Math 212br - Advanced Real Analysis

Taught by Yum-Tong Siu
Notes by Dongryul Kim

Spring 2018

This course was taught by Yum-Tong Siu, on Tuesdays and Thurdays from 2:30pm to 4pm. It drew material from Stein and Shakarchi's *Real analysis*, *Fourier analysis* and Evans's *Partial differential equations*, as well as famous papers. There were midterm and final take-home exams.

Contents

1 January 23, 2018

In this course, the emphasis will be on PDEs. Last term, we dealt with PDEs with constant coefficients, using the technique of Malgrange–Ehrenpreis. This fundamental solution is a distribution, but can be taken as a tempered distribution as well.

Definition 1.1. A **distribution** is an element of $\mathcal{D}'(\mathbb{R}^n)$, where $\mathcal{D}(\mathbb{R}^n)$ is the space of compactly supported complex-valued smooth functions, with the topology given by the direct limit of Fréchet spaces. A **tempered distribution** is an element of $\mathcal{S}'(\mathbb{R}^n)$, where $\mathcal{S}(\mathbb{R}^n)$ is the space of all rapidly decreasing smooth complex-valued functions on $\mathbb{R}^n$. This means that

$$\sup_{x\in\mathbb{R}^n} (1+|x|^2)^\beta |D^\alpha f(x)| < \infty$$

for all α and β.

The nice thing is that the Fourier transform maps $\mathcal{S}(\mathbb{R}^n)$ to $\mathcal{S}(\mathbb{R}^n)$. To get the fundamental solution as a tempered distribution, we need either the technique of Hörmander or of Bernstein–Sato polynomials.

We then look at a (single) PDE with variable coefficients, by freezing the coefficients and perturbation. If you $Lu = f$ is the equation we want to solve, and freeze the variable to get a L_0, this is an approximation of L_0. Then we can write

$$L = L_0 + (L - L_0) = L_0(1 + L - 0^{-1}(L - L_0))$$

and then we need to estimate the errors. But Hans Lewy came with an example where such a method does not work.

Example 1.2. On $\mathbb{R}^3$, with the variables x_1, y_1, x_2, let $z_1 = x_1 + iy_1$. Then the equation

$$\overline{L} = \frac{\partial}{\partial z_1} + i\bar{z}_1 \frac{\partial}{\partial x_2}$$

does not admit even a local smooth solution, where f is given by

$$F(z_1, z_2) = e^{-(\frac{z_2}{i})^{1/2}} e^{-(\frac{i}{z_2})^{1/2}}.$$

What is this equation? We have to look at the **Cayley transform**. In one complex variable, the map $w = \frac{z-i}{z+i}$ maps the upper half plane to the unit ball. In n variables,

$$w_n = \frac{i - z_n}{i + z_n}, \quad w_k = \frac{2iz_k}{i + z_n}$$

with $1 \le k \le n$ sends $\Im w_n < |w_1|^2 + \cdots + |w_{n-1}|^2$ to the ball. Then the boundary of $\Im w_n < \cdots$ look like $\mathbb{R}^3$, so we get an local chart of a boundary of B^3 by $\mathbb{R}^3$. This is how F should be taken as a function f on $\mathbb{R}^3$.

So we can't deal with PDEs with general variable coefficients. But we can say something once we impose something like ellipticity or hypoellipticity. For a long time, people were interested in isometric embeddings. Locally given an Riemannian metric, is it possible to embed it in large Euclidean space? People were trying to do this by looking at power series, but this is the same as freezing the coefficients, so people got stuck. Then Nash found out a way to deal with it, and Moser wrote it nicely. It is called the Nash–Moser implicit function theorem. Nash's method was to consider the difference quotient without differentiating, but then take care of the error terms.

1.1 PDE with constant coefficients

Consider an equation like $\Delta u = f$. If we take the Fourier transform, we get

$$-4\pi(\textstyle\sum_j \xi_j^2)\hat{u} = \hat{f}.$$

So we would like to say $\hat{u} = \hat{f}/(-4\pi \sum_j \xi_j^2)$, but there is a 0 on the denominator, so we are in trouble.

There are two tricks. The first is to go to the adjoint. Here, we want to show that $L : u \mapsto f$ is surjective, and so we can instead show that L^* is injective. So it suffices to show that $\|L^*\psi\| \geq c\|\psi\|$. When we take the Fourier transform, we need to show that $\|Q\hat{\psi}\|_{L^2} \geq c\|\hat{\psi}\|_{L^2}$. Now we can show some inequality to solve the equation. The second trick is to change the domain of integration and use the mean value property of holomorphic functions:

$$|F(0)|^2 \leq \frac{1}{2\pi}\int_{\theta=0}^{2\pi} |P(e^{i\theta})F(e^{i\theta})|^2 d\theta.$$

This shows the estimate. But nowadays, people change the domain of integration instead, by shifting it in the imaginary direction.

Then the problem is that this only gives u only as a distribution. If we want to get u as a tempered distribution, we need to divide $Q\hat{u} = \hat{f}$ by a polynomial. Hörmander's technique was to do this by some kind of a generalized L'Hôpital's rule. This is going to be bounding ψ by $Q\psi$. But there is another technique of Bernstein–Sato polynomials.

Consider the function $F = |P(x)|^2$, and for a complex number s we define F^s for $\Re s > 0$. The assignment

$$s \mapsto F^s = T(s) \in \mathcal{S}'(\mathbb{R}^n)$$

is holomorphic in the weak sense, that is, given any test function, it gives a holomorphic function. The question is to extend this meromorphically to all $\mathbb{C}$. This is in some sense similar to the Gamma function. Given a polynomial $g(x)$, there exists a partial differential operator L in $x_1, \ldots, x_n$ whose coefficients are in $\mathbb{C}[s, x_1, \ldots, x_n]$ such that $Lg^{s+1} = p(s)g^s$. (In the Gamma function, we had $\frac{\partial}{\partial x}x^{s+1} = (s+1)x^s$.) Then

$$g^s = \frac{1}{p(s)}(Lg^{s+1})$$

and so we can shift it by 1.

2 January 25, 2018

Here is Hans Lewy's example of a nonsolvable single PDE with variable coefficients in three real variables x_1, y_1, x_2: the equation $\overline{L}u = f$ where

$$\overline{L} = \frac{\partial}{\partial z_1} + i\bar{z}_1 \frac{\partial}{\partial x_2}, \quad f = e^{-(\frac{z_2}{i})^{1/2}} e^{-(\frac{i}{z_2})^{1/2}}.$$

This equation does not have a solution u even as a distribution near 0.

2.1 Using Bernstein–Sato polynomials

Definition 2.1 (Bernstein–Sato)**.** Given a polynomial $f(x_1, \ldots, x_n)$ with coefficients in $\mathbb{C}$, there exists a linear partial differential operator L with coefficients in $\mathbb{C}[x_1, \ldots, x_n, s]$ such that

$$L(f^{s+1}) = p(s) f^s$$

where $p(s)$ is monic. Such polynomials $p(s)$ forms an ideal, so we can look at the minimal monic $p(s)$. This polynomial $p(s)$ is called the **Bernstein–Sato polynomial** for f.

Example 2.2. For example, take $f(x_1, \ldots, x_n) = x_1^2 + \cdots + x_n^2$. Then

$$\left(\sum_{j=1}^{n} \frac{\partial^2}{\partial x_j^2} \right) f^{s+1} = 4(s+1)s(x_1^2 + \cdots + x_n^2)^s + 2n(s+1)(x_1^2 + \cdots + x_n^2)^s$$

and so

$$\frac{\Delta}{4} f^{s+1} = (s+1)(s + \tfrac{n}{2}) f^s$$

with $p(s) = (s+1)(s + \frac{n}{2})$.

The idea for proving this is to just take derivatives and try to cancel stuff out. Consider the non-commutative ring $D_n = \mathbb{C}[x_1, \ldots, x_n, \frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n}]$ with $\frac{\partial}{\partial x_j} x_l = \delta_{jl}$. This is called the **Weyl algebra**. We would like to look at modules over $D_n[s]$, which is also called a D-module. We can look at the sequence of modules

$$D_n[s] \cdot f^s \supseteq D_n[s] \cdot f^{s+1} \supseteq \cdots$$

and the question is whether this stabilizes. This is something like the Noetherian property.

The industry of D-modules is big, and is still ongoing. So what are people doing with them? People wanted to look at many equations, so that's one thing. The roots of Bernstein–Sato polynomials are also interesting.

Theorem 2.3 (Malgrange?)**.** *The roots of a Bernstein–Sato polynomial are negative rational numbers.*

So assume that we are happy with existence of such a polynomial. Our goal is to define $1/Q$ as a tempered distribution. Let's assume, without loss of generality that $f = |Q|^2$ so that $f \geq 0$ on $\mathbb{R}^n$. The idea is to look at

$$f^s = e^{s \log f}.$$

This is well-defined on $\Re(s) > 0$, and you can also check that it is well-defined on $\Re(s) > -\epsilon_f$ for some $\epsilon_f > 0$ sufficiently small. Now the Bernstein–Sato theorem gives us a polynomial $p(s)$ and a differential operator L such that

$$f^s = \frac{1}{p(s)} L(f^{s+1}).$$

So the assignment $s \mapsto f^s \in \mathcal{S}'(\mathbb{R}^n)$ can be extended meromorphically to all of $s \in \mathbb{C}$.

Take a Laurent series expansion at $s = -1$:

$$f^s = \sum_{k=-l}^{\infty} T_k (s+1)^k$$

for $l \in \mathbb{N} \cup \{0\}$. If we have this expansion, we have

$$f^{s+1} = f f^s = \sum_{k=-\ell}^{\infty} (f T_k)(s+1)^k,$$

where f^{s+1} is holomorphic at $s = -1$. This shows that $T_k = 0$ for $k < -1$ and $f T_0 = 1$. So T_0 is what we want.

Now how do we get this Laurent expansion? We write

$$\frac{1}{p(s)} = \sum_{\lambda=-k} a_\lambda (s+1)^\lambda, \quad L = \sum_{\nu=0}^{\ell} L_\nu (s+1)^\nu.$$

Then

$$f^s = \frac{1}{p(s)}(f^{s+1}) = \sum_{\nu=-k}^{\infty} T_\gamma (s+1)^\gamma.$$

So to get T_γ, we just expand the left hand side. It turns out that

$$T_\gamma = \sum_{\lambda+\nu+\mu=\gamma} a_\lambda L_\nu \Big(\frac{1}{\mu!} (\log f)^\mu \Big).$$

In particular the fundamental solution is obtained explicitly as

$$T_0 = \sum_{\nu=0}^{k} \sum_{\mu=0}^{k-\nu} a_\lambda L_\nu \Big(\frac{1}{\mu!} (\log f)^\mu \Big).$$

2.2 The example of Hans Lewy

After this example, the whole field of PDE changed, because before this example everyone was trying to freeze the coefficients. This example is not just random example that came from trial and error. The Fourier series works on S^1, and the Fourier transform works on $\mathbb{R}$. We can explain their relation by rescaling, but we can also explain the relation from the Cayley transform. If you think about the transform

$$w = \frac{z+i}{z-i},$$

this map sends the real line to the circle. But then, what is the relation between the Fourier transforms? Here, we would have to do something interesting. A function f on the boundary of the disk looks like

$$f(\theta) = \sum_{n=-\infty}^{\infty} c_n e^{in\theta} = \sum_{n=-\infty}^{-1} c_n e^{in\theta} + \sum_{n=0}^{\infty} c_n e^{in\theta}.$$

So we would have look at half of the transform. This will correspond to something like looking at the Fourier transform only on the nonnegative real line, or something like this.

Let us look at the parametrized Cayley transform on $\mathbb{C}^n$. If we write $(z_1, \dots, z_n) = (z', z_n)$, then the region we want to transform is $y_n = \Im(z_n) > |z'|^2$. Then this region

$$U = \{y_n > |z'|^2\}$$

will be biholomorphic to $\{|z|^2 < 1\}$. The formula for this biholomorphic transformation is

$$w_n = \frac{i - z_n}{i + z_n}, \quad w_k = \frac{2iz_k}{i + z_n}.$$

The boundary of U has real dimension $2n - 1$. In Hans Lewy's example, we have $n = 2$ and is identified with (z_1, x_2). The differential operator is a vector field of type $(0, 1)$ on U.

3 January 30, 2018

So far we have looked at different ways of solving linear PDEs with constant coefficients:

(1) Malgrange–Ehrenpreis

(2) vertical shifting of the line of integration for the Fourier transform

(3) Hörmander's divison of tempered distributions (using l'Hôpital's rule)

(4) using Bernstein–Sato polynomials to meromorphically extend $|f|^s$

The question of meromorphically extending $|f|^s$ was first asked by Gelfand in 1954 at the Amsterdam ICM. Berstein and Sato independently came up with ways of extending to $L(s)f^{s+1}b(s)f^s$, for a polynomial $b(s)$. There is also a fifth technique:

(5) Hörmander's completion of squares and commutation

3.1 Cohomology and local insolvability

We were talking about Hans Lewy's example

$$\overline{L} = \frac{\partial}{\partial z_1} + i\bar{z}_1 \frac{\partial}{\partial x_2}$$

on $\mathbb{R}^3$. There exists a C^∞-function f that rapidly decreasing at ∞ or 0, such thta $Lu = f$ cannot be solved locally at 0 for u being a distribution. There are also problems at other points, and then people started to answer questions about when a nice single equation is solvable.

One of the motivation is from global nonsolvability with real variabels to local, with complex variables. Global insolvability is easy to construct. Take, for instance, the polar coordinates. If we consider

$$d\theta = d\tan^{-1}\frac{y}{x} = \frac{xdy - ydx}{x^2 + y^2}$$

on $\mathbb{R}^2 - \{0\}$, there is no solution on $U - \{0\}$ for a neighborhood U of 0. The important thing is that locally, everything is convex so that topology cannot play any role.

But in the complex case, things can be locally non-convex. In Euclidean space, convexity means the following thing. For any function r, the domain $\{r < 0\}$ is convex when the Hessian

$$\left(\frac{\partial^2 r}{\partial x_j \partial x_k}\right)\Big|_{dr=0} > 0.$$

In the complex case, convexity is defined by

$$\left(\frac{\partial^2 r}{\partial z_j \partial \bar{z}_j}\right)\Big|_{dr=0} > 0.$$

The important thing is that in the complex case, convexity is preserved by biholomorphism. (This is because coordinate transformation is linear only in z_j.)

Now consider the subspace of $\mathbb{C} - \{0\}$ given by

$$\{(z,w) : |z|, |w| < b\} \setminus \{(z,w) : |z|, |w| \leq a\}.$$

To compute its sheaf cohomology, we can use Čech cohomology, and then the function $\frac{1}{zw}$ is one that corresponds to $d\theta$ that does not have a global solution. We can use this and the convexity of the ball in $\mathbb{C}^2$ to construct Lewy's example.

After this, people tried find criteria for local solvability of a single equation. Franc̨ois Trèves and Louis Nirenberg got some cases. They had some good estimates on the constant coefficient case, and then freezing coefficients to get good approximations. In the mid 70s, Beals and Fefferman got a complete solution.

Consider a unit ball B in $\mathbb{C}^2$. This has real dimension $2n-1$. This can locally be given a structure of a group, by first approximating it by $\mathbb{R}^{2n+1}$ and then using addition.

Consider the Cayley transform, given by

$$w_n = \frac{i - z_n}{i + z_n}, \quad w_k = \frac{2iz_k}{i + z_n}$$

for $1 \leq k \leq n-1$. The $(w_1, \dots, w_n)$ in the ball corresponds to $(z_1, \dots, z_n)$ in the upper half plane. This is because

$$-1 + \sum_{k=1}^{n} |w_k|^2 = \frac{1}{|i + z_n|^2}\left(|i - z_n|^2 - |i + z_n|^2 + \sum_{k=1}^{n-1} |2iz_k|^2\right).$$

The boundary is cut out by the function $\rho = |z'|^2 - \Im(z_n)$, and so the tangential vectors are of the form

$$\xi = a\frac{\partial}{\partial \bar{z}_1} - 2iaz_1\frac{\partial}{\partial \bar{z}_2}.$$

Then the differential operator comes from $L = \frac{\partial}{\partial \bar{z}_1} - 2iz_1\frac{\partial}{\partial \bar{z}_2}$, and then this gives the operator L we want.

3.2 Fourier analysis over complex variables

Consider a function

$$f(\theta) = \sum_{n=-\infty}^{\infty} c_n e^{in\theta}$$

on the boundary of the disk $\mathbb{D}$. We can fill the inside of the disk by looking at

$$\sum_{n=0}^{\infty} c_n r^n e^{in\theta}.$$

This is the sort of the fundamental theorem of calculus (or mean value property) that works by multiplication. In particular, the sum $\sum_n r^{|n|}e^{in\theta}$ gives the Dirichlet kernel.

