Partial Differential Equations Notes

Professor: Jalal Shatah Eduardo Corona

Spring 2009

Contents

Part I
Introduction

1 1st order PDEs: The Method of Characteristics

Lecture 1 (15/01/09)

1.1 Linear PDE

We can start by thinking of a set of particles moving in $1D$, their motion given by the equation:

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

Additionally, there might be a quantity, or "piece of information" that is carried by these particles and that we are interested in measuring (e.g. charge, mass, velocity, etc), denoted by u. If this information is not changing with time, then we have the following:

$$0 = D_t u = \frac{d}{dt} u(\gamma(t)) = \frac{du}{dt} + v\frac{du}{dx}$$

This is a typical advection equation.

We can now consider the inverse process: we have a function u such that it satisfies the advection equation:

$$\left\{ \begin{array}{c} u_t + vu_x = 0 \\ u(x,0) = f(x) \end{array} \right\}$$

For each point $x_0 \in \mathbb{R}$, we can then solve the ODE to obtain $x = \gamma(t, x_0)$:

$$\left\{ \begin{array}{c} \dot{x} = v(x,t) \\ x(0) = x_0 \end{array} \right\}$$

By our hypothesis, $D_t u = 0$, and so:

$$u(x,t) = f(x_0)$$

$$x_0 = \gamma^{-1}(t,x)$$

$$u(x,t) = f(\gamma^{-1}(t,x))$$

Example 1 *We want to solve*

$$\left\{ \begin{array}{c} u_t + u_x = 0 \\ u(0,x) = \sin x \end{array} \right\}$$

$v \equiv 1$, and the solution to the ODE $\dot{x} \equiv 1$ with $x(0) = x_0$ is $x = t + x_0$. Hence,

$$x_0 = x - t$$

$$u(x,t) = \sin(x-t)$$

This is basically telling us that the sine wave is being transported with speed 1.

Example 2 *Modifying the previous equation, we are now interested in:*

$$\left\{ \begin{array}{c} u_t + u_x = u \\ u(0,x) = f(x) \end{array} \right\}$$

Again, the trajectory is given by $x = x_0 + t$, but $D_t u = u$. So,

$$u(x,t) = f(x-t)e^t$$

Example 3 *Now, we try to solve a general linear* 1*st order PDE using this method:*

$$a(x,y)u_x + b(x,y)u_y = c(x,y)$$

We could assume $a \neq 0$ *and/or* $b \neq 0$ *and solve it like the previous examples. However, it is best to solve an ODE directly. For that matter, we introduce a new parameter* τ*:*

$$\left\{ \begin{array}{c} \frac{dx}{d\tau} = a(x,y) \quad , \quad \frac{dy}{d\tau} = b(x,y) \\ (x,y)|_{\tau=0} = (x_0,y_0) \end{array} \right\}$$

By solving this ODE, I get a curve $(x,y) = (\alpha,\beta)(\tau)$*, and the equation simplifies to:*

$$D_\tau u = c(\alpha,\beta)$$

Hence, an integration formula yields:

$$u = u_0 + \int_0^\tau c(\alpha(\sigma),\beta(\sigma))d\sigma$$

The last step involves solving for (x_0,y_0) *and plugging them in.*

1.2 Nonlinear 1st order PDEs

We now proceed to the next level of complication: a semilinear 1*st* order PDE. That is,

$$a(x,y,u)u_x + b(x,y,u)u_y = c(x,y,u)$$

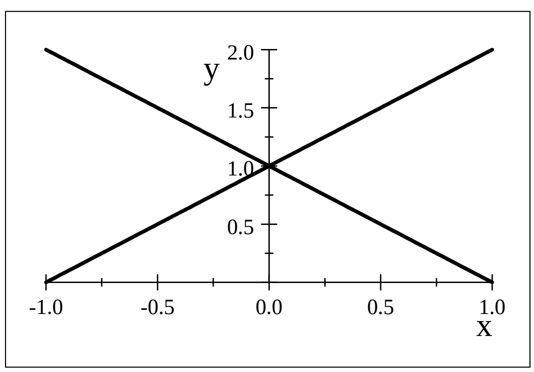
Example 4 *We start with a simple example:*

$$\left\{ \begin{array}{c} u_t + uu_x = 0 \\ u(0,x) = f(x) \end{array} \right\}$$

Applying our method, we then have to solve the ODE system

$$\left\{ \begin{array}{c} \dot{x} = u \\ \dot{u} = 0 \end{array} \right\}$$

Consider two points $x_0 < x_1$ *(say* $x_0 = -1$ *and* $x_1 = 1$*) with* $f(x_0) = 1$ *and* $f(x_1) = -1$*.* $x+1$



On each curve, u *is constant and equal to* $f(x_i)$*, and so the characteristic curves can intersect if (as in this case) their slopes have opposite signs. This creates a conflict at the points of intersection: because the solution is given implicitly, we run into an issue (can't define* u *uniquely).*

This example makes us think about the problem we have been posing: we want to obtain a global solution from local data. In the case where characteristic curves intersect, it is clear that the solution exists only for a little bit of time, at best.

Now, if we blindly solve for the system, we get:

$$x = ut + x_0$$

$$u = f(x - ut)$$

From this implicit equation, we must solve for u. Let $g(x, t; u) = u - f(x - ut)$. By the Implicit Function Theorem, in order to be able to solve for u (locally), we require:

$$g_u = 1 + tf'(x - ut) \neq 0$$

So, for instance if $f' > 0$ or t is small, a solution always exists.

This equation we have just studied is important, since it is a prototype for conservation laws.

1.2.1 What are conservation laws?

Let $\chi(x, t)$ describe the family of characteristic curves, q a quantity that is conserved over time (mass, charge, momentum, energy, etc), given by a density $u(x, t)$:

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

Now, we take a derivative with respect to time, and let $\dot{\chi} = v$ (velocity):

$$0 = \int_{\chi(x_0,t)}^{\chi(x_1,t)} u_t(x, t)dx + vu|_{\chi(x_0,t)}^{\chi(x_1,t)}$$

Since x_0 and x_1 are arbitrary, this implies:

$$\left\{ \begin{array}{c} u_t + (vu)_x = 0 \\ u(x, 0) = f(x) \end{array} \right\}$$

Sometimes v depends on u (quasilinear case) or is given (linear case).

Case 5 *If v is given, we have:*

$$u_t + vu_x = -v_x u$$

The ODE is $\dot{x} = v$, and once we solve it, we plug into $D_t u = -v_x u$, which is separable:

$$\ln(u) = -\int v_x(x, t)dt$$

$$u(x, t) = \exp\left(-\int v_x(x, t)dt\right) f(x_0(x, t))$$

We can rewrite the solution to this equation in an interesting way. The solution to our system is given on the curves $x(t, x_0)$. But then, differentiating $\dot{x} = v$ yields:

$$\left\{ \begin{array}{c} \frac{d\dot{x}}{dx_0} = v_x(x, t)\frac{dx}{dx_0} \\ \frac{d\dot{x}}{dx_0}(0) = 1 \end{array} \right\}$$

And so,

$$\frac{dx}{dx_0} = \exp\left(\int v_x(x, t)dt\right)$$

So,

$$u = \frac{1}{\frac{dx}{dx_0}} f(x_0)$$

This gives us insight to generalize this result to higher dimensions:

$$u = \frac{1}{|J|} f(x_0)$$

Which makes sense, since the answer comes directly from the change of variables between x_0 and x (this is strongly related with the difference between eulerian and lagrangian coordinates in fluid dynamics).

Now, the fact that first order PDEs can be solved explicitly using these methods is an anomaly: this is almost never the case for higher order PDEs (we then turn to numerical methods...).

