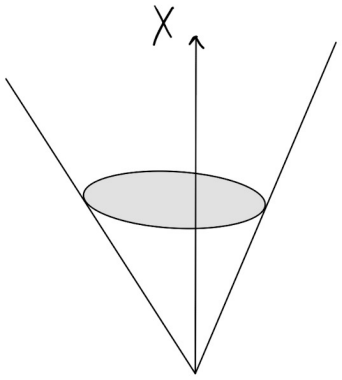
$$= \partial_t (\underbrace{\partial_t u \cdot \partial_k u}_{\text{density}}) - \underbrace{\frac{1}{2}\partial_k \left(\partial_t u\right)^2 - \partial_j \left(\partial_j u \partial_k u\right) + \frac{1}{2}\partial_k \left(\partial_j u\right)^2}_{\text{divergence of a flux}}$$

We get a new, conserved quantity, the momentum

$$P_k = \int \partial_t u \cdot \partial_k u dx$$

This tells you in what direction the energy is moving around. Conservation says that if the energy is moving around in one direction, it will be moving in that same direction forever.

More generally, consider $\Box u \cdot Xu$, where $X = \sum X^\alpha \partial_\alpha$ This gives a conserved quantity E_X, which is positive definite if the vector field X is **forward time-like**.



Remark 1.202.
We can put the energy and the momentum into one conserved quantity, called the energy-momentum tensor,

$$T^{\alpha\beta}(u) = \partial^\alpha u \partial^\beta u - \frac{1}{2} m^{\alpha\beta} \partial^\gamma u \partial_j u$$

where

$$\partial^\alpha = m^{\alpha\beta}\partial_\beta u$$

For the wave equation, this looks like

$$\partial^0 u = -\partial_0 u, \quad \partial^j u = \partial_j u.$$

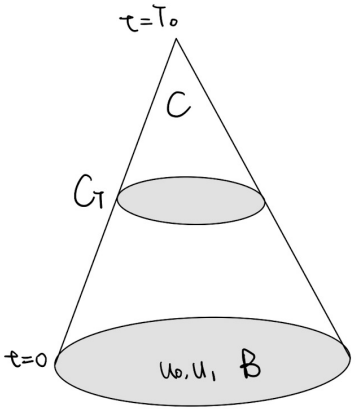
This is a divergence free tensor:

$$\partial_\alpha T^{\alpha\beta} = 0 \quad \forall \beta \quad \text{if} \quad \Box u = 0.$$

If $\beta = 0$, this is the energy, and if $\beta = j \neq 0$, this is the momentum P_j.

1.26.3 Finite speed of propagation via energy estimates

The finite speed of propagation is a robust phenomenon. We can show this by providing a proof which does not rely on the fundamental solution and only requires energy estimates. If we have a ball B for our inital data, and a cone C, we want to show that (u_0, u_1) in B uniquely determines the solution in C.



This is the same as saying that if $(u_0, u_1) = (0, 0)$ in B, then $u = 0$ in C. Suppose we want to show that $u = 0$ in the slice C_T of the cone. We saw the following density flux relation for the energy:

$$\partial_t e(t, x) = \partial \left(\partial_t u \cdot \partial_u \right)$$

Integrate over $C_{[0,T]}$, the section of the cone up to time T.

$$\int_{C_T} e - \int_{C_0} + \int_{\partial C_{[0,T]}} e \cdot \nu_t - p_j \nu_x = 0$$

Moving the middle term to the right hand side, this tells us that

$$\text{Energy}(t = 0) = \text{Energy}(t = T) + \text{ Flux (boundary)}$$

The former term is the part that is left in the cone, and the latter term is the part that goes out. If the energy at time $t = 0$ is 0, then these two terms must both equal 0.

The remaining objective is to show that the Flux term is nonnegative. What does it mean that the slope of the cone is -1 ? This means that the outward pointing normal is $\nu = (1, \omega)$ with $|\omega| = 1$. Then

$$\begin{aligned} e \cdot \nu_t - p_j \cdot \nu_j &= \frac{1}{2} \left(u_t^2 + |\nabla_x u|^2 \right) - \partial_t u \cdot \partial_j \cdot \omega_j \\ &\overset{?}{\geqslant} 0 \end{aligned}$$

We can use Cauchy-Schwarz twice to say

$$\begin{aligned} |\partial_t u \cdot \partial_j u \cdot \omega_j| &\leqslant |\partial_t u| \cdot |\partial u| \\ &\leqslant \frac{1}{2} \left(|\partial_t u|^2 + |\partial u|^2 \right) \end{aligned}$$