But we want to fill in the disk. So we look at the Bergman kernel, so that for L^2-holomorphic $f \in \mathscr{O}_{L^2}(\mathbb{D})$,

$$f(z_0) = \int_{\mathbb{D}} B(z_0, w) f(w).$$

This $B(z_0, w)$ is the Bergman kernel can be explicitly written down.

Now we want to make consider space of holomorphic function on $\mathbb{D}^2$. The **Hardy space** is the Hilbert space of functions with

$$\sum_n |c_n|^2 = \lim_{r \to 1-} \sum |c_n|^2 r^{2n}$$

finite. We can bring this to the upper half plane. Here, we consider holomorphic functions f such that $\sup \hat{f}$ is in $[0, \infty)$ and $\int_{\mathbb{R}} |\hat{f}|^2 < \infty$ as we integrate along $\mathbb{R} + i\tau$ for $\tau \to 0$.

Theorem 3.1 (Stein, p. 126)**.** *Let g and $\hat{g}$ be functions on $\mathbb{R}$ with moderate decay, i.e.,*

$$|g|, |\hat{g}| \leq \frac{A}{1 + x^2}$$

as $x \to \infty$. Then g can be extended to a bounded function on $\Im(z) \geq 0$ that is holomorphic on $\Im(z) > 0$ if and only if $\hat{g}(\xi) = 0$ for $\xi < 0$.

So we have several spaces now. There is the space $\mathscr{O}_{L^2}(\mathbb{D})$, and Riesz representation on the space gives the Bergman kernel. There is also the space $H^2(\mathbb{D})$, and using the Riesz representation theorem gives the Szegö kernel. Then

$$f(z) = \frac{1}{2\pi i} \int_{|w|=1} \frac{f(w)dw}{w - z} = \frac{1}{2\pi} \int \frac{f(w)d\theta}{z - \overline{w}}.$$

This recovers the interior from the boundary. For the upper half plane, the Szegö kernel looks like

$$\frac{1}{2\pi i} \frac{1}{u - z}.$$

Now we are interested in the situation of $\Im(z_n) > |z'|^2$. For z_n, this is a shifted upper half plane that is parametrized by z'.

4 February 1, 2018

Last time we have been talking about the Hans Lewy example. The Cayley transform gave a biholomorphic map between

$$\sum_j |w_j|^2 \leq 1$$

and the space $U = \{z : \Im z_n > |z'|^2\}$. Then there was a nontrivial cohomology class. If we denote by $\mathscr{A}_M^{0,k}$ the space of smooth $(0,k)$-forms,

$$\omega = \sum_{1 \leq p \leq n-1} a_{j_1 \cdots j_p} d\bar{z}_{j_1} \wedge \cdots \wedge d\bar{z}_{j_p},$$

we have a complex

$$0 \to \mathbb{C} \to \mathscr{A}_M^{0,0} \xrightarrow{\partial_b} \mathscr{A}_M^{0,1} \xrightarrow{\partial_b} \cdots \xrightarrow{\partial_b} \mathscr{A}_M^{0,n-1} \to 0.$$

Here, ∂_b is defined in the following way. We first regard it as a function on $\mathbb{C}^n$, and then take $\bar{\partial}$, and show that it is actually independent of the choice of the extension. Then we get

$$\bar{\partial}_b \omega = \sum (\bar{\partial} a_{j_1 \cdots j_p}) d\bar{z}_{j_1} \wedge \cdots d\bar{z}_{j_p}.$$

For $n = 2$, we have

$$0 \to \mathbb{C} \hookrightarrow \mathscr{A}^{0,0} \xrightarrow{\bar{\partial}_b} \mathscr{A}^{0,1} \to 0.$$

Then for $u \in \mathscr{A}^{0,0}$ and $f d\bar{z}_1 \in \mathscr{A}^{0,1}$, we would like to solve the equation

$$\bar{\partial}_b u = f d\bar{z}_1.$$

The operator L is chosen sot that this is equivalent to $Lu = f$.

4.1 Nonsolvability of Hans Lewy's equation

In this case, $\bar{L} = \partial_b$. The explicit f is

$$f(z_2) = \exp\Big(-\Big(\frac{z_2}{i}\Big)^{1/2} - \Big(\frac{i}{z_2}\Big)^{1/2}\Big).$$

First of all, the square root is well-defined because we have a branch. The square is introduced in the following reason. We need to use the fact that $e^{-\lambda}$ decays rapidly as $\Re(\lambda) \to \infty$. So we want that the real part goes to infinity fast. So if we take the square root, we can achieve this. By the other factor, we also have this at 0. That is, f is rapidly decreasing at ∞ and at 0 as well.

The more important reason is that it cannot be analytic at $z_2 = 0$. The function f is holomorphic in U, and C^∞ up to $\overline{U}$, but it cannot be extended holomorphically at 0 to $\mathbb{C}^2$.

Proposition 4.1. *For any distribution u with compact support, if $\bar{L}u = f$ in a neighborhood of 0 in ∂U, then f has a holomorphic extension to a neighborhood of 0.*

The main idea of this proof is that holomorphicity can be checked by taking the dual, by testing against something and showing that it is zero. This is not surprising, because a lot of cohomology theory works by the dual, like Poincaré duality.

Consider the open subset $U \subseteq \mathbb{C}^n$ the region of vertically shifted z_n by $|z'|^2$. We actually work in ∂U. But the contradiction should be at the boundary 0, so we need to relate ∂U to U in the study of holomorphic functions. The tool is the Cauchy–Szegö kernel.

4.2 Cauchy–Szegö kernel

Because it is a vertically shifted upper half-plane, let us focus on this variable and let this be $\mathbb{H} + i\eta$ for $\eta > 0$. In the case of a disk, we had that for

$$f(z) = \sum_{n \geq 0} c_n z^n,$$

c_n is the Fourier transform and f is the inverse Fourier transform. So to get the interior values, we first take the Fourier transform, introduce a factor of r^n, and the take the inverse Fourier transform.

On the upper half plane, we do the same thing. Consider a function $G_0(x)$ on the boundary $\mathbb{R} = \partial\mathbb{H}$. The inverse Fourier transform is given by

$$G_0(x) = \int_{\xi \in \mathbb{R}} \hat{G}_0(\xi) e^{2\pi i \xi x} d\xi.$$

Then to extend it to the complex plane, we simply set

$$G_0(z) = \int_{\xi \in \mathbb{R}} \hat{G}_0(\xi) e^{2\pi i \xi z} dz.$$

Let us do this with the vertical shift now. The boundary values will can be written as $x \mapsto G_0(x + i\eta)$. Then

$$\hat{G}_0(\xi) = \int_{x \in \mathbb{R}} G_0(x + i\eta) e^{-2\pi i \xi x} dx.$$

Then we get

$$\begin{aligned} G(z) &= \int_{\xi \in \mathbb{R}} \left(\int_{x \in \mathbb{R}} G_0(x + i\eta) e^{-2\pi i \xi x} dx \right) e^{2\pi i (z - i\eta)\xi} d\xi \\ &= \int_{\xi \in \mathbb{R}} \left(\int_{x \in \mathbb{R}} G_0(x + i\eta) e^{-2\pi i (x + i\eta)\xi} dx \right) e^{2\pi i z \xi} d\xi. \end{aligned}$$

So given a function F holomorphic on U with boundary value F_0, we consider the "shifted" Fourier transform of F_0 as

$$f(z', x_n) = \int_{u_n \in \mathbb{R}} F_0(z', u_n + i|z'|^2) e^{-2\pi i x_n (u_n + i|z'|^2)} du_n,$$

and from this we can recover F as

$$F(z', z_n) = \int_{\lambda=0}^{\infty} f(z', \lambda) e^{2\pi i \lambda z_n} d\lambda.$$

We want to determine the kernel. There is a trick to to write down the Cauchy–Szegö kernel, coming from the Gamma function. We first polarize $\rho(z) = |z'|^2 - \Im(z_n)$ to get $r(z, w) = \frac{i}{2}(\bar{w} - z_n) - z'\bar{w}$ and define

$$S(z, w) = \frac{(n-1)!}{(4\pi r(z, w))^n}.$$

Then the claim is that

$$F(z) = \int_{(w', u_n) \in \mathbb{C}^{n-1} \times \mathbb{R}} S(z, w) F_0(w', u_n + i|w'|^2) dm(w', u_n).$$

Now here is the punchline. Define $Cf = \int_{\partial U} Sf$. Given f on ∂U, we can define and we can do something like a Cauchy integral formula. If f is smooth and rapidly decreasing on ∂U and we can form Cf, and also that $f = \bar{L}u$ for some distribution around 0, then Cf is holomorphic in a neighborhood of 0. If we manage to prove it, then Cf is a holomorphic extension of f around 0. That is precisely what we made sure do not happen.

Why does Cf extend to around 0? Assume $n = 2$. If u had compact support, we would have had

$$\begin{aligned} C(f) &= (f(w), S(z, u_2 + i|w_1|^2))_{(\partial U, dm(\mathbb{R}^{2n-1}))} \\ &= (\bar{L}_W u, S(z, u_2 + i|w_1|^2)) = -(U, \bar{L}_W S(z, u_2 + i|w_1|^2)) = 0 \end{aligned}$$

because S is anti-holomorphic in w. But this is going to be not true, but true only on a neighborhood. Then this works up to a cut-off function, and this is holomorphic in z. Then we get $C(f)$ as holomorphic near 0.

5 February 6, 2018

We have seen how to solve constant coefficients PDEs, and then looked at an example of a variable coefficient equation that cannot be solved. Then there is the question of which equations can be solved. Of course, not all can be solved, and there is the Nirenberg–Trèves criterion. This involves a Lie bracket.

Everybody knows the equation

$$\frac{\partial u}{\partial x} = P, \quad \frac{\partial u}{\partial y} = Q,$$

and there is the compatibility condition $\frac{\partial Q}{\partial x} = \frac{\partial P}{\partial y}$. But for one equation, what is the obstruction? In the complex case, the obstruction comes from $[L, \overline{L}]$. When I met him in the 70s, Sato claimed that the world depends on this compatibility condition, and looked at the space of solvable equations, called the Sato Grassmanian.

5.1 Existence of Bernstein–Sato polynomials

This is done purely algebraically, by counting the dimension. If I have an operator $A : V \to V$, there is a characteristic polynomial, and by dimension argument, there is a relation. Here, we are going to count dimension over the Weyl ring

$$D_N = \mathbb{C}[x_1, \dots, x_N, \partial_1, \dots, \partial_N].$$

Bernstein's contribution is to count using Hilbert polynomials, and the use of Fourier transforms.

The technical difficulty here is that the ring is noncommutative. The noncommutative part is given by

$$\left[x_j, \frac{\partial}{\partial x_k}\right] = -\delta_{jk}.$$

So these should all be regarded as operators. What they do operate on? Let k be a field with characteristic 0. We consider the D_N-modules.

First let us look at counting. We introduce a filtration with index n given by

$$D^n = \left\{ \sum_{|\alpha|+|\beta| \le n} a_{\alpha\beta} x^\alpha \partial_x^\beta \right\}.$$

Then we have $\cdots \subseteq D_n \subseteq \cdots$, and we can count $\dim_k D^n$ easily. Then the assignment $n \mapsto \dim_k D^n$ is a polynomial in n, which are known as Hilbert polynomials. But we have more problems, because the bracket gives you something in lower dimension. Then quotienting out gives a graded ring

$$\bigoplus_{n=1}^{\infty} D^n / D^{n-1}.$$

Let M be a finitely generated graded D-module. Then we can again look at the assignment

$$n \mapsto \dim_k M^n$$

that is a polynomial for sufficiently large n. We can ask about the degree of this polynomial $d(M)$, and also the leading coefficient of the polynomial $e(M)/d(M)!$. (This is to make sure that $e(M)$ is an integer.) Bernstein used this two number (d, e) to count.

Proposition 5.1. *Given a graded D-module, either $M = 0$ or $d(M) \geq N$ (where $k = \mathbb{C}$).*

The idea is that if $d(M) < N$, given $f_1, \ldots, f_k$ there exist many differential equations with polynomials coefficients satisfied by f_j. Then you should be able to cancel them out to get nothing. In our case, the number we are interested in is $D_N f^s$. Then it turns out that $d(M) = N$. After this, $e(M)$ is the only number that matters, and so we can use this as a notion of counting.

Proof. We apply induction on N. Consider a D_N-module M such that $d(M) < N$. Let $t = x_N$, and we will want to quotient by t. First, we claim that there exists some α such that the operator $t - \alpha$ is not invertible. This is because otherwise we can consider M as a module over $k(t)$, in which case M is too large.

Next, we show that $(t - \alpha)M = M$ for all $\alpha \in k$. This is because if $L = M/(t - \alpha)M$ then $d(L) < N - 1$ and so $L = 0$. So there should be $\ker(t - \alpha)$ should be nonzero for some $\alpha \in k$. Replace $t - \alpha$ by t so that we assume $\alpha = 0$. Consider the submodule

$$L = \{f \in M : t^n f = 0 \text{ for large } n\}$$

This should be nonzero, and $d(L) \leq d(M) < N$. This means that we can replace L by M.

Now we use the idea of Fourier transform. This process should switch the role of t and ∂_t. Note that $\ker(\frac{\partial}{\partial t} - \alpha) = 0$ for all $\alpha \in k$, because $t^n f = 0$ and $(\frac{\partial}{\partial t} - \alpha) f = 0$ implies that

$$0 = \Big[t^n, \frac{\partial}{\partial t} - \alpha\Big] f = n t^{n-1} f$$

and similarly we can reduce the degree.

There is a Fourier transform automorphism

$$\rho : D_N \to D_N; \quad \rho(x_j) = x_j, \quad \rho(\partial_j) = \partial_j, \quad \rho(t) = \partial_t, \quad \partial_t = -t.$$

Using this, we can give another module structure on L, which satisfies

$$\ker\Big(\frac{\partial}{\partial t} - \alpha\Big) = 0$$

for all α. This is what we said that is impossible. ∎

Now we have $d(M) = N$, and we look at $e(M)$. This shows that there are only a finite number of submodules. So the submodules $Df^s, Df^{s+1}, \ldots$ should stabilize.

5.2 Using reflection to get holomorphicity

We had this Cauchy–Szegö kernel. How is this related to $\bar{\partial}_b$? The contradiction came from the fact that if $\overline{L}u = f$ then $Cf = f$ and Cf can be holomorphically extendable. Then f is not analytic so gives a contradiction. The function Cf was defined by taking the inner product with $S(z, w)$ that is holomorphic in z and anti-holomorphic in w.

Now what is the Schwartz relfection? If you have a function f on $H \cap B_1$, and f is 0 on the real line segment, then you say that f can be holomorphically extended as $f(\bar{z}) = \overline{f(z)}$. Then you use the Cauchy integral to show that it is holomorphic. The Cauchy–Szegö kernel is supposed to give describe this process, but with

$$\mathcal{U} = \bigcup_{z'} H + |z'|^2.$$

6 February 8, 2018

We were talking about the Bernstein–Sato polynomial. Let k be a field of characteristic zero, and let $D(k) = k[x_1, \dots, x_n, \partial_1, \dots, \partial_n]$. For a polynomial $p \in k[x_i]$, we want an operator $L \in D[s]$ such that

$$Lp^{s+1} = b(s)p^s,$$

with minimal degree for $b(s)$. This is done by looking at the sequence

$$D(k(s))p^{s+1} \subseteq D(k(s))p^s \subseteq \cdots .$$

We want this to stabilize, because once $D(k(s))p^{s+k} = D(k(s))p^{s+k-1}$ then $p^{s+k-1} = Lp^{s+k}$.

You can check that these $D(k(s))p^{s+1}$ has $d \le N$, which implies $d = N$ by the inequality we proved last time. Also, you can give an explicit bound on e by just looking at polynomials, with a specific graded structure. At each quotient

$$\frac{D_N(k(s))p^{s+k}}{D_N(k(s))p^{s+k+1}},$$

this quotient should be either 0 or have $d \ge N$. In the nonzero case, it should have $e = 1$. This shows that either the quotient is 0 or e should increase by 1. This cannot happen infinitely often, so it should stabilize.

6.1 Solving differential equations with variable coefficients

The first breakthrough was by Hans Lewy, who found this example of $\overline{L}u = f$ that is not solvable. So people started to think about when a linear equation is solvable.

First, given a general operator $Lu = f$, you can write $L = L_1 + iL_2$ and then if L_1 and L_2 are linearly independent and $[L_1, L_2] \subseteq \mathbb{R}L_1 + \mathbb{R}L_2$ at every point, then we can integrate using Frobenius's theorem and make it to $\bar{\partial}_{x_1+ix_2}u = f$. Then you can use Cauchy's kernel.