Before proceeding, it is first relevant to answer some basic questions:

1. What kind of problems are solvable? What does it mean for a problem to be **well-posed** (make sense)?

We have seen problems for which we have a global, unique solution (e.g. the linear transport PDE). A problem which is ill-posed is such that either the solution via characteristic curves contradicts the data, or in which characteristic curves intersect.

2 What do we mean by a solution?

Although I write a differential equation with derivatives, the solution in its most general form is that of an integral equation. This is known as a weak solution, and the notion is well rooted in the fact that most conservation laws and variational inequalities are originally cast in this form, and then the differential equation per se can be derived assuming enough regularity. The theory of weak solutions is made rigorous through the study of Distributions and Sobolev Spaces.

3 Classification of PDEs into elliptic, parabollic, hyperbollic, etc

Lecture 2 (22/01/09)

On the previous lecture, we introduced the method of characteristic curves to handle linear and semilinear 1st order PDEs. For the most general quasilinear case, we introduce a parameter τ, and solve the following system of ODEs:

$$\left\{ \frac{dx}{d\tau} = a(x,y,z), \frac{dy}{d\tau} = b(x,y,z), \frac{dz}{d\tau} = c(x,y,z) \right\}$$

The solution for a particular point $(x_0, y_0, u(x_0, y_0))$ is a curve in the (x, y, z) plane.

Claim 6 *The surface $z = u(x, y)$ generated by these characteristic curves then is a solution to the PDE. Conversely, if $z = u(x, y)$ is a given classical solution to the PDE and $\gamma(\tau)$ is a characteristic such that at $\tau = 0$, $z_0 = u(x_0, y_0)$, then $\gamma(\tau)$ lies completely on the surface.*

Proof. (Sketch) (1) The tangent to the characteristic curve at the point (x', y', z') is given by (a, b, c), by our assumption. We also know that the surface $z = u(x, y)$ has a normal vector $(u_x, u_y, -1)$. But then, it follows that:

$$(u_x, u_y, -1) \cdot (a, b, c) = 0$$

Which returns the PDE.

(2) Now, let $\gamma(\tau) = (x, y, z)(\tau)$ be the characteristic curve. We know it touches the surface at (x_0, y_0, z_0), but we do not know if $z(\tau) = u(x(\tau), y(\tau))$ for $\tau > 0$.

Let $w(\tau) = z(\tau) - u(x(\tau), y(\tau))$. Now,

$$\begin{aligned}
\frac{dw}{d\tau} &= c(x,y,z) - u_x a - u_y b \\
&= c(x,y,z) - c(x,y,u(x,y)) \\
&= c(x,y,z) - c(x,y,z-w)
\end{aligned}$$

This is then an ODE for w, for which $w \equiv 0$ is a solution. By the existence and uniqueness of solutions, it then follows that $w(\tau) \equiv 0$. ■

Now, when can I construct a solution of the PDE from the ODE? We can deduce this from the conditions for the Implicit Function Theorem: we originally have the surface $(x, y, z)(\tau, s)$, and I want to write it as $z = u(x, y)$. This is possible if and only if:

$$\det \begin{pmatrix} x_\tau & y_\tau \\ x_s & y_s \end{pmatrix} \neq 0$$

I can compute this at the initial point, with $(x, y) = (f(s), g(s))$, $z = h(s) = u(x, y)$:

$$\det \begin{pmatrix} a & b \\ f_s & g_s \end{pmatrix} \neq 0$$

This essentially says the vectors (a, b) and (f_s, g_s) cannot be parallel. In other words, **the PDE is solvable by this method (locally) if and only if the characteristic curves "take off" from the curve of initial values.** In other words, we may not impose initial data on a curve that is characteristic (or with a tangent vector which is parallel to a characteristic at some point).

Theorem 7 *Given a quasilinear equation $au_x + bu_y = c$ (with a, b, c sufficiently smooth), $C = (f, g)(s) \in C^1$ and "Cauchy data" $u|_C = h(s)$, then there exists a unique solution in a neighborhood of C provided the "non-characteristic" condition holds.*

$$\det \begin{pmatrix} a & b \\ f_s & g_s \end{pmatrix} \neq 0$$

Example 8 *For $u_t + uu_x = 0$, $u|_{t=0} = h(x)$, the curve $\{t = 0\}$ is never characteristic, since $\frac{dt}{ds} \equiv 1$.*

Example 9 *$u \equiv 1$, for $0 \leq x \leq 1$. This is also solvable, but we can only tell what u is in the region swept by the characteristics $\{(x, y) : x - 1 \leq y \leq x\}$.*

1.3 Generalization to higher dimensions

For $x \in \mathbb{R}^d$ we have the following quasilinear PDE (where A and C are sufficiently smooth):

$$A(x, u)^\top \nabla u = C(x, u)$$

We can then obtain the $d + 1$ dimensional ODE system:

$$\left\{ \frac{dx}{d\tau} = A, \frac{dz}{d\tau} = C \right\}$$

As before, we solve this from an initial point $(x_0, u(x_0))$, and call it a characteristic curve. Everything follows verbatim, except now the non-degeneracy / non-characteristic condition is more complicated.

A Cauchy problem now consists of providing initial data on a $d - 1$ dimensional hypersurface, which can be either written as $X(0, y_1, .., y_{d-1})$ (determined by $d - 1$ parameters) with nonzero Jacobian or as a level set $\varphi(x) = 0$ with $\nabla \varphi \neq 0$.

We then solve the ODE on each point of this surface S, obtaining $\{X(\tau, y), Z(\tau, y)\}$. Although the IFT again yields a condition (the Jacobian being nonzero), but it is usually very burdensome to check.

Instead, in the case where the surface is given by a level set, it is better to usea change of variables $(\alpha_1, .., \alpha_d) = (x, \varphi(x))$. So,

$$A^\top \nabla u = \sum_{i,k} A^i \frac{\partial u}{\partial \alpha_k} \frac{\partial \alpha_k}{\partial x_i}$$

Then our hypersurface becomes $\alpha_d = 0$, and the non-degeneracy condition now reads:

$$\sum_i A^i \frac{\partial \varphi}{\partial x_i} \neq 0$$

Since the normal to the hypersurface is $\nabla \varphi$, this is a straightforward extension of the non-characteristic condition in $1D$: characteristics must have a non-trivial normal component to the surface S.

1.3.1 Power Series

The first time you see an ODE, what do you try to do? We can either integrate, or try to come up with a power series solution. Lets try this for PDEs:

Example 10 *Lets solve (again) $u_t + u_x = 0$ with $u(0,x) = \sin(x)$. We first write $u(x,t) = \sum_{n=0}^{\infty} g_n(x)\frac{t^n}{n!}$, and from the PDE and initial conditions:*

$$u_t|_{t=0} = -u_x|_{t=0} = -\cos(x)$$

Differentiating once on time, we then have $u_{tt}(0) = -u_{xt}(0) = -\sin(x)$, and we can continue this process. Thus,

$$\begin{aligned} u(x,t) &= \sin(x) + (-1)\cos(x)t + \frac{(-1)^2}{2}[-\sin(x)]t^2 + \frac{(-1)^3}{6}[-\cos(x)]t^3 + ... \\ &= \sin(x-t) \end{aligned}$$

Example 11 *Now, we consider initial data $u(s,-s) = \sin(s)$. Now, I need to find u_r, u_{rr} on the (r,s) plane (rotating $45°$), so the coefficient on the ODE must not vanish! Writing $r = x - y$ and $s = x + y$, we see that the coefficient vanishes for $u_x - u_y = 0$.*

Example 12 *Lets consider the second order equation $u_{tt} - u_{xx} = 0$ (wave equation). Then, after the change of variables,*

$$u(s,\varphi) = u_{\varphi\varphi}(\varphi_t^2 - \varphi_x^2) + ... = 0$$

And we need to determine $u_{\varphi\varphi}$ from the lower order terms. It is then clear that the non-characteristic condition reads:

$$\varphi_t^2 - \varphi_x^2 = 0$$

$$|\varphi_t| = |\varphi_x|$$

And the characteristic curves are lines with slopes ± 1.