So the story becomes interesting in the case when Frobenius has some linear dependency. For instance, consider

$$\frac{\partial}{\partial x_j} + ix_1^k \frac{\partial}{\partial x_2}.$$

If k is even, you can change variables as $x_j = t^{k+1}/k + 1$, so you can solve it. But for k odd, you can't do this. An amazing thing was when Nirenberg–Trèves attempted to write down a condition in 1963.

First they did a coordinate transformation to normalize the operator to

$$L_0 = \frac{\partial}{\partial t} + \sum_{k=1}^{n} b^k(x,t) \frac{\partial}{\partial x_k}.$$

Then

$$\Big[\frac{\partial}{\partial t}, \sum_k b^k \frac{\partial}{\partial x^k}\Big] = \sum_k (\partial_t b^k) \frac{\partial}{\partial x_k}.$$

The condition they gave was

(P) $\vec{b}(x,t) = |\vec{b}(x,t)| \cdot \vec{v}(x)$

This formulation can be generalized to n variables using null bicharacteristic. Given a differential operator L, we can write $Lu = f$ formally as $P\hat{u} = \hat{f}$. Here, $P(x,D)$ would be something that is expressed in terms of x and $D_j = \frac{1}{i}\frac{\partial}{\partial x_j}$. We replace D_j with ξ_j, and this will give a symbol $P(x,\xi)$.

Let $A = \Re P$ and $B = \Im P$. A curve in (x,ξ) is called **null bicharacteristic** if

$$\begin{cases} \dot{x}_j = \frac{\partial A}{\partial \xi_j}, \\ \dot{\xi}_j = -\frac{\partial A}{\partial x_j} \end{cases}$$

in $A = 0$. Now the new condition will read:

(P) The sign of B does not change on any null bicharacteristic curve in A.

Let's first check that these two conditions are equivalent for dimension 2. If we have

$$L_0 = \frac{\partial}{\partial t} + i \sum b_k(x,t) \frac{\partial}{\partial x_k} = i\Big(\frac{1}{i}\partial_t\Big) - \sum b_k \Big(\frac{1}{i}\frac{\partial}{\partial x_k}\Big),$$

then the symbol will be

$$P = \tau + i \sum_{k=1}^n b_k \xi_k.$$

Then $A = \tau$ and $B = \sum_{k=1}^n b_k(x,t)\xi_k$, and null bicharacteristic curves can only change in t. It is then clear that the sign of B cannot change at all.

7 February 13, 2018

Nirenberg–Tréves in Comm. Pure&App. Math. **16** (1963), 331–351, conjectured a solvability criterion for differential equations.

Let L be a partial differential operator of order $m \geq 1$. Consider the principal symbol P obtained by replacing $\frac{1}{i}\frac{\partial}{\partial x_j} \leftrightarrow \xi_j$. The operator is called **principal type** if

$$\nabla_\xi p(x,\xi) \neq 0$$

on $p(x,\xi) = 0$ and $\xi \neq 0$. A null bicharacteristic is a Hamilton vector flow

$$\dot{x}_j = \frac{\partial \Re p}{\partial \xi_j}, \quad \dot{\xi}_j = -\frac{\partial \Re p}{\partial x_j}.$$

The criterion for local solvability around x^* is that $\Im p$ does not change sign on $U \times (\mathbb{R}^n - \{0\})$ where U is a neighborhood of x^*.

Now any first-order differential equation can be written as

$$L = \frac{\partial}{\partial t} + i \sum_k b^k(x,t) \frac{\partial}{\partial x^k}$$

on $\mathbb{R}^{n+1}$. Then the principal part of $-iL$ is

$$P = \tau + i \sum_{k=1}^n b^k(x,t)\xi_k.$$

In this case, $\Re(p) = \tau$ and $\Im(p) = \vec{b} \cdot \vec{\xi}$ and so the criterion is equivalent to that

$$\vec{b}(x,t) = |\vec{b}(x,t)| \vec{v}(x)$$

has direction that is t-independent.

7.1 Riesz representation revisited

The idea is to modify the Riesz representation theorem. This states that if $\Phi : H \to \mathbb{C}$ is continuous, there exists an u such that $\Phi(\varphi) = (\varphi, u)_H$. To show that L is surjective, it suffices to show an estimate

$$\|\varphi\| \leq C\|L^*\varphi\|.$$

If we want to solve $Lu = f$, and we have this estimate, we first consider the map $L^*\varphi \mapsto (\varphi, f)$ with

$$|(\varphi, f)| \leq \|\varphi\|\|f\| \leq C\|L^*\varphi\|$$

and then extend it to $H \to \mathbb{C}$ with the same bound. Then you can use Riesz representation.

In this case, we will only be able to get

$$\|\varphi\|_{L^2} \le C\|L^*\varphi\|_{L^2_1}.$$

If we run the same argument, we are going to get a $u \in L^2_1$ such that

$$(\varphi, f) = (L^*\varphi, u)_1 = (L^*\varphi, u) + (DL^*\varphi, Du) = (\varphi, Lu + LD^*Du)$$

formally. Then you have $L(1 + D^*D)u = f$, but this is hard to justify. This is why we use modify the Riesz representation.

Consider $L^2_1 \subseteq L^2$, and

$$\|\psi\|_{H_{(-1)}} = \sup_{\|\psi\|\le 1} |(\varphi, \psi)|.$$

We can understand this space by using the Fourier transform. Then differentiation becomes coordinate multiplication, and Fourier transform preserves the inner product. Then $\|\varphi\|_1 \le 1$ is like

$$\int |\hat{\varphi}|^2(1 + |\xi|^2) \le 1,$$

and then its dual is like

$$\int |\hat{\psi}|^2(1 + |\xi|^2)^{-1} \le 1.$$

Then when $\|\varphi\|_0 \le C\|L^*\varphi\|_1$ shows that $Lu = f$ can be solved for $f \in L^2$ with $u \in L^2_{-1}$. This space $H_{-1} = L^2_{-1}$ can be understood as something sitting inside $\mathcal{S}'$.

7.2 First-order case of Nirenberg–Trèves

Now let

$$L^* = A + iB + c$$

so that $A = \frac{\partial}{\partial t}$, $B = \vec{b}(x,t)\vec{\partial}_x$, and $c = c(x,t)$ is complex-valued.

The first trick is to make c purely imaginary by using a weight function. The weight function one uses is to choose h such that $\frac{\partial h}{\partial t} = \Re(c)$. Then

$$\begin{aligned} e^{-h}L^*(e^h\varphi) &= e^{-h}(-\partial_t(e^h\varphi) + ib\cdot\partial_x(e^h\varphi) + ce^h\varphi) \\ &= -\partial - t\varphi + ib\cdot\partial_x\varphi - \varphi\tfrac{\partial h}{\partial t} + c\varphi + i\varphi\vec{b}\cdot\vec{\partial}_x h. \end{aligned}$$

So we have purely imaginary c now.

Solvability is local with respect to compact support on the right hand side. So we can only worry about test functions

$$\varphi \in C_0^\infty(\Omega)$$

where Ω is a fixed neighborhood of 0 in $\mathbb{R}^{n+1}$.

Let's first look at the trivial case $B = 0$. We may assume $\Omega \subseteq (-\frac{\delta}{2}, \frac{\delta}{2})$. Then

$$\varphi(x,t) = \int_{-\frac{\delta}{2}}^{x} \partial_t \varphi dt$$

and so

$$|\varphi(x,t)|^2 \leq \left| \int_{-\frac{\delta}{2}}^{x} (\partial_t \varphi) dt \right|^2 \leq \delta \int_{-\frac{\delta}{2}}^{\frac{\delta}{2}} |\partial_t \varphi|^2 dt.$$

So the troublesome term is the cross terms, measured by

$$C_1 = [A, B].$$

We can first compute

$$\begin{aligned} \|L^* \varphi\|^2 &= \|(A + iB + c)\varphi\|^2 \\ &= \|A\varphi\|^2 + \|B\varphi\|^2 + \|c\varphi\|^2 + 2\Re(A\varphi, iB\varphi) + 2\Re(A\varphi, c\varphi) + 2\Re(iB\varphi, c\varphi). \end{aligned}$$

Then

$$2\Re(A\varphi, iB\varphi) = (A\varphi, iB\varphi) + (iB\varphi, A\varphi) = -i(B^* A\varphi, \varphi) + i(A^* B, \varphi, \varphi).$$

Clearly $A^* = -A$, but

$$\int (B^* \varphi)\overline{\psi} = \int \varphi \overline{\vec{b} \cdot \partial_x \psi} = -\int (\vec{b} \cdot \partial_x \varphi)\overline{\psi} - \int (\operatorname{div}_x \vec{b})\varphi\overline{\psi}$$

and so $B^* = -B - \operatorname{div}_x \vec{b}$. Then

$$2\Re(A\varphi, i\beta\varphi) = -i(-BA\varphi - (\operatorname{div}_x \vec{b})A\varphi, \varphi) - i(AB\varphi, \varphi) = -i(C_1\varphi, \varphi) - i(\varphi, d\varphi)$$

where $d = -\operatorname{div}_x \vec{b}$. Now $\Re(A\varphi, c\varphi)$ is bound by $\|A\varphi\|\|\varphi\|$ and $\Re(iB\varphi, c\varphi)$ is bound by $\|B\varphi\|\|\varphi\|$.

Now how do we take care of C_1? We have

$$C_1 = AB - BA = \partial_t(\vec{b} \cdot \partial_x) - (\vec{b} \cdot \partial_x)\partial_t = (\partial_t \vec{b}) \cdot \partial_x.$$

How do we dominate $(C_1\varphi, \varphi)$? Here the idea is that we can dominate $|\vec{b}_t \cdot \partial_x \varphi|^2$ by $(\vec{b} \cdot \partial_t \varphi)(\vec{b}_t \cdot \partial_t \varphi)$ (after Fourier transform). But $\vec{b}_{tt}$ looks harder to control.

7.3 Technical lemma

Introduce a neighborhood $\Omega = (-\alpha, \beta) \times N_0$ where $\alpha, \beta > 0$ and N_0 is a neighborho of 0 in $\mathbb{R}^N$. For $J = (-\alpha, \beta)$ we let

$$\sigma(x) = \sup_{t \in J} |\vec{b}_{tt}(x,t)|.$$

Let

$$\vec{m}(x,t) = \vec{b}(x,t) + \sigma(x)\vec{v}(x)$$

where $\vec{b}(x,t) = |\vec{b}(x,t)|\vec{v}(x)$.

Lemma 7.1. *Let $\epsilon > 0$ and $C_\epsilon = 2(1+\frac{2}{\epsilon^2})$. Let $J = (-\alpha, \beta)$ and $K_\delta = (-\alpha+\epsilon, \beta-\epsilon)$. Let N_0 be a convex open neighborhood of 0 in $\mathbb{R}^n$. Then for all $(x,t) \in N_0 \times K_\epsilon$ and $\vec{\zeta} \in \mathbb{C}^n$,*

$$|\vec{b}_t(x,t)\cdot\vec{\zeta}|^2 \le C_\epsilon(\vec{b}(x,t)\cdot\vec{\zeta})\overline{\vec{m}(x,t)\cdot\vec{\zeta}}.$$

Proof. Let $\varphi(x,t) = |\vec{b}(x,t)|$ and fix $x \in N_0$. Since $\vec{v}(x) = 1$, we have $\rho(x,t)$ is C^∞. Then we can write

$$\vec{b}(x,t) = \rho(x,t)\vec{v}(x), \quad \vec{b}_t(x,t) = \rho_t(x,t)\vec{v}(x), \quad \vec{b}_{tt}(x,t) = \rho_{tt}(x,t)\vec{v}(x).$$

Then Taylor expansion gives

$$\rho(x,t) + h\rho_t(x,t) + \frac{1}{2}h^2\rho_{tt}(x, t\theta h) = \rho(x,t+h) \ge 0.$$

Now we want to do something like looking at discriminant. We move the middle term to one side to get

$$|\rho_t(x,t)| \le \frac{\rho(x,t)}{|h|} + \frac{1}{2}|h|\sigma(x).$$

Now we choose h suitably.

If there exists a $0 < \delta < \epsilon$ such that

$$\frac{1}{\delta}\rho(x,t) = \frac{1}{2}\delta\sigma(x)$$

then we can plug in $|h| = \delta$ and get

$$\rho(x,t)^2 \le 2\rho(x,t)\sigma(x).$$

Then we get the lemma.

In the other case, we have

$$\frac{\rho(x,t)}{\epsilon} \ge \frac{1}{2}\epsilon\sigma(x)$$

and so

$$\rho_t^2(x,t) \le 2\Big(1+\frac{2}{\epsilon^2}\Big)\rho(x,t)(\rho(x,t)+\sigma(x)).$$

Then we get the same inequality. □

If we set $M = \vec{m}(x,t)\cdot\vec{\partial}_x$ with $\vec{m} = \vec{b} + \sigma(x)\vec{v}$, we get

$$[A, M] = AM - MA = AB - BA = C_1$$

because $\sigma(x)$ has no t-dependence.

8 February 15, 2018

Last time we were talking about the solvability criterion of Nirenberg–Trèves. To solve this, we needed the solvability estimate

$$\|\varphi\|_0 \le C\|L^*\varphi\|_1$$

for all $\varphi \in C_0^\infty$. To do this, we introduced the space $\mathcal{H}_{(-1)}$.

Definition 8.1. The space $\mathcal{H}_{(-m)}(\Omega)$ consists of all linear combinations of derivatives of L^2 functions up to order m. Then for $T \in \mathcal{H}_{(-m)}$, we define

$$\|T\|_{\mathcal{H}_{(-m)}(\Omega)} = \sup\{\|T\varphi\| : \varphi \in C_0^\infty(\varphi), \|\varphi\|_m \le 1\}.$$

Under the L^2-norm, this is sort of the dual of L^2_m.

For nonsovlability, we are going to use the following strategy. Suppose that $Lu = f$ is solvable on Ω. Then there exists an $C > 0$ and $k, l \ge 1$ such that

$$\left|\int_\Omega gv\right| \le C\biggl(\sum_{|\alpha|\le k} \sup_{x\in\Omega}|D_x^\alpha g|\biggr)\biggl(\sum_{|\beta|\le l} \sup_{x\in\Omega}|D_x^\beta L^* u|\biggr)$$

for all $g, v \in C_0^\infty(\Omega)$. So we want to find a sequence of functions so that the left hand side stays finite while the right hand side goes to infinity.

8.1 Solving the first-order equation without sign change

Let us write

$$L^* = A + iB + c = \partial_t + i\vec{b}(x,t)\partial_x + c$$

where c can be set to be purely imaginary. We want to prove an estimate like $\|\varphi\|_0 \le C\|L^*\varphi\|_1$. First we have

$$\|\varphi\|_0 \le \delta\|A\varphi\|_0$$

by fundamental theorem of calculus. Here, we note that

$$\|L^*\varphi\|_0^2 = \|A\varphi\|^2 + \|B\varphi\|^2 + \|c\varphi\|^2 + \cdots$$

where the cross terms $\cdots$ are those we want to estimate. We have

$$2\Re(\alpha\varphi, iB\varphi) = -i((C_1\varphi,\varphi) + (A\varphi, d\varphi))$$

where $d = -\operatorname{div}_x \vec{b}$. This is essentially the only thing we need to worry, because $\|\varphi\|$ is dominated by $\|A\varphi\|$ after making δ small enough, and we can use inequalities like

$$2|(A\varphi, c\varphi)| \le \epsilon\|A\varphi\|^2 + \frac{1}{\epsilon}\|c\varphi\|^2.$$

So at the end, we get an estimate like

$$\|A\varphi\|^2 + \|B\varphi\|^2 \le \|L^*\varphi\|^2 + K(\|C_1\varphi\| + \|\varphi\|)\|\varphi\|.$$

The hard thing is dealing with $C_1\varphi$. But the lemma we had last time gives

$$\|C_1\varphi\|^2 \le C_\epsilon(B\varphi, M\varphi)$$

where $M = \vec{m}(x,t)\partial_x$. We now have

$$(L^*\varphi, M\varphi) = (A\varphi, M\varphi) - i(B\varphi, M\varphi) + (c\varphi, M\varphi).$$

But B and M point in the same direction, so $(B\varphi, M\varphi)$ is real. So

$$\Re(B\varphi, M\varphi) = -\Im(L^*\varphi, M\varphi) - \Im(A\varphi, M\varphi) + \Im(c\varphi, M\varphi).$$

Now note that

$$-2\Im(A\varphi, M\varphi) = i((M^*A - A^*M)\varphi, \varphi) = i(C_1\varphi, \varphi) + i(A\varphi, k\varphi)$$

where k is a scalar. Also,

$$-2\Im(M\varphi, c\varphi) = i((\bar{c}M\varphi, \varphi) - (M^*(c\varphi, \varphi))) = (i(Mc)\varphi, \varphi) - (ic\varphi, k\varphi).$$

If we put things together, we get

$$\|C_1\varphi\|^2 \le K\|\varphi\|(\|M^*L^*\varphi\| + \|A\varphi\| + \|\varphi\|).$$

There is an extra $|(C_1\varphi, \varphi)|$ appears on the right hand side, but we can absorb it with $\epsilon\|C_1\varphi\|^2 + \frac{1}{\epsilon}\|\varphi\|^2$.

But here, note that $M^* = -M + k$ where you can show that k is L^∞ because $\vec{m}$ is Lipschitz. So you can show that

$$\|M^*L^*\varphi\| \le K\|L^*\varphi\|_1$$

because M^* just has one derivative. So we get

$$\begin{aligned}\|A\varphi\|^2 + \|B\varphi\|^2 &\le \|L^*\varphi\|^2 + K(\|C_1\varphi\| + \|\varphi\|)\|\varphi\|, \\ \|C_1\varphi\|^2 &\le K\|\varphi\|(\|L^*\varphi\|_1 + \|A\varphi\| + \|\varphi\|).\end{aligned}$$

9 February 20, 2018

We had to prove solvability and also nonsolvability. For the solvability part, we wanted to prove the estimate

$$\|\varphi\|_0 \leq C\|L^*\varphi\|_1$$

where

$$L^* = \frac{\partial}{\partial t} + ib\vec{b}(x,t)\frac{\partial}{\partial x} + c = A + iB + c.$$

Then we were trying to solve the equation under the assumption

$$\vec{b}(x,t) = |\vec{b}(x,t)|\vec{v}(x).$$

We know that A is always solvable by the fundamental theorem of calculus. Local solvability means that we can shrink the domain as we want, and so we even have $\|\varphi\|_0 \leq \delta\|A\varphi\|_0$. When we expand $\|L^*\varphi\|^2$, we get pure terms and mixed terms, and the point is that the mixed terms were given by $\|C_1\varphi\|$ where $C_1 = [A, B] = \vec{b}_t\partial_x$. To deal with this, we used the $B^2 - 4AC$ argument

$$\|\vec{b}_t \cdot \vec{\zeta}\| \leq C_\epsilon(\vec{b} \cdot \vec{\zeta})\overline{((\vec{b} + \sigma\vec{v}) \cdot \vec{\zeta})}$$

where $\sigma(x) = \sup_t|\vec{b}_{tt}(x,t)|$. So if we let

$$M = \vec{m}\partial_x = B + \sigma\vec{v} \cdot \partial_x$$

then

$$\|C_1\varphi\|^2 \leq K(B\varphi, M\varphi).$$

Now we would like to replace $B\varphi$ by $L^*\varphi$, $A\varphi$, $c\varphi$. Because $[A, M] = C_1$, we can take care of $(A\varphi, M\varphi)$.

9.1 Insolvability of a differential equation

In the homework, there is the function

$$u(x) = e^{\tau(ix_1 - \frac{1}{2}|x|^2)}$$

on $x = (x_1, x_2, x_3)$ and a self-adjoint second-order partial differential operator with variable coefficients which are quadratic polynomials. Here, the dual is not injective, so it should not be surjective. This means that there is global nonsolvability for $\mathcal{S}(\mathbb{R}^3)$.

This holds the key for nonsolvability. We want to show that the estimate

$$\int fv \leq C \sup_{|\alpha|\leq k}|D^\alpha L^* f| \sup_{|\beta|\leq N}|D^\beta v|$$

fails. So we want to make the left hand side go to infinity where the right hand side is bounded. Here, if we use this function $e^{\tau(ix_1 - \frac{1}{2}|x|^2}$, we will get something like the Fourier transform of f.

Theorem 9.1 (3.1)**.** *Assume $\alpha > 0$ and $\vec{\xi} \in \mathbb{R}^n$ with $|\vec{\xi}| = 1$ such that $\vec{b}(0,t)\cdot\vec{\xi} > \alpha|\vec{b}(0,t)|$ for $t > 0$ and $\vec{b}(0,t)\cdot\vec{\xi} > \alpha|\vec{b}(0,t)|$ for $t < 0$. (Also $\vec{b}(0,0) = 0$.) Assume for some $C > 0$,*

$$\sum_j |\vec{b}_{x_j}(0,t)\cdot\vec{\xi}| \leq C|\vec{b}(0,t)|$$

for all $|t| < T$. Then L is not solvable at $(0,0)$.

We are going to construct a v such that L^*v is small but (v,f) is not small. That is, we want to violate

$$\int vf \leq C \sup|D^\alpha f| \sup|D^\beta L^* v|.$$

We seek $L^*u = O(|x^\rho|)$ by solving Cauchy–Kowalevski

$$u = -ix\cdot\vec{\xi} - |x|^2 + \sum_{j=1}^q u_j(x,t)$$

up to a certain order. Then we are going to rescale $v = e^{\tau u(\frac{x}{\tau},\frac{t}{\tau})}$.

Let's do this more carefully. Let us write

$$L^*u = L_0 u - g(x,t)$$

where $L_0 = \frac{\partial}{\partial t} + i\vec{b}\frac{\partial}{\partial x}$. We week u with $L_0 u = 0$ up to order $O(|x|^q)$ and also u with $L_0 u - g = 0$ up to some order $O(|x|^q)$. We use Cauchy–Kowalevski in t, with initial value $t = 0$ being $-ix\cdot\vec{\xi} - |x|^2$. Let us write

$$b^k = \sum_j b_j^k$$

where b_j^k is homogeneous of degree j. Then if we look at the first order, we should get

$$(u_0)_t + ib_0^k\Big(-i\xi_k + \frac{\partial u_1}{\partial x_k}\Big) = 0.$$

If we look at the second order term, we should get

$$(u_1)_t + ib_1^k\Big(-i\xi_k + \frac{\partial u_1}{\partial x_k}\Big) + ib_0^k\Big(-2x_k + \frac{\partial u_2}{\partial x_k}\Big) = 0.$$

In general, we are going to get

$$(u_j)_t + ib_j^k\Big(-i\xi_k + \frac{\partial u_1}{\partial x_k}\Big) + ib_{j-1}^k\Big(-2x_k + \frac{\partial u_2}{\partial x_k}\Big) + i\sum_{l=2}^j b_{j-l}\frac{\partial u_{l+1}}{\partial x_k} = 0.$$

But we are really interested in u_0, not $(u_0)_t$. To get u_0, we need to integrate. But when we integrate b^k in t, we get increasing as t goes away from 0. This is why we are able to control.

We actually want something more. Because we need $\int fv$ to be something like the Fourier transform of f, we want

$$\Im u \sim i\vec{x} \cdot \vec{\xi}.$$

Because we want L^*u part to vanish, we want

$$\Re u \leq -\frac{1}{8}|x|^2 - \frac{1}{2}\alpha\left|\int_0^t |\vec{b}(0,s)|ds\right|,$$

both for $t < 0$ and $t > 0$.

Suppose we have this u such that $\|L_0 u\| = O(|x|^q)$ and $\|L_0 h - g\| = O(|x|^q)$. Then we can define

$$v_\tau = \tau^{n+2+k} e^{\tau u + h}\varphi, \quad f_\tau = \frac{1}{\tau^h} F(\tau x, \tau t),$$

where φ is the cut-off function. We then have

$$\frac{1}{\tau}\int f_\tau v_\tau - \int F(x,t) e^{\tau u(\frac{x}{\tau},\frac{t}{\tau}) + h(\frac{x}{\tau},\frac{t}{\tau})} \varphi\Big(\frac{x}{\tau}, \frac{t}{\tau}\Big) dx dt.$$

Then the only thing left is $\hat{F}(\xi, 0)$, because φ is a cut-off.

10 February 22, 2018

For nonsolvability, we first looked at the Hans Lewy example $\bar{\partial}_b$, which depended on the kernel. After this become known, people used the failure of estimates to prove nonsolvability. This means that

$$\int fv \leq C\biggl(\sum_{|\alpha|\leq k} \sup|D^\alpha f|\biggr)\biggl(\sum_{|\beta|\leq l} \sup|D^\beta L^* v|\biggr)$$

for $f, v \in C_0^\infty$, fails. The idea is to use the adjoint, and get a solution for L^*. But we only need approximate solution of L^*.

For differential equations with real analytic coefficients, there is the Cauchy–Kowalevski. So we can get an approximate solution (up to order q) to L^*. The factor $e^{ix\cdot\xi}$ will make $\int fv$ something like a Fourier transform. So we set f as a function with nonzero Fourier transform, and then we use a cutoff function to make L^*v really small, not only at a neighborhood.

10.1 Introduction to higher order equations

Let me briefly talk about Hörmander's work on higher order differential equations. If $\xi_1, \dots, \xi_\ell$ are nondegenerate vector fields, and

$$[\xi_\mu, \xi_n] \in \langle \xi_1, \dots, \xi_\ell \rangle$$

then we can integrate this and make $\xi_j = \frac{\partial}{\partial x_j}$. But this is only for first-order PDEs. There is an issue if the vectors are degenerate, and there is also the problem that this can only deal with first-order. If we write $D_j = \frac{1}{i}\frac{\partial}{\partial x_j}$, then the differential operator can be written as

$$P(x, D).$$

Then we can look at the symbol as

$$P(x, \xi) = e^{-i\xi\cdot\xi} P(x, D) e^{ix\cdot\xi}.$$

This can be done in a coordinate-free way. The symbol is important because we have integration by parts.

We can look at the principal symbol $P_m(x, \xi)$ and its complex conjugate (of coefficients) $\overline{P}_m(x, \xi)$. We have previously taken the commutator $[\overline{L}, L]$. So we can similarly define the commutator

$$C_{2m-1}(x, D) = \overline{P}(x, D)P(x, D) - P(x, D)\overline{P}(x, D).$$

We say that that the differential operator is of **principal type** if the zeros of $P_m(x, \xi)$ have order 1. That is, when $\nabla_\xi P_m(x, \xi) \neq 0$ for any $\xi \neq 0$ with $P_m(x, \xi) = 0$.

Theorem 10.1 (Hörmander)**.** *The differential equation is solvable at $x = 0$ only when $C_{2m-1}(x, \xi) = 0$ on $P_m(x, \xi) = 0$.*

There is some linear algebra statement. Define the **Siegel upper half space** as the space of $n \times n$ symmetric matrix with complex entries Z, such that $\Im(Z)$ is positive definite. Then you can show that for $\vec{a} \cdot \vec{f} \in \mathbb{C}^2$, there exists an $Z \in \mathrm{Siegel}_n$ such that $Z\vec{a} = \vec{f}$ if and only if

$$\Im(\vec{f} \cdot \vec{a}) > 0.$$

After all this, we are going to look at Hörmander's criterion on hypoellipticity.

10.2 Getting the estimate on Cauchy–Kowalevski

We want to show that

$$\frac{\partial}{\partial t} + i\vec{b}(x,t) \cdot \partial_x$$

is not solvable, when the direction changes. Then we can pick a vector $\vec{\xi}$ such that $\vec{b} \cdot \vec{\xi}$ changes sign. Technically, we assume that

$$\begin{cases} \vec{b}(0,t) \cdot \vec{\xi} > \alpha|\vec{b}(0,t)| & \text{for } t > 0, \\ -\vec{b}(0,t) \cdot \vec{\xi} > \alpha|\vec{b}(0,t)| & \text{for } t < 0. \end{cases}$$

Also we assume that

$$\sum_j |\vec{b}_{x_j}(0,t) \cdot \vec{\xi}|^2 \leq C|\vec{b}(0,t)|.$$

We want an approximation solution to

$$u = -ix \cdot \xi - |x|^2 + \sum_{j=0}^{q-1} u_j(x,t)$$

so that $|L_0 u| \leq \mathrm{const}|x|^q$. Then we are going to set something like $v = e^{\tau u(\frac{x}{\tau}, \frac{t}{\tau})}$.

We do this Cauchy–Kowalevski with initial data $e^{-ix\cdot\xi - |x|^2}$ at $t = 0$. One thing we might worry about is the terms u_0, u_1, u_2 messing up the new input. But if we write down the equation, we get

$$\begin{cases} (u_0)_t + ib_0^k(-i\xi_k + \frac{\partial u_1}{\partial x^k}) = 0, \\ (u_1)_t + ib_1^k(-i\xi_k + \frac{\partial u_1}{\partial x^k}) + ib_0^k(-2x^k + \frac{\partial u_2}{\partial x^k}) = 0. \end{cases}$$

Here, $b = \sum b_i$ is the expansion of b in x. Now the $b_0^k \xi_k$ part is exactly $\vec{b}(0,t) \cdot \vec{\xi}$, and so we get

$$|\Re u_0(t)| \leq \frac{1}{2}\alpha \int_0^t |\vec{b}(0,s)| ds,$$

after integrating along t. Here, $\frac{\partial u_1}{\partial x^k}$ is a constant in x, so we can ignore this for t small.

Next, we have

$$\left|(u_1)_t + ib_1^k \frac{\partial u_1}{\partial x_k}\right| \leq C|b_0(t)|^{1/2}|x| + c_1|b_0(t)||x|$$

and then integrating again gives

$$|u_1| \leq c_2|x|\left|\int_0^t |b_2(s)|^{1/2} ds\right| \sum |\vec{b}_{x_j}(0,t)\vec{\xi}|^2 \leq c_1|x|\epsilon^{1/2}\left|\int_0^t |b(s)| ds\right|^{1/2}$$

for $|t| < \epsilon$.

11 February 27, 2018

We are going to talk about the higher order case, solved by Hörmander. This can be regarded as an analogue of Frobenius integrability. Then we can find coordinates $y_1, \dots, y_k$ such that $X_j = \frac{\partial}{y_j}$. There are several issues to think about:

- What if $X_1, \dots, X_k$ are not linearly independent?
- What about higher orders?

11.1 Hörmander's criterion

Let us consider

$$P(x, D) = \sum_{|\alpha| \le m} a^\alpha(x) D^\alpha, \quad D_j = \frac{1}{i} \frac{\partial}{\partial x_j},$$

where a^α are complex-valued functions. Take the principal symbol $P_m(x, \xi) = \sum_{|\alpha|=m} a^\alpha(x)\xi^\alpha$. We can also define $\bar{P}_m(x, \xi) = \sum_{|\alpha|=m} \bar{a}^\alpha(x)\xi^\alpha$. Then we can define the commutator

$$C(x, D) = \overline{P}(x, D)P(x, D) - P(x, D)\overline{P}(x, D)$$

of order $2m - 1$, and consider its principal symbol $C_{2m-1}(x, \xi)$.

Theorem 11.1 (Hörmander, necessary condition)**.** *If $P(x, D)u = f$ is always solvable, then $C_{2m-1}(x, \xi) = 0$ at $\xi \neq 0$ when $P_m(x, \xi) = 0$.*

What does it has to do with the change of sign? This theorem actually preceded Nirenberg–Trèves. Note that $C_{2m-1}(x, \xi) = 0$ is of odd order. So if this is not true, there exists some $\xi \neq 0$ such that $P_m(x, \xi) = 0$ and $C_{2m-1}(x, \xi) \neq 0$.

To prove this, we are going to do the same thing, but we need to do some linear algebra. This is about moving a fixed vector to the variable factor with the angular component the movement.

Lemma 11.2. *Let $\vec{a} \in \mathbb{C}^n - \{0\}$ be the initial fixed vector, and let $\vec{f} \in \mathbb{C}^n$ be the variable vector. Then there exists an $A = (\alpha_{kj}) \in \mathrm{GL}(n, \mathbb{C})$ symmetric with $\Im A > 0$ such that $A\vec{a} = \vec{f}$ if and only if $\Im(\vec{f}, \vec{a}) > 0$.*

Such A are the elements of the Siegel upper half space. This is the universal covering of the moduli space for

$$\mathbb{C}^n/\text{lattice} \hookrightarrow \mathbb{P}^N$$

projective abelian varieties.

Proof. Let us first show necessity. We have

$$(f, a) = \sum_k f_k \bar{g}_k = \sum_{k,j} \alpha_{k,j} a_j \bar{a}_k.$$

Then

$$\Im(f,a)=\sum_{j,k}(\Im\alpha_{k,j})(b_jb_k+c_jc_k).$$

For sufficiency, we use the projection operator for the real Hilbert space. First look at the special case. The spherical coordinate is essentially equal to $n=1$. So first assume that $a\in\theta\mathbb{R}^n$. We want to find $A=B+iC$ such that $Ba=g$ and $Ca=h$, where we write $f=g+ih$. Let us first define

$$h_1=h-\frac{(h,a)a}{2(a,a)}.$$

Then we are going to define

$$Cx=\frac{(h,a)x}{2(a,a)}+\frac{(x,h_1)h_1}{(a,h_1)}.$$

Then you can check that $Ca=h$ and C is symmetric and positive definite. Then $Ba=g$ can be done because the only requirement is that B is symmetric.