From this last example, we can glean that the "simbol" of the highest order terms in our PDE gives out the non-characteristic condition, by replacing u with φ, and the order as an exponent. For example, for the equation $au_{xx} + bu_{xy} + cu_{yy} = 0$ it would be the quadratic $a\varphi_x^2 + b\varphi_x\varphi_y + c\varphi_y^2$.

Lecture 3 (27/01/09)

1.4 Weak Solutions: Shock waves and jump conditions

1.4.1 What is a weak solution?

Lets recall one of our first examples for quasilinear equations. This example is especially relevant, since it comes from a conservation law in fluid dynamics. It goes by the name of the inviscid Burgers equation, and is often applied to the modeling of gas dynamics and traffic flow. It is also a prototypical example of an equation for which the solution can develop discontinuities (shock waves):

$$\left\{ \begin{array}{c} u_t + uu_x = 0 \\ u(x,0) = f(x) \end{array} \right\}$$

For this equation, we had obtained an implicit solution of the form $u(x,t) = f(x - ut)$, and concluded that it was necessary to have $1 + tf'(x - \xi t) \neq 0$ for $t \geq 0$. Hence, if $f'(x - \xi t)$ is negative, then this will vanish at $t = 1/f'$.

The key point here is to realize that, since this models the behavior of a physical phenomena, there IS a solution. Hence, there must be a problem with the way we are seeking solutions.

We then have to go back to the integral form of the conservation law, and find out what we mean by a weak solution for this equation. The general form of a conservation law is the following: the change in time of a quantity (mass, charge, momentum, etc) for which we have a density function $u(x,t)$ is due to the flux of that quantity on the boundary (in $1D$, given by a flux function F evaluated on a and x). That is:

$$\frac{d}{dt}\int_a^x u(y,t)dy = F(u(x,t)) - F(u(a,t))$$

Equivalently, we can write the following equation using a test function $v \in D(\mathbb{R}^2)$:

$$\int_0^\infty \int_{-\infty}^\infty uv_t - F(u)v_x dxdt = 0$$

The equivalence of these two can be shown by the use of test functions of the form $v = \nu_1(x) \otimes \nu_2(t)$, where ν_1 is a smoothed $\chi_{[a,x]}$ and ν_2 a smoothed $\chi_{[\tau,t]}$, for a and τ arbitrary.

If u is regular, we can then push the differential sign in / use integration by parts respectively and obtain the PDE:

$$u_t - F'(u)u_x = 0$$

So, for our problem we have $F(u) = -u^2/2$.

1.4.2 The Rankin-Huguenot jump condition

Now, what happens if u is not differentiable? For instance, lets consider the case where $u(x,t)$ is piecewise smooth, with a jump discontinuity on the curve $x = s(t)$. We will show in a moment that, in that event, u is a weak solution if it satisfies the PDE whenever it is C^1, and if it jumps, it only does so on curves which satisfy the Rankin-Huguenot jump condition:

$$\frac{ds}{dt} = \frac{F(u_+) - F(u_-)}{u_+ - u_-}$$

Where $u_+(x_0)$ and $u_-(x_0)$ are the limits when $x \to x_0$ from each side of $x = s(t)$.

Claim 13 *If u is a weak solution as defined by the conditions above, then it satisfies the integral form of the corresponding conservation law.*

Proof. Clearly, in regions which do not include discontinuities, the PDE implies the integral equation and there is nothing to show. Let Ω be a ball centered around a point in $(s(t),t)$, Ω_+ the fraction on the right side, and Ω_- on the left. Then, for $\varphi \in D(\Omega)$

$$\begin{aligned} 0 &= \int_0^\infty \int_{-\infty}^\infty u\varphi_t - F(u)\varphi_x dxdt \\ &= \int_{\Omega_+} u\varphi_t - F(u)\varphi_x dxdt + \int_{\Omega_-} u\varphi_t - F(u)\varphi_x dxdt \end{aligned}$$

Now, integrating by parts on each region:

$$\begin{aligned} \int_{\Omega_+} u\varphi_t - F(u)\varphi_x dxdt &= -\int_{\Omega_+} [u_t - F(u)u_x]\varphi dxdt + \int_C u\nu_t - F(u)\nu_x \\ &= \int_C u_+\nu_t - F(u_+)\nu_x dl \end{aligned}$$

Since $[u_t - F(u)u_x]$ vanishes a.e. on Ω_+ (u is C^1 and the integral equation here is equivalent to the PDE). On Ω_- we have the same situation, with $\nu_- = -\nu_+$ (normal vectors are opposite). Hence,

$$\begin{aligned} &\int_{\Omega_+} u\varphi_t - F(u)\varphi_x dxdt + \int_{\Omega_-} u\varphi_t - F(u)\varphi_x dxdt \\ = &\int_C [u_+ - u_-]\nu_t - [F(u_+) - F(u_-)]\nu_x dl \end{aligned}$$

Since this holds for all test functions, it implies that:

$$[u_+ - u_-]\nu_t - [F(u_+) - F(u_-)]\nu_x = 0$$

Now, for a curve given by $x = s(t)$, the tangent vector is given by $\tau = (\dot{s}, 1)$, and so the normal vector is $(\nu_x, \nu_t) = (-1, \dot{s})$. Hence,

$$[u_+ - u_-]\dot{s} = [F(u_+) - F(u_-)]$$

Which returns the stipulated jump condition. ■

Remark 14 *We can treat a generalization of this result using an identical process as the one shown above. We now want to apply the same procedure to the equation:*

$$S'(u)u_t - F'(u)u_x = 0$$

And moreover, we want to study the discontinuity on a curve given by $(x(\tau), t(\tau))$. In an analogue fashion, a weak solution for this equation is defined as u such that, for all $\varphi \in D(\mathbb{R}^2)$,

$$0 = \int_0^\infty \int_{-\infty}^\infty S(u)\varphi_t - F(u)\varphi_x dxdt$$

Following the same procedure, we conclude that:

$$[S(u_+) - S(u_-)]\nu_t - [F(u_+) - F(u_-)]\nu_x = 0$$

Now $\tau = (\dot{x}, \dot{t})$, and $\nu = (-\dot{t}, \dot{x})$ and so:

$$[S(u_+) - S(u_-)]\dot{x} - [F(u_+) - F(u_-)]\dot{t} = 0$$

(or more generally, we need $[S(u_+) - S(u_-)]dx - [F(u_+) - F(u_-)]dt$ to be an exact differential). In the case the curve is given by $x = s(t)$, this reduces to:

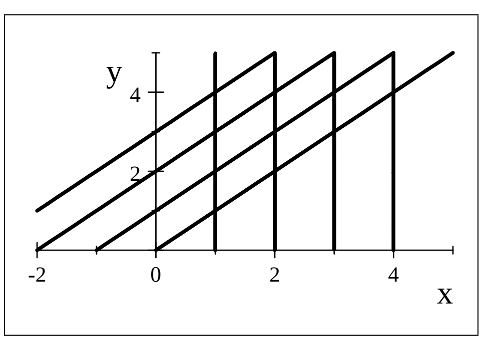
$$\frac{ds}{dt} = \frac{F(u_+) - F(u_-)}{S(u_+) - S(u_-)}$$

and our previous example corresponds to $S(u) = u$.

1.4.3 Are weak solutions unique?

Through several examples using the Burgers equation, we will see that even if we go back to the conservation law and come up with weak solutions, these might not be unique. We will study a way to determine a unique solution (known as the viscosity solution) which is more significant from a physical / modelling standpoint, and we will also see different phenomena such as shocks or rarefaction waves appear.