In the general case, $\alpha\notin\theta\mathbb{R}^n$, we would like to make

$$\alpha=i\frac{\Im(f,a)}{(a,a)}I+\beta.$$

If we let

$$f_1=f-ai\frac{\Im(f,a)}{(a,a)},$$

then we have $\Im(f_1,a)=0$ and we want to now achieve $\beta a=f_1$ and β symmetric and real. Consider all $z\in\mathbb{C}^n$ such that $z=\beta a$ for β symmetric real. This set is a $\mathbb{R}$-linear set, so we can write $\{z=\beta a\}$ is the intersection of $\Im(z,g)=0$ for a collection of g. Then $\Im(\xi,g)(a,\xi)=0$ for all $\xi\in\mathbb{R}^n$, and this shows that g should be a real multiple of a. $\square$

So we are going to repeat the same argument for nonsolvability, that Nirenberg–Trèves took from Hörmander. We claim by Taylor expansion,

$$C_{2m-1}(x,\xi)=i\sum_{j=1}^{m}(P_m^{(j)}(x,\xi)\bar{P}_{m,j}(x,\xi)-P_{m,j}(x,\xi)\bar{P}_m^{(j)}(x,\xi)).$$

Here,

$$P_{m,j}(x,\xi)=\partial_jP_m(x,\xi),\quad P_m^{(j)}=\partial_\xi P_m(x,\xi).$$

This is because if we write

$$P(D)(au)=\sum_\alpha(D^\alpha a)Q_\alpha(D)u$$

then $P(\xi+\eta)=\sum_\alpha\xi^\alpha Q_\alpha(\eta)$. So we can write $Q_\alpha(\eta)=\frac{1}{\alpha!}P^{(\alpha)}(\eta)$. We now compute

$$\bar{P}(x,D)P(x,D)u=\sum_\beta\bar{P}(x,D)(a^\beta D^\beta u)=\sum_{\alpha,\beta}\frac{D^\alpha a^\beta}{\alpha!}\bar{P}^{(\alpha)}D^\beta u.$$

and then you can do the other product as well.

Assume there exists $\xi \in \mathbb{R}^n - \{0\}$ such that $P_m(0,\xi) = 0$ but $C_{2m-1}(0,\xi) < 0$. Then

$$\Im\left(\sum_j P_{m,j}(x,\xi)\bar{P}_m^{(j)}(x,\xi)\right) < 0$$

at $x = 0$. Applying the lemma, we find a symmetric matrix $(\alpha_{j,k})$ such that $\Im(\alpha_{jk}) > 0$ and

$$P_{m,j}(0,\xi) + \sum_{k=1}^{n} \bar{P}_m^{(k)}(0,\xi)\alpha_{j,k} = 0.$$

This will take the place of $\vec{b} \cdot \vec{\xi}$ having a sign change.

12 March 1, 2018

Hörmander's idea was to make the estimate fail, using Cauchy–Kowalevski. But in the higher order case, the expressions easily become complicated, and Hörmander had this linear algebra lemma.

Lemma 12.1. *Let $\vec{a} = (a_1, \dots, a_n) \in \mathbb{C}^n$ and $\vec{f} \in \mathbb{R}^n$. Then there exists an $A = (\alpha_{jk})$ that is symmetric and $\Im > 0$ such that $A\vec{\alpha} = \vec{f}$, if and only if $\Im(\vec{f}, \vec{a}) > 0$.*

The symmetric condition will be needed in the second derivative, and also $\Im > 0$ will be used in the exponential quadratic decay. Let me be precise. Let P be a differential operator of order m, and it takes the form of

$$\Big(\frac{\partial}{\partial x_1}\Big)^m + \cdots$$

then we can treat $x_2, \dots, x_n$ as parameters. So if we find a function w such that the differential equation takes this form in a coordinate system with $y_1 = w$, and also $P(x, \operatorname{grad} w) \neq 0$, then we can do something like Cauchy–Kowalevski. Here the set $P(x, \operatorname{grad} w) = 0$ is called the **characteristic set**. Here is the geometric picture. Consider a constant coefficient differential operator, and consider $P_m(\xi) = 0$. This gives a cone in the ξ-space, and the $w = \text{const}$ hypersurfaces intersecting the cones transversely is the good condition.

12.1 Applying Hörmander's lemma

Now let us start showing the necessary condition for solvability. Let $P(x, D)$ be of order m, and consider

$$C(x, D) = \overline{P}(x, D)P(x, D) - P(x, D)\overline{P}(x, D)$$

be of order $2m - 1$. Then we have shown

$$C_{2m-1}(x, \xi) = i \sum (P_m^{(j)}(x, \xi)\overline{P}_{m,j} - P_{m,j}(x, \xi)\overline{P}_m^{(j)}(x, \xi)).$$

If the equation $P(x, D)u = f$ is solvable at 0, we want to show that $C_{2m-1}(0, \xi) = 0$ on $P_m(0, \xi) = 0$. We were showing that if there exists $\xi \neq 0$ such that $P_m(0, \xi) = 0$ and $C_{2m-1}(0, \xi) < 0$, then the estimate fails. That is, given k, N we want to find a sequence $f_\tau, v_\tau \in C_0^\infty(\Omega)$ such that

$$\frac{1}{\tau} \int f_\tau v_\tau$$

is bounded from below but $\sup_{|\alpha| \le k} |D^\alpha f|$ and $\sup_{|\beta| \le l} |D^\beta L^* v|$ is bounded.

We want to find a v such that $L^* v = 0$ as a jet, i.e., up to some high order. Now we are going to choose

$$v = \exp(iw), \quad w = \langle x, \xi \rangle + \frac{1}{2} \sum_{j,k=1}^n \alpha_{jk} x_j x_k.$$

Here, α_{jk} should satisfy the condition that α_{jk} is symmetric, and $\Im(\alpha_{jk}) > 0$. We also want $L^* w = 0$ modulo $|x|^q$. Another thing we want is $P_m(x, \operatorname{grad} w) = 0$, because we are going to take $L^*(e^{i\tau w(\frac{x}{\tau})}\varphi)$ later.

So how do we get this? Hörmander's idea is to choose α_{jk} smartly using the linear algebra lemma. We do a linear change of coordinates so that the coefficient of $\frac{\partial^m}{\partial x_n^m}$ is nonzero. Then we use the initial data $W = \langle x, \xi \rangle + \frac{1}{2} \sum_{j,k=1}^{n-1} \alpha_{jk} x_j x_k$. If we differentiate the equation $P_m(x, \operatorname{grad} w) = 0$ with respect to x_j, then we get

$$P_{m,j}(x, \operatorname{grad} w) + \sum_{k=1}^{n} P^{(k)}(x, \operatorname{grad} w) \frac{\partial^2 w}{\partial x_j \partial x_k} = 0.$$

Now note that at $x = 0$, what we have is precisely what we are trying to do. Because $\frac{\partial^2 w}{\partial x_j \partial x_k} = \alpha_{jk}$, this can be chosen precisely when

$$C_{2m-1}(0, \xi) = \Im(P_{m,\bullet}, P_m^{(\bullet)}) < 0.$$

But we still need to show some estimates. We started with the initial data

$$W = \langle x, \xi \rangle = \frac{1}{2} \sum_{j,k=1}^{n-1} x_j x_k + \text{higher order}$$

at $x_n = 0$ and used Cauchy–Kowalevski, and got this w. We had $\frac{\partial^2 w}{\partial x_j \partial x_j} = \alpha_{jk}$ for $1 \leq j, k \leq n-1$ at 0, and we also had the other term

$$P_{m,j}(x, \operatorname{grad} w) + \sum_{k=1}^{n} P^{(k)}(x, \operatorname{grad} w) \frac{\partial^2 w}{\partial x_j \partial x_k} = 0.$$

This actually shows that $\alpha_{jk} = \frac{\partial^2 w}{\partial x_j \partial x_k}$ at the origin, even for $1 \leq j, k \leq n$.

13 March 6, 2018

We had to solve a $L^* v_\tau$ approximately 0 to order $2r$. Here, we use Fourier series. We write

$$L^*(\psi e^{i\tau w(x)}) = \sum_{\nu \le m} c_\nu(x) \tau^\nu e^{i\tau w(x)}.$$

and we want the leading coefficients to be nonzero.

13.1 Construction of counterexamples to the inequality

So we start outed with

$$w = ix\xi + \sum_{i,j=1}^n a_{ij} x_i x_j$$

and constructed the solution w for $P_m(x, \operatorname{grad} w) = 0$. Then we are going to let

$$v_\tau = \tau^{n+1+k} e^{i\tau w} \Big(\varphi_0(x) + \frac{\varphi_1(x)}{\tau} + \cdots + \frac{\varphi_r}{\tau^r} \Big).$$

Let us apply L^* to this. If we have $L^* = \bar{P}(x, -D)$, then the top degree is

$$\bar{P}(x, -D)(\varphi_0 e^{i\tau w}) = A\varphi_0 \tau^m e^{i\tau w} + \bigg(\bigg(\sum_{j=1}^n A_j D_j \bigg) \varphi_0 + B\varphi_0 \bigg) \tau^{m-1} e^{i\tau w} + \cdots .$$

Here, we can easily see $A = \bar{P}_m(x, -\partial_x w)$ and $A_j = -\bar{P}_m^{(j)}(x, -\partial_x w)$. Now we do the same thing for φ_1 and so on. Then we can carefully compute

$$\begin{aligned} e^{-i\tau w} \bar{P}(x, -D) \Big(\varphi_0 + \frac{\varphi_1}{\tau} + \frac{\varphi_2}{\tau^2} + \cdots \Big) \\ &= (A\varphi_0)\tau^m + ((\sum A_j D_j)\varphi_0 + B\varphi_0)\tau^{m-1} + (B_{0,2}\varphi_0)\tau^{m-2} + (B_{0,3}\varphi_0)\tau^{m-3} + \cdots \\ &\qquad + (A\varphi_1)\tau^{m-1} + ((\sum A_j D_j)\varphi_1 + B\varphi_1)\tau^{m-2} + (B_{0,2}\varphi_1)\tau^{m-3} + \cdots \\ &\qquad\qquad \cdots . \end{aligned}$$

We want this to be $O(|x|^{2r})$. But we note that we have already set w so that $A = O(|x|^{2r})$. So we only need to find φ_i so that

$$\begin{cases} (\sum A_j D_j)\varphi_0 + B\varphi_0 = O(|x|^{2r}), \\ (\sum A_j D_j)\varphi_1 + B\varphi_1 + B_{0,2}\varphi_0 = O(|x|^{2r}), \\ (\sum A_j D_j)\varphi_2 + B\varphi_2 + B_{0,2}\varphi_1 + B_{0,3}\varphi_0 = O(|x|^{2r}), \\ \cdots . \end{cases}$$

This can be solved inductively by Cauchy–Kowalevski, because one of the coefficients $A_j = \bar{P}_m^{(j)}(0, \xi)$ has to be nonzero.

So for this choice of v_τ, we compute

$$L^* v_\tau = \tau^{r-N} e^{i\tau w} \sum_{\mu=0}^{m+r-1} \frac{1}{\mu} a_\mu$$

where $a_\mu(x) = O(|x|^{2(r-\mu)})$. We also have $\Im w \geq a|x|^2$ for some $a.0$, so we get that $|e^{i\tau w}| \leq e^{-\tau a|x|^2}$. So as $\tau \to \infty$, we get

$$\sup_{|\beta|\leq n} |D^\beta L^* v_\tau| \to 0.$$

Now choose $F \in C_0^\infty(\Omega)$ such that $\hat{F}(-\xi) \neq 0$. We set

$$f_\tau(x) = \frac{1}{\tau^k} F(\tau x).$$

Then we sill have

$$\sup_{|\alpha|\leq k} |D^\alpha f_\tau(x)|$$

is bounded. On the other other hand,

$$\frac{1}{\tau} \int f_\tau v_\tau = \int F(x) e^{i\tau w(\frac{x}{\tau})} \sum_{\nu=0}^{r-1} \frac{\varphi_\nu(\frac{x}{\tau})}{\tau^\nu} \to \hat{F}(-\xi)\varphi_0(0).$$

This does not go to 0.

The moral of this entire story is that the characteristic is really important. The contribution of Nirenberg and Trèves was that you can use the sign change to get an additional sign, instead of setting it.

Now we are going to talk about hypoellipticity. People like Kolmogorov showed that a differential equation like

$$\frac{\partial^2 u}{\partial x^2} + x\frac{\partial u}{\partial t} - \frac{\partial u}{\partial t} = f.$$

Hörmander observed that if you write $X_1 = \frac{\partial}{\partial x}$ and $X_0 = x\frac{\partial}{\partial y} - \frac{\partial}{\partial t}$, then we can write it as $(X_1^2 + X_0)u = f$. Then

$$[X_1, X_0] = \frac{\partial}{\partial t}$$

and so $X_0, X_1, [X_1, X_0]$ generate the tangent space at 0.

Theorem 13.1. *If $X_0, X_1, \ldots, X_r$ are smooth vector fields, and if iterated brackets generate all vectors, then*

$$\sum_{j=1}^{r} X_j^2 + X_0$$

is solvable at 0 *and is hypoelliptic.*

14 March 8, 2018

The zeros of the principal symbol is really important. This is always going to be some cone, because $P_m(x,\xi)$ is a homogeneous polynomial. If it is a point, this is operator is called elliptic and everything is very nice. The next case seems to be when it has zeros of multiplicity 1, but we have seen that this is the worst case. Hörmander observed that if we have a linear one term like

$$\frac{\partial}{\partial t} - \Delta,$$

the heat kernel, this is nice. Here, note that the zero set is a line, but with multiplicity 2. This is the second next case.

14.1 Hörmander's hypoellipticity

In general, consider a differential operator

$$\sum_{i,j=1}^{n} a_{ij}\frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^{n} a_i \frac{\partial u}{\partial y_i} + au + \sum_{i=1}^{n} b_i \frac{\partial u}{\partial x_i} - \frac{\partial u}{\partial t} = f,$$

where $(a_{ij}) > 0$, on $\mathbb{R}^{2n+1}$. If this can be written as

$$\sum_{j=1}^{r} X_j^2 + X_0$$

for some vector fields X_j, and $[[[X_j, X_k], X_l], \dots, X_m]$ generate the tangent plane, then the equation is solvable. Moreover, we get some gain: if $Lu = f$ and $f \in C_0^k(\Omega)$ then $u \in C^{k+\epsilon}(\Omega)$ where ϵ^{-1} is roughly the number of brackets we need to generate the plane.

The rough idea is to take a detour. Basically we want to show some a priori estimate, and we want to compare values between points. But the problem is that sometimes we don't have vectors that will send me somewhere directly. So we take a detour by looking at the Lie bracket.

14.2 Preparation for Riesz representation

Let $X_0, \dots, X_n$ be smooth real vector fields, and let c be a complex-valued smooth functions. Consider the operator

$$P = \sum_{j=1}^{r} X_j^2 + X_0 + c.$$

We want to solve $Pu = f$ for $f \in C_0^\infty(\Omega)$, and we want to write $\int v(\overline{Pv})$ in terms of $\sum_j |X_j v|^2$. We have $X_j^* = -X_j + a_j$ for some smooth a_j. Then we can write

$$\begin{aligned}
-\Re\int v\overline{Pv} &= -\Re\int\sum_{j=1}^{v} v\overline{X_j^2 v} - \Re\int v\overline{X_0 v} - \int 9\Re c)|v|^2 \\
&= -\Re\int\sum_{j=1}^{v}(X_j^* v)(\overline{X_j v}) - \Re\int X_0(\tfrac{1}{2}|v|^2) - \int(\Re c)|v|^2 \\
&= -\Re\int\sum_{j=1}^{r}(-X_j v + a_j v)X_j\overline{v} - \Re\int a_0\frac{|v|^2}{2} - \int(\Re c)|v|^2.
\end{aligned}$$

So we can write

$$-\Re\int vP\bar{v} = \sum_{j=1}^{r}\int|X_j v|^2 + \int d|v|^2$$

where $d = \frac{1}{2}\sum(X_j a_j - a_j^2) - \frac{1}{2}a_0 - \Re(c)$. Then we have

$$\sum_{j=1}^{r}\|X_j v\|^2 + \|v\|^2 \le C\|v\|^2 - \Re\int vP\bar{v}.$$

The left hand side looks like a L_1^2-norm. So we define

$$|||v|||^2 = \sum_{j=1}^{r}\|X_j v\|^2 + \|v\|^2.$$

We use the Riesz representation theorem. But we are going to be in the situation of $\|\varphi\| \le C\|L^*\varphi\|_m$ again. Then what we are going to get is something in L_{-1}^2.

Definition 14.1. $|||-|||'$ is dual to $|||-|||$ with respect to the $(-,-)_{L^2}$ inner product. That is,

$$|||f|||' = \sup_{v\in C_0^\infty(\Omega)}\frac{|\int fv|}{|||v|||}.$$

If we look at the real part, we have by definition

$$-\Re\int v\overline{Pv} \le |||v||| |||Pv|||.$$

But we need to estimate $X_0 v$ in terms of Pv, but only have

$$X_0 v = Pv - \sum_{j=1}^{n} X_j^2 v - c.$$

We know how to bound $X_j v$, but we don't know anything about $X_j^2 v$. So we look at the weaker norm $|||X_j v|||'$ of $X_j v$. Then we have

$$|||X_j f|||' = \sup\frac{\int\bar{g}(X_j f)}{|||g|||} = \sup\frac{\int(-X_j\bar{g} + a_j\bar{g})f}{|||g|||} \le C\|f\|\frac{(\|X_j g\| + \|g\|)}{|||g|||} \le C'\|f\|.$$

It follows that

$$\begin{aligned}\|\!|X_0 v|\!\|' &\le \|\!|Pv|\!\|' + \sum_j \|\!|X_j^2 v|\!\|' + C\|\!|v|\!\|' \\ &\le \|\!|Pv|\!\|' + C\sum_j \|X_j v\| + C\|v\| \\ &\le C'(\|v\| + \|\!|Pv|\!\|').\end{aligned}$$

On the other hand, recall that we had

$$\|\!|v|\!\|^2 \le C\|v\|^2 + C'\|\!|v|\!\|\|\!|Pv|\!\|'$$

and then by the small-constant large-constant,

$$\|\!|v|\!\|^2 \le C(\|v\|^2 + \|\!|Pv|\!\|'^2).$$

So if we add up, we get

$$\|\!|v|\!\|^2 + \|\!|X_0 v|\!\|'^2 \le C(\|v\|^2 + \|\!|Pv|\!\|'^2).$$

We now want to use the idea of taking the detour (or iterated Lie bracket). We want an estimate like

$$\|v\|_{(\epsilon)} \le C(\|\!|v|\!\| + \|\!|X_0 v|\!\|')$$

for all $v \in C_0^\infty(\Omega)$, for some $\epsilon > 0$. This $\|-\|_{(\epsilon)}$ is going to be the L_2-norm with all directional derivatives of order $\le \epsilon$ in some sense. If we have this, we would get

$$\|v\|_{(\epsilon)} \le C(\|v\| + \|\!|Pv|\!\|').$$

But then the diameter being less than η will give us some estimate like $\|v\| \le C\eta^\epsilon \|v\|_{(\epsilon)}$.

Let me introduce some notation. For a vector field X, we denote by e^{tX} the vector flow generated by it. Then for a function u, we denote

$$(e^{tX}u)(x) = u(f(x,t))$$

the function u evaluated at the flow at time t. We then define

$$|u|_{X,s}^\epsilon = \sup_{0 \le |t| < \epsilon} \frac{\|e^{tX}u - u\|_{L^2}}{|t|^s}$$

for ϵ sufficiently small. This is some kind of a Hölder norm for elements in L^2.

15 March 20, 2018

We are looking at the Hörmander's hypoellipticity. His important contribution was that the operator can be written as $P = \sum_{j=1}^n X_j^2 + X_0$, and the iterated brackets generate the tangent space. In fact, there is going to be a gain of derivative in all directions. The idea is to use the Sobolev norm with only specific directions of differentiation. Using this norm, we are going to use the Lie bracket to measure the failure of the exponential law, and then we will do some smoothing.

We were trying to use Riesz representation. We are going to use some duality with respect to the inner product without derivatives. So we defined

$$|\!|\!|v|\!|\!|^2 = \sum_{j=1}^r \|X_j v\|^2 + \|v\|^2$$

and $|\!|\!|v|\!|\!|'$ as the dual norm, so that we always have

$$|(u,v)| \le |\!|\!|u|\!|\!| |\!|\!|v|\!|\!|'.$$

For $P = \sum_j X_j^2 + X_0 + c$ where X_0, X_j are real and c might be complex, we showed that

$$-\Re \int v\overline{Pv} = \sum_{j=1}^r \int |X_j v|^2 + \int d|v|^2, \quad d = \frac{1}{2}\sum_{j=1}^r ((X_j a_j) - a_j^2) - \frac{1}{2}a_0 - \Re c$$

where $X_j^* = -X_j + a_0$. From this we got

$$|\!|\!|v|\!|\!| = \sum_{j=1}^r \|X_j v\|^2 + \|v\|^2 \le C\|v\|^2 - \Re \int v\overline{Pv}$$

for all $v \in C_0^\infty(\Omega)$. But

$$|\!|\!|X_j^2 v|\!|\!|' \le C\|X_j v\| \le |\!|\!|v|\!|\!|$$

and this implies by small-constant large-constant,

$$|\!|\!|X_j^2 v|\!|\!|' \le C(\|v\| + |\!|\!|Pv|\!|\!|').$$

Because $X_0 v = Pv - \sum_j X_j^2 v - c$, we get

$$|\!|\!|X_0 v|\!|\!|' \le C(\|v\| + |\!|\!|Pv|\!|\!|').$$

So the conclusion is

$$|\!|\!|v|\!|\!|^2 + |\!|\!|X_0 v|\!|\!|'^2 \le C(\|v\|^2 + |\!|\!|Pv|\!|\!|'^2).$$

15.1 Using iterated Lie brackets

Now we have control on the L^2-norm on the X_j-derivative, and the weak X_0-derivative norm. We want to generate from this the L^2-norm in all directions. Here, we are going to interpret the weak X_0-derivative norm as some kind of $L^2_{1/2}$ in the X_0-direction.

Let us write

$$X_I = [X_{\nu_1}, [X_{\nu_2}, [\dots, X_{\nu_k}]]] = \operatorname{ad} X_{\nu_1} \cdots \operatorname{ad} X_{\nu_{k-1}} \cdot X_{\nu_k}$$

for $I = (\nu_1, \dots, \nu_k)$. Let us define $s(\nu) = 1$ if $1 \le \nu \le r$ and $s(\nu) = \frac{1}{2}$ if $\nu = 0$, and define

$$\frac{1}{s(I)} = \frac{1}{s(\nu_1)} + \cdots + \frac{1}{s(\nu_k)}.$$

If X is a real vector field and u a smooth function, we are going to write

$$e^{tX} u = u(f(x,t))$$

where f is the flow generated by X. For instance, $Xu = \lim_{t\to 0} \frac{e^{tX}u - u}{t}$.

When we try to compute $e^{r_2} = e^{-y}e^{-x}e^{x+y}$, we get

$$r_2 = -\frac{1}{2}[x,y] + \text{higher order terms}.$$

So let $z_2 = -\frac{1}{2}[x,y]$, then we can look at $e^{r_3} = e^{-z_2}e^{r_2}$ and take away the degree 3 elements and $e^{r_4} = e^{-z_3}e^{r_3}$ and so on. The result is that

$$e^{-y}e^{-x}e^{x+y} = e^{z_2}e^{z_3}\cdots e^{z_k}e^{r_{k+1}}$$

where z_k is the sum of commutators of degree k.

Lemma 15.1. *For any $t > 0$, $0 < \sigma \le 1$, and $N \ge 2$,*

$$\|e^{t(X+Y)}u - u\| \le C\left(\|e^{tX}u - u\| + \|e^{tY}u - u\| + \sum_{j=2}^{N-1} \|e^{t^j z_j}u - u\| + t^{\sigma N}|u|_\sigma\right),$$

where $u \in C_0^\infty(\Omega)$ and

$$|u|_\sigma = \sup_{|h|<\epsilon} \frac{\|u(x+h) - u(x)\|_{L^2}}{|h|^\sigma}.$$

(Different ϵ yield different norms, but they are equivalent norms.)

There is a different trick involving coordinate transformations.

Lemma 15.2. *Let $x \mapsto g(x,t)$ be a family of diffeomorphisms, and assume that $g(x,t) - x = O(t^N)$ as $t \to 0$, where $N > 0$. Then*

$$\int |u(g(x,t)) - u(x)|^2 dx \le C|t|^{2Ns}|u|_s^2$$

for $0 < s \le 1$.

This is trivial if $g(x,t)$ are curves in the standard coordinate directions.

Proof. We want to compare $u(g(x,t)) - u(x+h)$ with $u(y+w) - u(y)$. We make the change of coordinates $y = x + h$ and $y + w = g(x,t)$. The determinant of the Jacobian is close to 1, so we may ignore this. Then

$$\int |u(y+w) - u(y)|^2 \leq C|w|^{2s}|u|_s^2.$$

$\square$

Using a similar argument, you can prove something like this. If X is replaced by φX for $\varphi \in C_0^\infty(\Omega)$, we have

$$|u|_{\varphi X,s} \leq C_\varphi |u|_{X,s}.$$

16 March 22, 2018

What we are doing is some sort of microlocal analysis. We not only control the number of differentials but also the direction. Hörmander used Hölder estimates here for fractional order. We have (real) vector fields $X_1, \dots, X_r$ and X_0, and we're looking at

$$P = \sum_{j=1}^{r} X_j^n + X_0 + c.$$

Here X_0 is something like the time variable. Then Hörmander's theorem is that this is hypoelliptic. The point of microlocal analysis is to write down the correct norm, which is

$$|||u|||^2 = \|u\|^+ \sum_{j=1}^{r} \|X_j u\|^2$$

in this case. After integration by parts, we got the a priori estimate

$$|||v|||^2 + |||X_0 v|||'^2 \leq C(\|v\|^2 + |||Pv|||'^2)$$

for all $v \in C_0^\infty(\Omega)$. This $|||X_0 v|||'$ is roughly $\|X_0^{1/2} v\|$ because X_0 is roughly two X_j and we're taking away one X_j.

Now the key is to go back to the Sobolev ϵ-norm in all directions, because we want to show that it is smooth in all directions. So we need something like

$$\|v\|_{(\epsilon)}^2 \leq C(\|v\|^2 + |||Pv|||'^2).$$

Here, we need this $\epsilon > 0$ to jack up differentiability of the solution, by taking care of commutators. So suppose we have something like $Pu = f$. Then

$$P(D^\alpha u) = D^\alpha f + [P, D^\alpha]u.$$

For instance, take L be of first order and $Lu = f$. Then

$$L\frac{du}{dt} = \frac{df}{dt} + \Big[L, \frac{d}{dt}\Big]u,$$

and a priori estimates will give

$$\frac{1}{\delta^\epsilon}\|u\|_{L_1^2} \leq \|u\|_{L_{1+\epsilon}^2} \leq C_1 \Big\|\frac{df}{dt}\Big\|_{L^2} + C\|u\|_{L_1^2}$$

for diameter $< \delta$. This shows that for δ sufficiently small, we get the upper bound on $\|u\|_{L_1^2}$.

16.1 Bounds on the Hölder norm

Now our goal is to get the estimate

$$\|v\|_{(\epsilon)}^2 \leq C(|||v|||^2 + |||X_0 v|||'^2).$$

Roughly, we are going to use

$$(\operatorname{ad} X_{\nu_1} \operatorname{ad} X_{\nu_2} \cdots \operatorname{ad} X_{\nu_{k-1}} X_{\nu_k})^{\frac{1}{m(\nu_1,\dots,\nu_k)}}.$$

Then these directions will generate the whole space, so $\epsilon = \min \frac{1}{m}$ will do the job.

We want to approximate the Hölder norm. This is defined as

$$\|v(x+t) - v(x)\|_{L^2} \leq |t|^\epsilon \|v\|_{(\epsilon)}.$$

There was a lemma that allowed us to do this for coordinate transformations in general.

Lemma 16.1 (4.2)**.** *Assume $g(x,t) - x = O(t^N)$ where g is smooth. Then*

$$\int |u(g(x,t)) - u(x)|^2 dx \leq C|t|^{2Ns}|u|_s^2.$$

Lemma 16.2 (4.1, rescaling)**.** *If $\varphi \in C^\infty(\Omega, \mathbb{R})$, then*

$$|u|_{\varphi X,x} \leq C|u|_{X,s},$$

where $|u|_{X,s}$ is the Hölder norm along X,

$$|u|_{X,s} = \sup_{|t|<\epsilon} \frac{\|e^{tX}u - u\|_{L^2}}{|t|^s}.$$

Now for $I = (\nu_1, \dots, \nu_k)$, let us denote $X_I = \operatorname{ad} X_{\nu_1} \cdots \operatorname{ad} X_{\nu_{k-1}} X_{\nu_k}$. Then we want to add X_I together to get all directions. We can write

$$e^{(X_I + X_J)} = e^{X_I} e^{X_J} e^{Z_2} e^{Z_3} \cdots e^{Z_{N-1}} e^{\gamma_N}.$$

Then using the fact that

$$S_1 \cdots S_k u - u = \sum_{j=1}^k S_1 \cdots S_{j-1}(S_j u - u),$$

we can show that

$$\|e^{t(X+Y)}u - u\| \leq C_{k,N}\left(\|e^{tX}u - u\| + \|e^{tY}u - u\| + \sum_{j=2}^{N-1} \|e^{t^j Z_t}u - u\| + t^{\sigma N}|u|_\sigma\right)$$

for $0 < \sigma \leq 1$ and $N \geq 2$ and $u \in C_0^\infty(K)$ where K is compact in Ω.

So now let us define $s_0 = \frac{1}{2}$, $s_j = 1$, and

$$s(I) = \frac{1}{\frac{1}{s(\nu_1)} + \cdots + \frac{1}{s(\nu_k)}}.$$

From above, we have

$$|u|_{X,s} \leq C\left(\sum_{j=0}^{r} |u|_{X_j,s} + \|u\|\right)$$

for $X \in T^s$, where $T^s(\Omega)$ is the subbundle of $T(\Omega)$ generated by X_I with $s(I) \geq s$.

We still need a smoothing procedure. We look at $L^2_{\alpha+\epsilon}$ and consider pseudo-differential operator with symbol $(1+\delta^2|\xi|^2)^{-1}$,

$$v \mapsto \int (1+\delta^2|\xi|^2)^{-1}\hat{v}(\xi)e^{2\pi i x\cdot\xi}d\xi.$$

Note that this would be an actual differential operator if $p(x,\xi)$ is a polynomial instead of $(1+\delta^2|\xi|^2)^{-1}$. We would also have the use some kind of a Fredrich's lemma, because we want estimates to pass on. This will allow us to compare $\|X_0 u\|'$ and $|u|_{X_0,\frac{1}{2}}$.

17 March 27, 2018

We defined the L^2 Hölder norm as

$$|u|_{X,s}^{\epsilon} = \sup_{0<|t|<\epsilon} \frac{\|e^{tX}u - u\|_{L^2}}{|t|^s}$$

for $u \in C_0^\infty(K)$. Then we had some weight

$$m(I) = \frac{1}{s(I)} = \sum_j \frac{1}{s_{\nu_j}}.$$

We want to estimate on $T^s(\Omega)$, which is generated by all X_I (linear coefficients in $C^\infty(\Omega)$, with $s(I) \geq s$). Here, the key estimate is that

$$|u|_{X,s} \leq C\left(\sum_{j=0}^{r} |u|_{X_j,s} + \|u\|\right)$$

for all $X \in T^s(\Omega)$. This is just a matter if iterated application of vector fields. Here, to get iterated brackets, we had to have

$$\|e^{t^{m(I)}X_I}u - u\| \leq C_1 t \sum_{j=0}^{r} |u|_{X_j,s_j} + c_2 t |u|_\sigma.$$

17.1 Handling different directions

Now we have all these different Hölder norms in different directions. For $t > 0$, Consider $f(t) = \|e^{tX_0}v - v\|$. Then we have

$$\frac{d}{dt}(f(t)^2) = 2(e^{tX_0}X_0 v, e^{tX_0}v - v).$$

Here, if $\|\!|-|\!\|$ turns out to be e^{tX_0}-invariant, then we would have an estimate

$$\frac{d}{dt}(f(t)^2) \leq C\|\!|X_0 v|\!\|' \cdot 2\|\!|v|\!\|.$$

Then we will be able to conclude that

$$f(t) \leq t^{\frac{1}{2}} C(\|\!|v|\!\| + \|\!|X_0 v|\!\|')$$

and so $|v|_{X_0,\frac{1}{2}} \leq C(\|\!|v|\!\| + \|\!|X_0 v|\!\|')$. Then we are done.