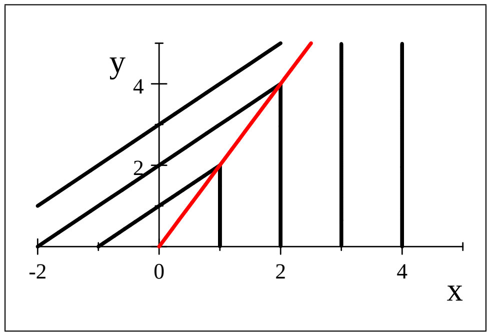
Example 15 *Let us solve the Burgers equation, with discontinuous initial data $u(0, x) = \chi_{[x<0]}(x)$. We first notice that the characteristics intersect, and thus, the method of characteristics does not yield an acceptable answer to the PDE. x*



Going back to the conservation law, we find that the jump condition reads:

$$\dot{s} = \frac{F(u_+) - F(u_-)}{u_+ - u_-} = \frac{\frac{1}{2}(0-1)}{0-1} = \frac{1}{2}$$

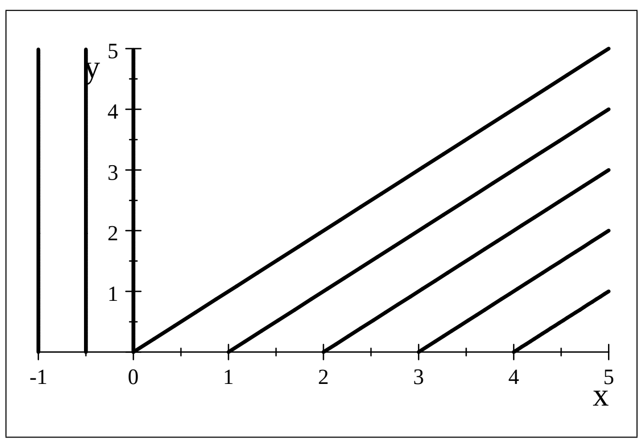
And so $s(t)$ must be of the form $x = \frac{t}{2} + C$. In this case, the only way to obtain a solution for all x and t is to introduce a jump / a "shock" at $x = 0$, and so $s(t) = \frac{t}{2}$.



This first example displays a shock: it can model cars that are moving at 2 different speeds (and have to break as they catch up) or the collision of sound waves.

Remark 16 *Another issue of defining weak solutions is that there is no unique way to go from the PDE to a more general integral equation. There are many ways to invent fluxes such that different integral equations will imply the same PDE given enough regularity of u.*

Example 17 *Now, let the initial data be $u(0, x) = \chi_{[x>0]}$. In this case, the characteristics do not intersect (except for being undefined for $x = 0$). Also, there is a whole area where we can define u to be a number of things, obtaining multiple weak solutions. x*



For example, we can define x to be identically 0 or 1 on each side of $\{t > 2x > 0\}$ (in which case we introduce a shock on the curve $s(t) = t/2$ given by the jump condition),We can instead opt for a solution which is differentiable except at $(0, 0)$, say

$$u(x, t) = \frac{x}{t}\chi_{\{0<x<t\}} + \chi_{\{x>t>0\}}$$

This solution looks like a fan, and corresponds to a "rarefaction" wave.

1.4.4 Which weak solution to pick?

There are several considerations which might sway our judgement towards picking one weak solution or another. In general, a more regular weak solution may be preferred, and we may also ask ourselves if I can view it as the limit of solutions with smooth data. We might also want criteria to determine whether a unique weak solution exists such that it fulfills extra conditions which are physically meaningful.

In the previous example, let u_ε be the solution to the Burgers equation with data which is identical to $\chi_{\{x>0\}}$ for $|x| > \varepsilon$, and is either a continuous (line segment) or even smooth (mollified) version for $|x| < \varepsilon$. We may compute the solution explicitly, and it is easy to see that as $\varepsilon \to 0$, u_ε can only tend to the "fan" / rarefaction wave solution. This is because the characteristics on $|x| < \varepsilon$ will continuously / smoothly come close to each other on the origin.

Entropy Condition

(From Evans) There is a natural condition, called the entropy or viscosity condition, which we can impose in order to produce physically meaningful weak solutions. It lies on the following observation: although we may expect to encounter the crossing of characteristics, resulting in discontinuities / shocks as we move forward in time, we also expect to be able to go backwards (moving along a characteristic) and recover the solution from initial conditions.

Hence, we require meaningful weak solutions to fulfill what is called an entropy condition, that is:

$$F'(u_-) > \dot{s}(t) > F'(u_+)$$

Since the slope of the characteristic in the (t, x) plane is precisely given by $F'(u)$, this ensures no characteristics can "spring" out of a curve where we have introduced a discontinuity. In particular, if F is uniformly convex ($F'' > \theta > 0$), this is equivalent to requiring $u_- > u_+$.

Lax-Oleinik Formula

Lets assume F is uniformly convex and $F(0) = 0$. We impose the initial data $u(x, 0) = g(x) \in L^\infty$, and we define:

$$h(x) = \int_0^x g(y)dy$$

Then, by the Hopf-Lax formula for the Hamilton-Jacobi equations, for $L = F^*$ (Fenchel transform)

$$w(x,t) = \min_y \left\{ tL\left(\frac{x-y}{t}\right) + h(y) \right\}$$

is the solution to the HJB equation:

$$\left\{ \begin{array}{c} w_t + F(w_x) = 0 \\ w(0,x) = h(x) \end{array} \right\}$$

Since h is differentiable a.e. (since it is absolutely continuous), we can differentiate with respect to x and then set $u = w_x$. But then, we find u is a solution for our original PDE.

Under these assumptions, it can be shown that the entropy solution is unique a.e., and that the Lax-Oleinik formula yields this solution.

1.5 Nonlinear Equations: Cauchy Kovalewski

In general, a first order nonlinear PDE is given by an equation of the form

$$F(x, u, \nabla u) = 0$$

So far, we have solved the linear and quasilinear cases. How do I solve this general case?

The answer is that, in a sense, every PDE is quasilinear with respect to higher order derivatives. Denoting $p_i = \frac{du}{dx_k}$

$$\frac{d}{dx_k} F(x, u, p) = \frac{dF}{dx_k} + F_u p_k + \sum \frac{dF}{dp_i} \partial_{ik} u = 0$$

Then, we have the characteristic curves are given by solving the system:

$$\left\{ \begin{array}{c} \frac{dx_i}{d\tau} = \frac{dF}{dp_i}, \frac{du}{d\tau} = \sum_k p_k \frac{dx_k}{d\tau} \\ \frac{dp_k}{d\tau} = -\frac{dF}{dx_k} - F_u p_k \\ F(x,u,p) = 0 \end{array} \right\}$$

(Insert example / from homework)

From our previous results, we know the solution by the method of characteristics can't always be computed or doesn't always make sense. The Cauchy Kovalewski theorem tells me how to solve the equation $w_t = F(t, x, w, \nabla w)$ in the case where F is analytic, and we have an initial condition $w(x, 0) = f(x)$ which is analytic as well. In this case, it concludes that there exists a local solution in a neighborhood Ω of $(0, x_0)$.

The method used on this theorem is a power series expansion. The question it asks is, given that I have this information, can I complete higher order derivatives? The key is that coefficients can be majorized, and so we can compute the power series around a point.

Generalizations: first, we can then apply this to the equation:

$$\partial_t^m w = G(t, x, w, \partial_t w, .., \partial_t^{m-1} w, D^\alpha w), \ |\alpha| \leq m, \alpha_1 < m$$

With $m-1$ pieces of information as initial data (Cauchy data), $w(x,0) = f_0, .., w^{m-1}(x,0) = f_{m-1}$. I can then change this into an instance of the first case, so that the original C-K theorem applies.