But it is clear that $\|\!|-|\!\|$ won't necessarily be invariant under that flow. So introduce another norm that is invariant. We do this by taking the average. Take $\sigma > 0$ such that $T^s(\Omega) = T(\Omega)$ for some $s > \sigma$. Consider the set $\mathcal{I}$ of all I with $\sigma m(I) \leq 1$ and $|I| < m(I) < 2|I|$. (So not all are X_0 but there exists an X_0.) Define

$$M(u) = \|\!|u|\!\| + \|\!|X_0 u|\!\|' + \sum_{I\in\mathcal{I}} |u|_{X_I,x(I)} + |u|_\sigma.$$

Then by the same argument, we get that

$$|u|_{X_0,\frac{1}{2}} \leq CM(u).$$

To see this, we consider

$$S^t u = \prod_{I \in \mathcal{I}} e^{t^{m(I)} X_I} \Phi_{t^{1/\sigma}} u.$$

Here, we are sort of using Freidrich's argument. The commutator and smoothing out and the differential operator is bounded by u. Using this, we can show that $\|S_t u - u\| \leq CtM(u)$.

17.2 Towards the implicit function theorem

There is Nash's original 1956 paper in *Ann. of Math.*, and there are 1969 notes by J. Schwartz called *nonlinear functional analysis.*

Theorem 17.1. *Let M be a compact Riemannian C^∞-manifold of dimension n, there exists a smooth embedding $M \hookrightarrow \mathbb{R}^N$ such that the metric is the pullback of the standard metric.*

The idea is to use something like Newton's method. Normally, the implicit function works in the following way. You move a little bit, and there still is a zero. But we can also use Newton's method, by iterating the approximations to get the zero.

18 March 29, 2018

There is one part in Hörmander's hypoellipticity that I have not explained. Because we are doing integration by parts, we actually need smoothness. So we need a smoothing operator. But we will not talk about this.

We now want to talk about the isometric embedding theorem of Nash. We are going to just do a special situation when M^n is a torus. This is in some sense an implicit function theorem. In the usual case, we have something like $F(x, y) = 0$ and $F(0, 0) = 0$ and $\frac{\partial F}{\partial y}(0, 0) \neq 0$. Then the conclusion is that near 0, there exists $y = y(x)$ such that $F(x, y(x)) = 0$. The usual proof is from the intermediate value theorem if you move x a bit, $F(x, -)$ should have a zero and then you do something.

Given a metric $g = \sum_{jk} g_{ij} dx_j dx_k$ that is C^∞ on M, we look at the space $\mathcal{G}$ of all smooth metrics on M. Also, let $\mathcal{F}$ be the set of all smooth embeddings $f : M \to \mathbb{R}^N$. Then you can pull back the metric, and solve

$$f^*(g_{\text{std}}) = g.$$

So we can define, for $f \in \mathcal{F}$ and $g \in \mathcal{G}$,

$$\Phi(f, g) = g - f^*(g_{\text{std}}).$$

Then we are solving $\Phi(f, g) = 0$, given a fixed f.

We can always look at Newton's method. We first approximate the function by 1st order and solve for the linearized equation, and iterate this. For the 2-variable case, we are approximating F by

$$F(x_0, 0) + \frac{\partial}{\partial y} F(x_0, 0) y_1 = 0.$$

This can be done in the context of Banach spaces. But here the problem is that $\mathcal{F}$ and $\mathcal{G}$ are Fréchet spaces, and you lose the differentiability at each step. So the key idea is to replace derivation by the difference quotient.

18.1 Newton's method in Banach spaces

Here, we would be looking at the differential. But the existence of the right inverse is going to be needed.

Theorem 18.1. *Let B be a Banach space, and let $f : \Omega \to B$, where Ω is the open unit ball in B. Suppose $f \in C^2$ in the sense of Fréchet. This means that*

$$\lim_{h \to 0} \frac{\|f(x + h) - f(x) - f'(x)h\|}{\|h\|} = 0.$$

Further, assume that the second derivatives are bounded by some $M \geq 2$, and for $u \in \Omega$ there exists a $L(u) \in \mathcal{B}(B)$ with bounded norm M, which is the right inverse of $(df)(u)$. If $|f(0)| \leq M^{-5}$, then there exists a $u \in \Omega$ such that $f(u) = 0$.

Proof. Inductively, we are going to write down

$$u_0 = 0, \quad u_{n+1} = u_n - L(u_n)f(u_n).$$

Let $\kappa = \frac{3}{2}$ and $\beta > 0$. Then we claim that $|u_n - u_{n-1}| \le e^{\kappa\kappa^n}$. We have

$$\begin{aligned} |u_{n+1} - u_n| &= |L(u_n)f(u_n)| \le M|f(u_n)| \\ &\le M|f(u_{n-1}) - df(u_{n-1})L(u_{n-1})f(u_{n-1})| + M^2|u_n - u_{n-1}|^2. \end{aligned}$$

Then we can choose β so that M can be absorbed. □

Now we want to replace B by $C^m(M)$, which is not a Banach space. So here is the method applied to a space with many Banach norms and smoothing operators. Here, when we smooth, we lose something. Suppose we want to smooth out $\varphi(x)$ with a weight function $h(x)$. Then the convolution is going to look like

$$\varphi_t(x) = \int_y \varphi(y)h_t(x-y)dy.$$

Here, is how you pay the price. First, by triangle inequality, we have $\|\varphi_t\|_0 \le \|\varphi\|_0$. But then we get something like $\|\varphi_t'\|_1 \le \frac{C}{t}\|\varphi\|_0$.

Assume there exists a family of smoothing operators $S(t)$ that is defined for $t \ge 1$, and goes to the identity as $t \to \infty$. Suppose there exist $M \ge 1$ and $m, \alpha \in \mathbb{N}$ such that for any $m - \alpha \le r \le \rho \le m$ the following holds. (Here, $|-|_\rho$ is the $C^\rho(M)$-norm.)

(i) bound in stronger norm: $|S(t)u|_\rho \le Mt^{\rho - r}|u|_r$ for all $u \in C^r(M)$.

(ii) bound in weaker norm: $|(1 - S(t))u|_r \le Mt^{r-\rho}|u|_\rho$ for all $u \in C^\rho(M)$.

(iii) bound in the weaker norm before integration: $|\frac{d}{dt}S(t)u|_r \le Mt^{r-\rho-1}|u|$ for all $u \in C^r(M)$.

(iv) approximation: $\lim_{t\to\infty}|(1 - S(t))u|_r = 0$ for all $u \in C^r(M)$.

Theorem 18.2. *Let Ω be a unit ball in C^m. Assume that $f : \Omega \to C^{m-\alpha}(M)$ is twice Fréchet differentiable with norm $\le M$, such that for all $u \in \Omega$, there exists a right inverse $L(u) \in \mathcal{B}(C^m, C^{m-\alpha})$ so that $df(u)L(u)h = h$ for all $h \in C^{m+\alpha}$, with bound M so that*

$$|L(u)h|_{m-\alpha} \le M|h|_m.$$

Also assume the technical inequality

$$|L(u)f(u)|_{m+9\alpha} \le M(1 + |u|_{m+10\alpha}).$$

If $|f(0)|_{m+9\alpha} \le \frac{1}{2^{40}M^{202}}$ then $f(u) = 0$ for some $u \in \Omega$.

Proof. Let $\kappa = \frac{3}{2}$ and we are going to fix $\beta, \mu, \nu > 0$. Set $t = e^{\beta\kappa^n}$ and define $S_n(e^{\beta\kappa^n})$. (Eventually, S_n is going to be the identity.) Let $u_0 = 0$ and inductively we want to construct u_n, such that $u_n \in \Omega$,

$$|u_n - u_{n-1}|_m \le e^{-\mu\alpha\beta\kappa^n}, \quad u_n \in C^{m+10\alpha}, \quad 1 + |u_n|_{m+10\alpha} \le e^{\nu\alpha\beta\kappa^n}.$$

The claim is that if you define

$$u_{n+1} = u_n - S_n L(u_n) f(u_n),$$

this works. Newton's method work as

$$\begin{aligned} |u_{n+1} - u_n| &= |S_n L(u_n) f(u_n)| \leq M e^{\alpha\beta\kappa^n} |L(u_n) f(u_n)|_{m-\alpha} \leq M^2 e^{\alpha\beta\kappa^n} |f(u_n)|_m \\ &\leq M^2 e^{\alpha\beta\kappa^n} |f(u_{n-1}) - df(u_{n-1}) S_{n-1} L(u_{n-1}) f(u_{n-1})|_m + M^3 e^{\alpha\beta\kappa^n} |u_n - u_{n-1}|_m^2. \end{aligned}$$

You can continue, and this will give you the right bounds. □

Now there are two things that we need to do now. The first thing is to construct the smoothing operators S_t. The other thing moving from local to global.

18.2 Constructing the smoothing operators

Here, the trick is to use a good choice of a weight function. We are going go make this weight function to make polynomials go to zero. Define a weight function a by specifying $\hat{a}$ instead, where this satisfies

$$0 \leq \hat{a} \leq 1, \quad \hat{a} \in C_0^\infty(\mathbb{R}^n), \quad \hat{a} \equiv 1 \text{ near } 0.$$

Then the moment will be

$$\int x^\alpha a(x) dx = D^\alpha \hat{\alpha}(0) = 0.$$

Also, it will automatically satisfy $\int a(x)dx = \hat{a}(0) = 1$.

Using this, we define

$$(S(t)u)(x) = t^n \int_{\mathbb{R}^n} a(t(x-y)) u(y) dy.$$

Here, the point is that if we expand u out by a Taylor series, then the polynomial terms vanish.

19 April 3, 2018

Last time we looked at the method of applying modified Newton's method using the change of norms. Normally, the solution is unique. But for the isometric embedding, there is no uniqueness. This is because imposed only the condition that $\frac{\partial}{\partial y}$ is right invertible. Concretely, we have for M a compact real smooth C^∞,

$$\mathcal{F} = \{f : M \xrightarrow{C^\infty} \mathbb{R}^{N_f}\} \to \begin{Bmatrix} C^\infty \text{ positive symmetric} \\ \text{covariant 2-tensors on } M \end{Bmatrix} = \mathcal{G}.$$

We want to show that this is surjective.

Let's try to understand this in the finite-dimensional case. Suppose you want to solve $y = (y_1, \dots, y_p)$ in terms of x, and it is not unique. Assume that we are solving for the relations $R_j(y_1, \dots, y_p, x) = 0$ for $1 \leq j \leq q$ (and $q < p$). Then the differentiation of $R = (R_j)$ is going to be $(\frac{\partial R_j}{\partial y_k})_{1 \leq j \leq q, 1 \leq k \leq p}$ which is not a square matrix. Then solving the linearization is the same as lifting a tangent vector in the x-space to a vector in $(p + 1 - q)$-space.

19.1 Continuity method using convex sets

But the best you can do is to get a local solution. If something is in the image, only a small open neighborhood is in the image as well. Note that the space $\mathcal{G}$ is a convex cone in the Fréchet space $\Gamma_{C^\infty}(\mathrm{Sym}^2 T^* M)$. We also know that the image E^∞ of $\mathcal{F}$ of $\mathcal{G}$ is also a cone, because we can embed $M \to \mathbb{R}^{N_1} \oplus \mathbb{R}^{N_2}$. Here, the cone E^∞ is equal to $\mathcal{G}$ if

(i) E^∞ is dense in $\mathcal{G}$, and

(ii) E^∞ contains an interior point.

First lets try to produce an interior point, in the set of realizable Riemannian metric. We start out with one embedding $M^n \to \mathbb{R}^s$. Now can we construct a right inverse of its variation? Nash's trick is to avoid differentiation in the process of pulling back the metric. Suppose that the embedding $w_1, \dots, w_s$ are functions on $\mathbb{R}^n$ with periods. (We're assuming M is a torus.) Then the pullback metric is

$$g = \sum_{l=1}^{s} \sum_{j,k=1}^{n} \frac{\partial w_l}{\partial x_j} \frac{\partial w_l}{\partial x_k} dx_j \otimes dx_k.$$

If we vary w to $w + h$, we will get

$$\delta g = \sum_{l=1}^{s} \sum_{j,k=1}^{n} \Big(\frac{\partial w_l}{\partial x_j} \frac{\partial h_l}{\partial x_k} + \frac{\partial h_l}{\partial x_j} \frac{\partial w_l}{\partial x_k} \Big) dx_j dx_k.$$

Here, the trick is to not differentiation $\frac{\partial h}{\partial x}$. The way to do this is to impose an additional condition

$$\sum_l \frac{\partial w_l}{\partial x_j} h_l = 0$$

for all $1 \le j \le n$. If we have this condition, then

$$\sum_{l=1}^{n} \frac{\partial^2 w_l}{\partial x_j \partial x_k} h_l + \sum_{l=1}^{n} \frac{\partial w_l}{\partial x_j} \frac{\partial h_l}{\partial x_k} = 0$$

It follows that

$$\delta g = -2 \sum_{l,j,k} \frac{\partial^2 w_l}{\partial x_j \partial x_k} h_l.$$

Now inverting this operator becomes a completely linear question.

The inverse we want to consider is the linear equation

$$\begin{cases} \sum_{l=1}^{s} \frac{\partial w_l}{\partial x_j} h_l = 0 & 1 \le j \le n \\ \sum_{l=1}^{s} \frac{\partial^2 w_l}{\partial x_j \partial x_k} h_l = g_{jk} & 1 \le j, k \le n. \end{cases}$$

This has s variables, and there are $n + \frac{n(n+1)}{2}$ equations. So if s is sufficiently large and the rank of this matrix is $n + \frac{n(n+1)}{2}$, we can find the minimal solution by using the "generalized Cramer's rule":

$$h_{\min} = B^t (BB^t)^{-1} \vec{b}, \quad \vec{b} = \begin{pmatrix} 0 \\ g_{jk} \end{pmatrix} \in C^\infty(M).$$

So we want make sure that for a good f, the matrix has maximal rank everywhere. Here we use the technique of mapping local coordinates. If we make a change of coordinates, the rank of the matrix doesn't change. So we choose a finite covering and charts, and just add the coordinates $x_1, \dots, x_s, x_1^2, x_1 x_2, \dots, x_{s-1} x_s, x_s^2$ (multiplied with a bump function) to the w_j. If we do this, we will get a embedding $w : M \to \mathbb{R}^s$ such that the matrix has full rank everywhere. Then applying the Newton's method with a family of norms, we will get that this embedding w is in the interior of E^∞.

We also need to show that the image E^∞ is dense in $\mathcal{G}$. Suppose not, so that $g \in \mathcal{G} - \overline{E}^\infty$. Then by Hahn–Banach, there exists a function Φ on $C^\infty(M)^{\frac{n(n+1)}{2}}$ such that $\varphi(E^\infty) \le 0$ and $\Phi(g) > 0$. Then Φ is a distribution, written as

$$\Phi(g) = \sum_{j,k=1}^{n} D_{jk} g_{jk}.$$

The condition we have is that

$$\sum D_{jk} \sum_{l} \frac{\partial w_l}{\partial x_j} \frac{\partial w_l}{\partial x_k} \le 0.$$

Locally, we use translation to smooth out D_{jk} to D_{jk}^ϵ. Then we see that any smoothing out D_{jk}^ϵ is semi-negative. If we apply this to g_{jk}, then we also get $\sum D_{jk}^\epsilon g_{jk} \le 0$. As we take $\epsilon \to 0$, we get a contradiction.

20 April 5, 2018

We were looking at the isometric embedding theorem. Here, we needed to change the norms in the topological vector space. To move between them, we had to define smoothing operators. This local picture is the Nash–Moser implicit function theorem.

20.1 Conjugacy problem

The second application of the Nash–Moser implicit function theorem is the conjugacy problem, which deals with stability of the planetary motions. If there are two planets, the motion is periodic. So the idea was to solve them independently, and then regard other interactions as perturbations.

For $f : \mathbb{R}/\mathbb{Z} \to \mathbb{R}/\mathbb{Z}$ a continuous map, its **rotation** is the constant ρ such that

$$\lim_{n\to\infty} \frac{1}{n}(f^n(x) - x)$$

converges uniformly to the constant function ρ.

Consider a rotation $R_\rho : x \mapsto x + \rho$, and take a small perturbation

$$\varphi : x \mapsto x + \rho + \eta(x).$$

Now we want a reparametrization of the circle that makes the perturbed map the same as the original rotation by ρ. That is, is there a function $H(x)$ such that $H^{-1}\varphi H = R_\rho$?

Historically, people tried to do this using Fourier series, but then there is the division by zero problem. If ρ is irrational, it is fine, but still there is a small divisor problem when we take the iteration f^n and so on. In the analytic case, Arnold managed to solve this using smaller and smaller norms by looking at the function on a neighborhood of $\mathbb{R}$ on the complex plane. Later, Yoccoz came up with a way of doing this.