Finally, suppose I don't have a preferred direction. then the equation looks like:

$$\sum a_\alpha(x) \partial^\alpha u + G(x, u, \{\partial^\beta u\}_{|\beta| \leq m-1}) = 0$$

And I provide initial (Cauchy) data on the hypersurface given by $\varphi(x) = 0$, in the form of normal (or oblique) derivatives up to order $m-1$. The question then is, can I find a power series solution? To find out, we first apply the usual change of variables, to separate tangential from normal derivatives. We obtain:

$$\left[\sum_{|\alpha|=m} a_\alpha (\nabla\varphi)^\alpha \right] \nabla_\nu^m u = \widetilde{G}(..)$$

Then the non-characteristic condition is given by $\sum_{|\alpha|=m} a_\alpha (\nabla\varphi)^\alpha \neq 0$, and under that assumption, $C - K$ tells me a local solution exists.

(Here Hantaek gave us a very messy lecture on C-K Theorem and a sketck of its proof. I will probably have to go over a book to replace it)

Lecture 5 (03/02/09)

1.5.1 Shatah's summary: Characteristics & the C-K Theorem

- In the simplest case, we impose Cauchy data on $\{x_d = 0\}$, consisting of the function and normal (or oblique) derivatives up to order $m-1$. Since we are looking for a power series solution, we need to be able to compute the tangential derivatives, and hence like in the quasilinear case, we only need $a_{dd} \neq 0$.
- In general, we apply a change of variables, such that one of them is $\tau = \varphi(x)$. Then $\tau = 0$ is a parametrization of our surface, and plugging this change of variables into the equation:

$$u_{x_i} = \widetilde{u}_\tau \varphi_{x_i} + \sum_\alpha \widetilde{u}_{y_\alpha} y^\alpha_{x_i}$$

$$u_{x_i x_j} = u_{\tau\tau} \varphi_{x_i} \varphi_{x_j} + \sum_\alpha u_{\tau, y_\alpha} (\varphi_{x_i} y^\alpha_{x_j} + \varphi_{x_j} y^\alpha_{x_i}) + \sum_{\alpha,\beta} u_{y_\alpha, y_\beta} y^\alpha_{x_i} y^\beta_{x_j} + \text{first derivatives} = 0$$

At $\tau = 0$, I am given initial data $\widetilde{u}(y, 0)$ and $\widetilde{u}_\tau(y, 0)$. Hence, this simplifies to:

$$\left[\sum_{i,j} a_{ij} \varphi_{x_i} \varphi_{x_j} \right] u_{\tau\tau} + \text{known stuff} = 0$$

And so, again, replacing the highest order differential operators by derivatives of φ gives us the expression that must vanish to fulfill the non-characteristic condition.

- I can do the same thing for higher order / systems of PDEs. We would now have:

$$\sum A_i \partial_i u + Bu = 0$$

And expanding in power series, we observe that in a similar fashion, we need to invert the matrix

$$\sum A_i \varphi_{x_i}$$

So, the non-characteristic condition now reads:

$$\det \left(\sum A_i \varphi_{x_i} \right) \neq 0$$

Which is always a 1st order nonlinear equation.

1.5.2 Characteristics as the "Carriers of Discontinuity"

The main importance of characteristics is that one can show discontinuities can only be carried through them. To illustrate this, we will give an example using the wave equation:

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

Let u be a solution which is piecewise C^2 and has a discontinuity across a curve. Then this curve $\varphi(x,y) = 0$ must be a characteristic. Now, let $h(\tau) = \frac{1}{2}\tau^2 \chi_{\{\tau>0\}}$. This function is clearly C^1, and has a unit jump at 0 on its second derivative (it's a 2nd antiderivative of the heaviside function). Now, let:

$$u = Nh(\tau) + R$$

Where N and R are nice (C^2 or better). Lets analize the behavior of u on both sides of a curve $y = c(x)$. We can then write:

$$u = u^- h(y - c(x)) + R(x)$$

And plugging it in,

$$u_{yy} - u_{xx} = u^- h''(y - c(x))[1 - c'(x)^2] + (\text{continuous or better}) = 0$$

The only way this can happen is if the term containing h'' (which is a Heaviside function) vanishes. Hence, we must have:

$$c'(x)^2 \equiv 1$$

Hence, we can only have a jump in the second derivative if it occurs on characteristic lines! This result also imples that, if there are no characteristics, no jumps are propagated (hence the smoothing effects of elliptic operators).

2 Classification of 2nd order linear PDEs (Hadamard)

The most general second order linear PDE is of the form:

$$[Au_{xx} + 2Bu_{xy} + Cu_{yy}] + Du_x + Eu_y + F = h$$

Now, the characteristic equation yields a quadratic:

$$A\varphi_x^2 + 2B\varphi_x\varphi_y + C\varphi_y^2 = 0$$

There are three distinct possibilities:

- $AC - B^2 > 0$, **HYPERBOLIC:** (the quadratic has two real roots) Then, the PDE has two characteristic curves, and the canonical form is $[\partial_\alpha^2 - \partial_\beta^2]u = [\partial_\alpha - \partial_\beta][\partial_\alpha + \partial_\beta] = G$. The characteristic variables are then $\gamma = \alpha - \beta$, $\sigma = \alpha + \beta$, and the equation under this change of variable has a leading term $\widetilde{u}_{\gamma\sigma} = G$.

- $AC - B^2 < 0$, **ELLIPIC:** (no real roots) Then, the PDE has no characteristic curves, and the canonical form is $[\partial_\alpha^2 + \partial_\beta^2]u = \Delta u = G$.
- $AC - B^2 = 0$, **PARABOLIC:** (the quadratic has one real root) Then, the PDE has one characteristic curve, and the canonical form is $\partial_\alpha^2 u = G$.

Not all equations fit in these categories, although there are ways of extending these to higher order PDEs. In general, each type of PDE has different properties that we want to study, and requires qualitatively different approaches to find solutions.

Lecture 6 (05/02/09)

3 Hyperbolic Equations

3.1 Homogeneous equation

As we mentioned on the classification of $2nd$ order linear hyperbolic equations, the arquetypical homogeneous equation looks like:

$$u_{tt} - u_{xx} = \Box(u) = 0$$

And the characteristics are lines with slope ± 1. Hence, we can pose a Cauchy problem on a curve that avoids being characteristic, such as $t = 0$:

$$u(x,0) = f(x) \quad u_t(x,0) = g(x)$$

How do we solve this?

- We can write this as a $1st$ order system, diagonalize it and solve it on each characteristic line. We can write $v = u_t$ and $w = u_x$, and so we have:

  $$\left\{ \begin{array}{c} v_t - w_x = h \\ v_x - w_t = 0 \end{array} \right\}$$

  Or equivalently,

  $$\begin{pmatrix} v \\ w \end{pmatrix}_t = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} v \\ w \end{pmatrix}_x + \begin{pmatrix} h \\ 0 \end{pmatrix}$$

  Diagonalizing this system, we find that $v + w$ and $v - w$ are each constant on one of the characteristics.

- Using characteristic coordinates, we factor de equation and obtain the change of variables:

  $$\widetilde{u}_{\gamma\sigma} = 0$$

  Which in turn implies that:

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

  Now, in order to determine F and G, we must use our initial data. Plugging it in, we find:

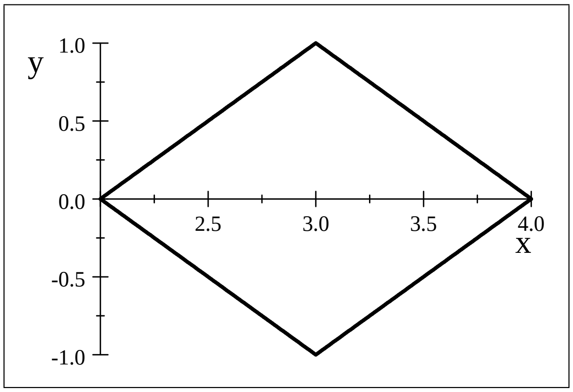
  $$\begin{aligned} u(x,0) &= F(x) + G(x) = f(x) \\ u_t(x,0) &= F'(x) + G'(x) = g(x) \end{aligned}$$

  The solution to this system, which is also known as the D'Alembert solution, is:

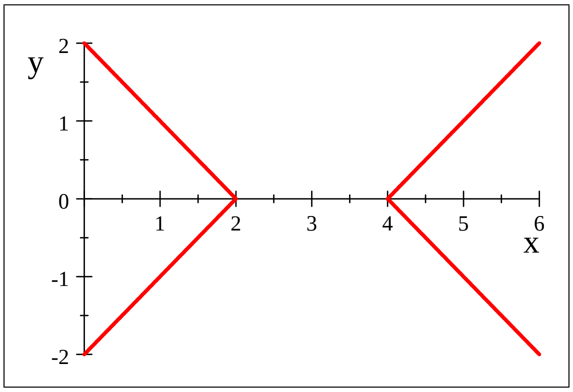
  $$u(x,t) = \frac{f(x+t) + f(x-t)}{2} + \int_{x-t}^{x+t} g(s)ds$$

Using this explicit solution, we can study two important concepts for the study of hyperbolic equations: domain of dependence and domain of influence.

Definition 18 *We say* (x,t) *is on the* ***domain of dependence*** *of* $\gamma \subset \{\varphi(x)=0\}$ *if* $u(x,t)$ *depends only on initial data inside* γ*. For example, for the wave equation, the domain of dependence of* $[a,b]$ *is the diamond* $\{(x,y) : \max(-x+a, x-b) > y > \min(x-a, -x+b)\}$.



Definition 19 *We say* (x,t) *is on the* ***domain of influence*** *of* $\gamma \subset \{\varphi(x)=0\}$ *if* $u(x,t)$ *is affected by changing initial data on* γ*. For example, for the wave equation, the domain of influence of* $[a,b]$ *is the region* $\{(x,y) \; : \min(-y+a, y-a) < x < \max(y-b, -y+b)\}$.



We also note that information "travels" or is propagated through characteristics at speed 1 (slope ± 1). Hence, a person / object moving at greater speeds would never see / hear it. This is the property of **finite propagation speed** of hyperbolic equations.

3.2 Nonhomogeneous equations

We are now interested in solving the nonhomogeneous wave equation:

$$u_{tt} - u_{xx} = h(x,t)$$

With the same initial data as before. As we know, it is a standad procedure to separate this into two parts, the homogeneous solution with inhomogeneous data (which we already calculated), plus a particular solution with 0 data:

$$\left\{ \begin{array}{c} u_{tt} - u_{xx} = h \\ u(x,0) = 0 \\ u_t(x,0) = 0 \end{array} \right\}$$

Using the same change of variables as before (characteristic variables), we obtain:

$$\widetilde{v}_{\xi\eta} = \widetilde{h}$$

Hence, integrating we obtain:

$$\begin{aligned}\widetilde{v} &= \int_0^{\xi}\int_0^{\eta} h(a,b)dadb + C\\ v &= \frac{1}{2}\int_0^t \int_{x-t+\tau}^{x+t+\tau} h(s,\tau)dsd\tau\end{aligned}$$

By transforming back to the original variables. Thus,

$$u = \frac{f(x+t)+f(x-t)}{2} + \frac{1}{2}\int_{x-t}^{x+t} g(s)ds + \frac{1}{2}\int_0^t \int_{x-t+\tau}^{x+t+\tau} h(s,\tau)dsd\tau$$

This again propagates at speed 1 through the characteristics.

3.3 Next level of complication: extra terms

Now, we want to solve:

$$\begin{aligned}\Box(u) + [a(x,t)u_x + b(x,t)u + c(u)] &= h(x,t)\\ \Box(u) &= F(x,t,u,u_t,u_x)\end{aligned}$$

Which is semilinear in (x,t). This has a solution, but it is not explicit. Lets assume for now we have this problem with inhomogeneous data. Then, the solution should satisfy:

$$u(x,t) = \frac{1}{2}\int_0^t \int_{x-t+\tau}^{x+t+\tau} F(s,\tau,u,u_t,u_x)dsd\tau$$

This is the essense of what is known as Duhammel's Formula / Principle: we treat the extra terms as a right hand side, and we write the solution in implicit form.

I am looking for a nice (smooth) solution. Initially, I can show the solution is small in an appropriate norm. Close to a point x_0, we take an interval of radius δ, and consider its domain of dependence. We then truncate this at $t = T$, and call this K_T. Now, let $\|u\|_{C^1(K_T)} \leq a$,

We also assume that, on K_T,

$$\|F(x,t,u)\|_{\infty,K_T} \leq C(\delta)\,\|u\|_{C^1(K_T)}$$

Now, bounding the derivative using the integral formula, we get: $\left|\frac{d}{dt}\int_0^t \int_{x-t+\tau}^{x+t+\tau} F(s,\tau,u,u_t,u_x)dsd\tau\right| \leq C_2(\delta,a)T$

And so, combining these two results, we have:

$$\left\|\frac{1}{2}\int_0^t \int_{x-t+\tau}^{x+t+\tau} F(s,\tau,u,u_t,u_x)dsd\tau\right\|_{C^1(K_T)} \leq C(\delta,a)T$$

We can then pick T small enough, so that this final bound is smaller than a. This makes the map that goes from u to this alleged solution a contraction in $C^1(K_T)$, and by the Banach Fixed Point Theorem, a unique fixed point exists (which is the local solution we were looking for).

3.3.1 The Newton Method

Based on this last result, we can set an iteration procedure, which is known as the Newton method. Essentially, starting with $u_0 = 0$, we iterate our map:

$$u_k = \frac{1}{2}\int_0^t \int_{x-t+\tau}^{x+t+\tau} F(s,\tau,u_{k-1})dsd\tau$$

Now, we have a uniform bound $\|u_k\|_{C^1(K_T)} \leq a$, and furthermore,

$$\begin{aligned} \|u_{k+1} - u_k\|_{C^1(K_T)} &\leq C(a)T \|F(u_k) - F(u_{k-1})\|_{C^1(K_T)} \\ &\leq C(a)M(a)T \|u_k - u_{k-1}\|_{C^1(K_T)} \end{aligned}$$

As we noted before, for T sufficiently small we have a contraction, and hence the sequence u_k is Cauchy in $C^1(K_T)$. Thus, there exists a unique u such that $u_k \to u$. We note here that u need not be C^2, it depends on F and at best we can assert it is a weak solution to the original differential equation.

3.4 Hyperbolic systems

Now, think about the hyperbolic system:

$$\left\{ \begin{array}{c} u_t + (1+u^2)u_x = v^2 \\ v_t - v_x = u^2 \end{array} \right\}$$

Can we still perform some kind of iteration procedure / fixed point argument? This is a quasilinear problem, and we can write it as $u_t + A(u)u_x = f(x)$, where u is a vector and $A(u)$ is a matrix. We can then find the condition for a surface to be characteristic: if we have the surface $\varphi(x,t) = 0$, it is characteristic if and only if

$$\det(A\varphi_x + \varphi_t I) = 0$$

That is, if we cannot invert the leading term in our expansion / cannot compute derivatives of u from data. Then, if this surface is given by the equation $x = s(t)$, this implies:

$$\det(A + \frac{ds}{dt} I) = 0$$

That is,

$$\left\{ \frac{ds}{dt} = \lambda(u) : \lambda(u) \in \Lambda(A(u)) \right\}$$

Are the characteristic curves for this equation. If all eigenvalues are real and distinct, and I have a full set of eigenvectors, we call the equation **strictly hyperbolic**. If in addition, A is symmetric, we call this a **symmetric hyperbolic** system. We note that the method of characteristics requires us to diagonalize this matrix, in order to work in characteristic coordinates.

3.4.1 Newton Method (revisited)

Now, we will see that the same iteration procedure applies for $1st$ order hyperbolic systems.

- Say we have:

$$\begin{aligned} u_t + A(x,t)u_x &= f(x,t,u) \\ u(x,0) &= u_0(x) \end{aligned}$$

  $t = 0$ is not a characteristic and the system is strictly hyperbolic. "So you say, this is too complicated, first I'll diagonalize it". After a change of variables,

$$\left\{ \frac{\partial u_a}{\partial t} + \lambda_a(x,t) u_a = f_a(x,t,u) \right\}_{a=1}^{N}$$

- We then solve the system of ODEs $\frac{dx_a}{dt} = \lambda_a(x,t)$, $x_a(0) = x_0$, $D_t u_a = f_a$.

If we start on x_0, this is ging to give me solutions "all over the place", since each coordinate travels on different curves. We cannot solve this as a system of ODEs. What do I do? I start with (x,t) a point close enough to the x axis, and ask what is $u(x,t)$. If all curves intersect it, then I am fine.

- So, choose τ small enough so that the Implicit Function Theorem guarantees that all characteristic curves γ_a intersect the x axis. That is: γ_a is a map from x_0 to x, and I need to invert it. I know $\frac{d\gamma_k}{dx_0} = 1$ at time $t = 0$, so for $t \leq \tau$ small enough, the derivative remains positive and this inversion is possible.

- We then restrict the region in which we are solving. If $M = \max_{t \leq \tau}\{\lambda_k\}$ and $m = \min_{t \leq \tau}\{\lambda_k\}$, for a small enough interval $[x_1, x_2]$ we can consider the trapezium Ω consisting of $t \in (0, \tau)$ and bounded by the line of maximum slope / speed M on x_1 and that of minimum slope / speed m on x_2 (that way, we ensure the solution inside depends only on the data in $[x_1, x_2]$, that is, we are inside this interval's domain of dependence).

- Now, my iteration procedure reads:

$$\left\{ \begin{array}{c} \frac{d\gamma_a}{dt} = \lambda_a(\gamma_a, t) \\ D_t u_a^m = f_a(x, t, u^{m-1}) \end{array} \right\}$$

  This defines a sequence u^m in Ω, and using the same argument as before we can make this into a contraction mapping for τ small enough. Hence, we have proved the existence of a solution with data on $t = 0$, and $u_0 \in C^1$ implies the local solution $u \in C^1(\Omega)$.

- Finally, to check local well-posedness, we have to check that the solution map $u_0 \to S(u_0) = 0$ is continuous from C^1 to itself. This of course depends strongly on the topologies of the spaces we take our data and our solution to lie on.

- We finally note that, if we now make A depend on u, we get a quasilinear equation. In that case, λ_a depends on u, but nothing essential in the previous argument changes.

Lecture 7 (10/02/09)

Part II
Background Material

Here are some results which are needed to build analytical tools and work with PDEs from a more general / rigorous standpoint:

3.5 Fourier Transforms

The Fourier Transform is one of the most useful operators in mathematical analysis. In order to extend it from the function spaces where it is usually defined (say, L^1) to L^2 and spaces of distributions, it is relevant first to introduce the space of rapidly decaying smooth functions:

Definition 20 *(Schwartz Space) We define* $S(\mathbb{R}^d) = \{f \in C^\infty(\mathbb{R}^d) : \left\|x^\beta D^\alpha f\right\|_\infty < C_{\alpha,\beta} \ \ \forall \alpha, \beta\}$

Proposition 21 *Let $f \in S$, then the Fourier Transform defined by*

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

is well defined, and moreover, $\widehat{f} \in S$. The converse is also true.

Proposition 22 *Two key properties of the fourier transform that are linked to the previous result are those which relate derivatives and multiplication by powers. That is,*

$$\begin{aligned} D_\xi^\alpha \widehat{f}(\xi) &= [\widehat{(-ix)^\alpha f}](\xi) \\ \widehat{D_x^\alpha} f(\xi) &= (i\xi)^\alpha \widehat{f}(\xi) \end{aligned}$$

Proposition 23 $\forall f \in S(\mathbb{R}^d)$,

$$\|f\|_{L^2} = \left\|\widehat{f}\right\|_{L^2}$$

Proof. The core of the proof is to use a limit with a Gaussian distribution function, using it as an approximate identity. That is, we have:

$$\left\|\widehat{f}\right\|_{L^2} = \lim_{\varepsilon \to 0} \int \int \int e^{-\varepsilon|\xi|^2} f(x) e^{-ix^\top \xi} \overline{f(y)} e^{iy^\top \xi} dx dy d\xi$$

Applying Fubini, we integrate $\exp(-\varepsilon |x|^2 + i(y-x)^\top \xi)$ with respect to ξ, and obtain $G_\varepsilon(x-y) = \frac{1}{(2\pi\varepsilon)^{d/2}} \exp(-\frac{|x-y|^2}{4\varepsilon})$, which is a Gaussian distribution function with variance ε. We can also see it as an approximate identity. Now,

$$\lim_{\varepsilon \to 0} \int \int G_\varepsilon(x-y) f(x) \overline{f(y)} dy dx = \int f(y) \overline{f(y)} dy = \|f\|_{L^2}$$

Since the convolution of f with an approximate identity in L^2 converges to f. ■

This result is usually known as the Plancherel theorem, and since S is dense in L^2, it allows us to define the fourier transform $F(f) = \widehat{f}$ for any $f \in L^2$.

3.6 Sobolev Spaces $H^k(\Omega)$

Now that we can use the fourier transform in L^2, we would like to define subspaces of functions which have derivatives (in a generalized sense) in L^2. In the case of the space of Schwartz, we saw bounding derivatives in space was equivalent to bounding powers in the fourier domain. Hence, it is natural to define:

Definition 24 *(Sobolev Space $H^s(\Omega)$) Given $s > 0$ (not necessarily an integer), we define the space $H^s(\Omega) = \{u \in L^2(\Omega) : (1 + \|\xi\|^2)^{s/2} \widehat{u}(\xi) \in L^2(\Omega)\}$. This is a Hilbert space, with the norm*

$$\|u\|_{H^s} = \left\|(1 + \|\xi\|^2)^{s/2} \widehat{u}\right\|_{L^2}$$

For $s = k$ integer, we can also define this space as $\{u \in L^2 : D^\alpha u \in L^2 \quad \forall |\alpha| \leq k\}$, and equip it with the equivalent norm:

$$\|u\|_{k,\Omega}^2 = \sum_{|\alpha| \leq k} \|D^\alpha u\|_{L^2}^2$$

It is clear that $H^0 = L^2$, and that H^s is a closed subspace of L^2 for $s > 0$. What would $H^s(\Omega)$ be for $s < 0$? (It is the dual space of $H_0^{-s}(\Omega)$).

Example 25 *$H^{-2}(\mathbb{R}^2)$ is the space of objects which on the Fourier domain are proportional / with growth like 1, since:*

$$\int_{\mathbb{R}^2} (1 + \|\xi\|^2)^{-2} dx < \infty$$

Similarly, objects in $H^{-3}(\mathbb{R}^2)$ would be asymptotically proportional to ξ on the fourier domain.

A fair question is if there are functions with these behavior, and the answer is no. To define these objects properly, we must interpret them as functionals of a dual space, which is the space of distributions (or generalized functions). We need these functions to study such things as differential operators of functions in L^2.

Lecture 8 (12/02/09)

3.7 Distributions

In the same way we have defined the H^s Sobolev spaces with negative s (which can be interpreted as dual spaces), we can define spaces of generalized functions by taking the dual of spaces such as S.

Definition 26 *The space S^* with the corresponding weak* topology is known as the space of tempered distributions, and it can be identified with slowly growing functions. By duality and using the weak Parseval lemma, we can define the Fourier Transform on S^* by:*

$$\langle F(T), \phi\rangle = \left\langle T, \widehat{\phi}\right\rangle \quad \forall \phi \in S$$

Definition 27 *The space $C_0^\infty(\Omega) = D(\Omega)$, known as the space of test functions, is the space of C^∞ funtions with compact support in Ω. We can immediately see that $D \subset S$.*

Definition 28 *The space D^* is what we now as the space of distributions, or generalized functions.*

We then have the following embeddings:

$$C_0^\infty \subset S \subset .. \subset H^1 \subset L^2 \subset H^{-1} \subset .. \subset S^* \subset D^*$$

In order to work with these spaces, we must first determine what is the topology and convergence in the space of test functions. This is a union of Fréchet spaces, and we say $\{\phi_n\} \subset D(\Omega)$ is convergent to ϕ if $\exists K \subset\subset \Omega$ (compact) such that $\text{supp}(u_n) \subset K$ $\forall n \geq n_0$ and $D^\alpha \phi_n \to D^\alpha \phi$ uniformly on K.

Proposition 29 *Then for every element $T \in D^*$, we have that $\forall K$ compact set of Ω, $\exists$ $N(K)$ such that*

$$|T(u)| \leq C_K \|u\|_{C^N(K)}$$

When this N is independent of K, we say T is a distribution of order N.

Proof. We show this by contradiction. Lets assume $\exists K_0$ a compact set such that $\forall N \in \mathbb{N}$ $\exists$ $u_N \in D(K_0)$ such that $\|u_N\|_{C^N(K_0)} = 1$ and $|T(u_N)| \geq N$. But then, $v_N = u_N / |T(u_N)|$ is such that:

$$\begin{aligned} \|v_N\|_{C^N} &\leq \frac{1}{N} \text{ (hence } v_N \to 0 \text{ in } D(\Omega)) \\ T(v_N) &= 1 \ \forall N \end{aligned}$$

Which contradicts the fact that $T \in D^*$. ■

Why do we like the Fourier Transform definition for Sobolev spaces? Lets show an embedding result: Let $s > 0$, $f \in H^s(\mathbb{R}^d)$. Then:

$$\begin{aligned} \|f\|_\infty &\leq \left\|\widehat{f}\right\|_1 \leq \left\|\frac{1}{(1+\|\xi\|^2)^s}(1+\|\xi\|^2)^s \widehat{f}\right\| \\ &\leq \left\|\frac{1}{(1+\|\xi\|^2)^s}\right\|_2 \left\|(1+\|\xi\|^2)^s \widehat{f}\right\|_2 \\ &= \left\|(1+\|\xi\|^2)^{-s}\right\|_2 \|f\|_{H^s(\mathbb{R}^d)} \end{aligned}$$

Now, $\left\|(1+\|\xi\|^2)^{-s}\right\|_2$ is finite if and only if $s > d/2$. In that case, we have a continuous embedding of $H^s(\mathbb{R}^d) \hookrightarrow C(\mathbb{R}^d)$. Using a similar procedure, we can also conclude that, when $s > d/2 + N$, we have an embedding $H^s(\mathbb{R}^d) \hookrightarrow C^N(\mathbb{R}^d)$. These results are very important, since they link Sobolev spaces (which we can work with on a very general setting) to classic spaces, and thus are useful to determine the regularity of weak solutions of PDEs.

Remark 30 *Now, let $T \in D^*$ of order N. Then, by the previous result, we have that for $u \in H^s(\mathbb{R}^d)$, with $s > d/2 + N$ we have:*

$$|T(u)| \leq C \|u\|_{C^N(K)} \leq \widetilde{C} \|u\|_{H^s(\mathbb{R}^d)}$$

Hence, $T \in H^{-s}$ for all such s. In particular, N counts the number of derivatives I need to use.

Proposition 31 *Any measures that assign finite values to compact sets (Radon measures) are distributions. For example, the dirac mass $\delta_0(\phi) = \phi(0)$.*

3.7.1 Weak derivative and other operations on D^*

We can now extend the notion of derivative to the space of distributions (as can be seen, most operations can be succesfully extended by duality arguments). This notion of weak derivative will give us a precise way to work with weak solutions of PDEs.

Definition 32 *Let $T \in D^*$. Then, the weak derivative $D^\alpha T$ is defined by its action on $\phi \in D(\Omega)$:*

$$\langle D^\alpha T, \phi \rangle = (-1)^{|\alpha|} \langle T, D^\alpha \phi \rangle$$

This is a direct generalization of what we would get using integration by parts. In this case, both notions of derivative coincide.

Other operations are allowed on D^*, as long as we are careful:

- (**multiplication by a function**) Let $\psi \in C^\infty$. Then I can define ψT by duality: $\langle \psi T, \phi \rangle = \langle T, \psi \phi \rangle$. Leibniz rules applies when considering this in combination with the weak derivative.
- (**convolution**) it is possible to extend the operation of convolution quite a bit, but it requires several steps. First, we remember some of the things we know about convolution:
  - If $f \in C_0^K$ and $g \in C_0^L$, then $f * g \in C_0^{\max\{K,L\}}$ and supp$(f * g)$ $\subset$supp(f)+supp(g). (In general, it always "inherits" the best behavior)
  - $\widehat{(f * g)} = \widehat{f}\widehat{g}$

Now, let $T \in D^*$. Can I convolve it with $\varphi \in D(\Omega)$? If it made sense to write the integral, we would have:

$$\int T(y)\varphi(x-y)dy = \langle T, \tau_x(\varphi) \rangle$$

Now, suppose $\varphi \in C_0^k$; I can talk about $T(\varphi * \phi)$. However, if I remove the compact support, this is no longer true.

Support of a Distribution In order to further extend the notion of convolution, we need to know what it means for a distribution to vanish, or equivalently, what the support of a distribution means. Again drawing intuition from the cases in which the integral makes sense, we know that $\int fg = 0$ whenever the supports of f and g are disjoint. Hence, we say T is supported on U closed set if $\langle T, \phi \rangle = 0 \ \forall \phi \in D(U^c)$.

Now, if T has compact support, we can talk about $T * S$, with S another distribution. We can do this because $T * \varphi \in D(\Omega)$, and so $T * S(\varphi) = S(T * \varphi)$ is well defined.

Proposition 33 *If T is a distribution supported at $x = 0$, then*

$$T = \sum_{|\alpha| \leq N} C_\alpha D^\alpha \delta_0$$

An application of this is the following: Let $u \in S^*$, and $\xi^2 \widehat{u} \equiv 0$. Hence, $\widehat{u}$ is only supported at 0, so we have:

$$u = \sum_{|\alpha| \leq N} C_\alpha D^\alpha \delta_0$$

We would like to conclude that $C_2, .., C_k = 0$. We have:

$$\int \xi^2 \widehat{u} \widehat{\phi} = 0$$

Now, we can construct ϕ such that $\widehat{\phi}(0) \neq 0$ and $\phi'(0) = \phi''(0) = ..0$. Then, by plugging it in, we can show $C_2 = C_3 = .. = 0$ (ξ^2 only survives when applying 2 derivatives or less). Hence, $\widehat{u} = C_0\delta + C_1\delta'$. (Check later)