Definition 20.1. We say that $\rho \in \mathbb{R}$ is of type (K, ν) if there exists $K > 0$ and $\nu > 0$ such that

$$\left|\rho - \frac{m}{n}\right| > \frac{K}{|n|^\nu}$$

for all $(m, n) \in \mathbb{Z} \times (\mathbb{Z} - \{0\})$.

Roth showed that if ρ is irrational and algebraic, then it is of type $(K, 2+\epsilon)$ for all ϵ.

Theorem 20.2 (Arnold). *If ρ is of type (K, ν), then there exists $\epsilon(K, \nu, \sigma) > 0$ such that if $\phi(x) = x + \rho + \eta(x)$ has rotation number ρ and $\|\eta\|_\sigma < \epsilon(K, \nu, \sigma)$ (this η is real-analytic and is the restriction of some holomorphic function on the strip $\{|\Re z| < \sigma\}$ with the sup norm $\|\eta\|_\sigma$ on the strip) then there exists a real-analytic change of variables $\xi = H(x)$ for which ϕ is conjugate to R_ρ.*

Let us write $\xi = H(x) = x + h(x)$. What we want is $\phi \circ H = H \circ R_\rho$, so we can write this as

$$\xi + h(\xi) + \rho + \eta(\xi + h(\xi)) = \xi + \rho + h(\xi + \rho).$$

So we have

$$h(x + \rho) - h(x) = \eta(x + h(x)).$$

If we look at its linearized equation $h(x + \rho) - h(x) = \eta(x)$, we can take its Fourier transform and write

$$\hat{h}(n)e^{2\pi i n\rho} - \hat{h}(n) = \hat{\eta}(n)$$

What Arnold did is to show that if ρ has some mild condition, you can still solve the small divisor theorem.

On the other hand, Yoccoz showed that this is actually not possible under some condition. For α an irrational, we define $\alpha_0 = \{\alpha\}$ and $\alpha_1 = \{\frac{1}{\alpha_0}\}$ and so on. (This is taking the continued fraction.) Define

$$\Delta_0(\alpha) = 10, \quad \Delta_{n+1}(\alpha) = \begin{cases} \frac{1}{\alpha_n}(\Delta_n(\alpha) - \log\frac{1}{\alpha_n} + 1) & \Delta_n \geq \log\frac{1}{\alpha_n} \\ e^{\Delta_n(\alpha)} & \Delta_n \leq \log\frac{1}{\alpha_n}. \end{cases}$$

Then define

$$\begin{aligned} \mathcal{H}_0 &= \{\alpha \in \mathbb{R} - \mathbb{Q} : \text{for all } n \geq n_0, \Delta_n(\alpha) \geq \log\tfrac{1}{\alpha_n}\}, \\ \mathcal{H} &= \{\alpha \in \mathbb{R} - \mathbb{Q} : g\alpha \in \mathcal{H}_0 \text{ for all } \mathrm{GL}_2(\mathbb{Z})\}. \end{aligned}$$

Yoccoz in 1989 showed that the conjugation problem is always solvable if and only if $\alpha \in \mathcal{H}$, where α is the rotation number.

20.2 Arnold's theorem

Now let us go back to the estimates. We have

$$\hat{h}(n) = \frac{\hat{\eta}(n)}{e^{2\pi i n\rho} - 1}$$

for $n \neq 0$. If ρ is of type (K, ν), we get

$$|e^{2\pi i n\rho} - 1| \geq \frac{4K}{|n|^{\nu - 1}}.$$

Now our function f is analytic and periodic, so we can write it as a holomorphic function on a neighborhood of the unit circle S^1. Then $\hat{\eta}(n)$ is the Laruent series coefficient. In any case, we will have the estimate

$$|\hat{\eta}(n)| \leq \|\eta\|_\sigma e^{-2\pi\sigma|n|}.$$

Going back, we have

$$h(z) = \sum_{n \neq 0} \frac{\hat{\eta}(n) e^{2\pi i n z}}{e^{2\pi i n \rho} - 1}.$$

For $\|\Im t\| < \sigma - \delta$, the sum can be bounded above by

$$\sum_{n \neq 0} \frac{|n|^{\nu - 1}}{4K} \|\eta\|_\sigma e^{-2\pi\sigma|n|} e^{2\pi|n|(\sigma-\delta)} \leq \frac{\Gamma(\nu)}{K(2\pi\delta)^\nu}.$$

We estimate this by comparing it with the Γ-function. This is not a complete solution because we have to handle $\hat{\eta}(0)$. Still we have

$$\|h\|_{\sigma-\delta} \leq \frac{\Gamma(\nu)}{K(2\pi\delta)^\nu} \|\eta\|_\sigma.$$

This is then going to be the solution to the linearized problem.

21 April 10, 2018

We were looking at the conjugacy problem. The technique used here is to shrink the strip at every step, and use these weaker norms. We are going to see another technique of shrinking the domain. Here, Moser's idea is that we can use a cutoff function that is 1 on a shrunken domain. We are going to need Harnack's principle, which can handle elliptic nonlinear partial differential equations. These equations will look like

$$F(x, u, Du, D^2u) = 0$$

where $\frac{\partial F}{\partial D^2 u}$ is positive definite. De Giorgi only assumed here that the coefficients are only measurable, but they are uniformly elliptic.

21.1 Iterating linearized conjugations

We were trying to prove the following theorem.

Theorem 21.1 (Arnold)**.** *Let ρ be irrational of type (K, v), i.e., $|\rho - \frac{m}{n}| > \frac{K}{|n|^\nu}$ for all integers m, n. Then there exists a $\epsilon = \epsilon(K, \nu, \sigma) > 0$ such that if $\phi(x) = x + \rho + \eta(x)$ has rotation ρ and $|\eta|_\sigma < \epsilon$ (i.e., η is holomorphic on $S_\sigma = \{|\Im(z)| < \sigma\}$ with $\eta(x + 1) = \eta(x)$ and $|\eta|_\sigma = \sup_{S_\sigma} < \epsilon$) then there exists a real-analytic change of coordinates $\xi = H(x)$ such that ϕ is conjugate to $x \mapsto R_\rho(x) = x + \rho$.*

Proof. If we write $H(x) = x + h(x)$, the equation we need to solve is

$$h(x + \rho) - h(x) = \eta(x + h(x)).$$

First we look at the linearization. Then we get

$$\hat{h}(n) = \frac{\hat{\eta}(n)}{e^{2\pi i n\rho} - 1}$$

for $n \neq 0$. There is a small divisor problem, but we got that

$$\|h\|_{\sigma-\delta} \leq \frac{\Gamma(v)}{K(2\pi\delta)^v} \|\eta\|_\sigma.$$

Now we need to invert $H(x) = x + h(x)$. For $H^{-1}(z) = z - h(z) + g(z)$, we get

$$g(z + h(z)) = \int_{s=0}^{1} h'(z + sh(z))h(z)ds$$

and so we get

$$\|g\|_{\sigma-4\delta} \leq \frac{2\pi\Gamma(v)^2}{K^2(2\pi\delta)^{2v+1}} \|\eta\|_\sigma^2.$$

Note that we have not taken care of $\hat{\eta}(0)$. If we plug in $H^{-1}\phi H$, we get

$$\tilde{\phi}(x) = x + \rho + \hat{\eta}(0) + \int_0^1 \eta'(x + sh(x))h(x)ds$$
$$+ \int_0^1 h'(x + \rho + s(h(x) + \eta(x + h(x)))(h(x) + \eta(x + h(x))))ds.$$

Define $\tilde{\eta}(x)$ by $\tilde{\phi}(x) = x + \rho + \tilde{\eta}$. Because the rotation number of ϕ is precisely ρ, there exists a x_0 such that $\eta(x_0) = 0$. So $\tilde{\eta}(x_0) = 0$ for x_0. Using this, we can approximate $\tilde{\eta}$ by using this point x_0. At the end, we will get

$$\|\tilde{\eta}\|_{\sigma-6\delta} \leq \frac{16\pi\Gamma(v)^2}{K^2(2\pi\delta)^{2v+1}}\|\eta\|_\sigma^2.$$

Now we can iterate this. At each step, you can use

$$\delta_n = \frac{\sigma}{36(1+n^2)}$$

and this gives the right result. □

21.2 Harnack inequality

Consider

$$\sum_{\nu,\mu=1}^n \frac{\partial}{\partial x_\nu}\Big(a_{\nu\mu}(x)\frac{\partial u}{\partial x_\mu}\Big) = 0 \qquad (\dagger)$$

for $x = (x_1, \dots, x_n)$ where $a_{\nu\mu} = 0$ are only Lebesgue measurable, but has $\frac{1}{\lambda} \leq a_{\mu\nu} \leq \lambda$ on a domain $D \subseteq \mathbb{R}^n$. Let $D' \subset\subset D$. If u is a weak solution of (†) on D and $u > 0$ on D, then the inequality states that

$$\max_{D'} u \leq C \min_{D'} u$$

for some $C = C(D, D', \lambda)$. This can be found in Moser, *On Harnack's theorem for elliptic differential equations*, Comm. P.&App. **14** (1961).

Here, when we say weak solution, we meant that $u \in L_1^2(D)$ and for all $\phi \in L_1^2(D)$, we have

$$\int_D (\phi_x, au_x) = 0.$$

The idea is to integrate by parts to estimate L^2 of Du in terms of L^2 of u. We can interpret the maximum and the minimum as the L^∞ norm of u^l for $l > 0$ and $l < 0$. Then we can relate the two by using $\log u$ and the John–Nirenberg inequality, which reverses the Hölder inequality.

22 April 12, 2018

We started looking at De Giorgi's ellipticity. For the operator $\sum_{k=1}^n \partial_j(a_{jk}\partial_k u)$ for a_{jk} only L^1_{loc} but $\frac{1}{\lambda} \le (a_{jk}) \le \lambda$, De Giorgi proved some Hölder regularity for u.

22.1 Proof of Harnack's inequality

Theorem 22.1 (Harnack's inequality)**.** *Let $D' \subset\subset \mathbb{R}^n$ be a bounded domain. Let u be a (weak) solution of the elliptic equation in divergence form on D whose matrix of coefficients for the 2-order terms have matrix entries bounded between $\frac{1}{\lambda}$ and λ for some $\lambda > 1$. If $u > 0$ on D, then there exists a $c = c(n, D, D')$ such that*

$$\sup_{D'} u \le c \inf_{D'} u.$$

Here, the weak solution is defined as

$$\int (\varphi_x, au_x) = 0$$

for all compactly supported $\varphi \in L^2_1(D)$. We apply integration by parts to bound L^2 of $D(u^k)$ by L^2 of u^k. The point is that we need to take a cutoff function to avoid the occurrence of boundary terms. This forces us to shrink the domain.

Consider $v = f(u) = u^k$ for $k \ge 2$. Then for a cutoff function η, we let

$$\phi = ku^{2k-1}\eta^2.$$

Then $\phi_x = k(2k-1)u^{2k-2}u_x\eta^2 + 2ku^{2k-1}\eta\eta_x$. Then our condition gives

$$\begin{aligned} 0 = \int \phi_x a u_x dx &= \int k(2k-1)u^{2k-2}u_x a u_x \eta^2 + 2ku^{2k-1}\eta\eta_x a u_x \\ &= \int \frac{2k-1}{k}\eta^2 (u^k)_x a (u^k)_x + \int 2u^k \eta\eta_x (u^k)_x. \end{aligned}$$

Then we end up with

$$\frac{2k-1}{k}\frac{1}{\lambda}\int_D (\eta v_x)^2 \le 2\int_D |(\eta_x v, a\eta v_x)|,$$

and Cauchy–Schwartz gives

$$\Big(\frac{2(2k-1)}{k\lambda^2}\Big)^2 \int_D (\eta v_x)^2 \le \int (\eta_x v)^2.$$

We have $\eta = 1$ on some smaller domain D'. Here, $\eta_x = 0$ but η_x is big outside D'. So we need to keep track of shrinking of domain. Take $D = Q(h)$ be a cube with side h centered at 0, and take $D' = Q(h')$. Then

$$\int_{Q(h')} v_x^2 dx \le c_3 \Big(\frac{2k}{2k-1}\Big)^2 \Big(\frac{2}{h-h'}\Big)^2 \int_{Q(h)} v^2 dx.$$

If we apply Sobolev to the left hand side, we gain some power, and we find that there is some constant $\kappa > 1$ such that

$$\left(\fint v^{2\kappa}\right)^{1/\kappa} \le \beta\left(h'^2 \fint v_x^2 + \fint v^2\right).$$

If we put this in, we get

$$\left(\fint_{Q(h')} v^{2\kappa}\right)^{1/\kappa} \le c\left(\left(\frac{h}{h'} - 1\right)^{-2}\left(\frac{2k}{2k+1}\right)^2 + 1\right)\fint_{Q(h)} v^2 dx.$$

Fix $p > 1$. Every time, we jack up by a factor of $\kappa > 1$. So for $p_\nu = \kappa^\nu p$ we use $h = h_\nu = 1 + 2^{-\nu}$ and $h' = h_{\nu+1} = 1 + 2^{-\nu-1}$. Then if we iterate this inequality, we are going to get something like

$$\sup_{D'} u = \|u\|_{L^\infty(D')} \le C\|u\|_{L^p(D)}.$$

But we still need to relate the positive u^k to the negative u^{-k}. To do this, we now use $v = \log u$. Then we are using $\phi = \eta^2/u$. In this case,

$$(\phi_x, au_x) = -\eta^2(v_x, av_x) + 2\eta(\eta_x, av_x)$$

as before, and because a is positive definite, we get

$$\int_Q \eta^2 v_x^2 dx \le 4\lambda^4 \int_Q \eta_x^2.$$

On the other hand, if we have the Poincaré inequality

$$\int_Q (v - v_Q)^2 \le C \int v_x^2 dx \le C.$$

(This can be seen from looking at the Fourier expansion.)

Now the John–Nirenberg inequality gives, for $w \in L^2(Q(1))$, if $\int_Q (w - w_Q)^2 \le 1$ then for every $Q = Q(h) \subset Q(\frac{1}{2})$, there exist $\alpha, \beta > 0$ such that

$$\left(\int_{Q(\frac{1}{2})} e^{\alpha w}\right)\left(\int_{Q(\frac{1}{2})} e^{-\alpha w}\right) \le \beta^2.$$

23 April 17, 2018

Today we are going to wrap up the Harnack inequality and the Hölder estimates. Consider

$$L = \sum_j \partial_j \sum_k a_{jk}(x)\partial_k$$

where (a_{jk}) are measurable but bounded between $1/\lambda$ and λ. Then the Harnack inequality tells us that if $u > 0$ is a solution on D and $D' \subset\subset D$ then

$$\sup_{D'} u \le c_{\lambda,D,D',u} \inf_{D} u.$$

For $v = \log u$, the John–Nirenberg inequality gives

$$\left(\int_{D'} |u|^p\right)\left(\int_{D'} |u|^{-p}\right) \le C$$

for some $p > 0$. Now Moser observed that if you take $v = u^{(p+1)/2}$ for $p \ne -1$, we get $(av_x)_x = 0$ and so $(\phi_x, av_x) = 0$. Then $v_x = \frac{p+1}{2}u^{(p-1)/2}u_x$ and so

$$v_x^2 = \frac{(p+1)^2}{4}u^{p-1}u_x u_x.$$

When we use $(\phi_x, au_x) = 0$ for $\phi = v\eta^2$, we can bound $\int v_x^2$ by $\int v^2$.

23.1 De Giorgi's argument for the Hölder estimates

For $Lu = 0$ with $u > 0$, let us write

$$M(r) = \sup_{|x|\le r} u, \quad \mu(r) = \inf_{|x|\le r} u.$$

Now take $D = \{|x| < r\}$ and $D' = \{|x| < r' = \frac{r}{2}\}$. Write $M = M(r), \mu = \mu(r)$ and $M' = M(r'), \mu' = \mu(r')$. Then $M - u$ and $u - \mu$ on D are both the solutions for $Lu = 0$ as well. Then we can apply the Harnack inequalities here and get

$$\begin{aligned} \sup_{D'}(M-u) &= M - \mu' \le c(M - M') = c\inf_{D'}(M-u), \\ \sup_{D'}(u-\mu) &= M' - \mu \le c(\mu' - \mu) = c\inf_{D'}(u-\mu). \end{aligned}$$

So if we add them up, we get

$$\operatorname{osc}_{D'} u \le \alpha \operatorname{osc}_D u,$$

where $\operatorname{osc} = M - \mu$ and $0 < \alpha < 1$ is given by $\frac{c-1}{c+1}$.

The point of the argument is that this already implies 1/2-Hölder. This is because

$$\operatorname{osc}\left(\frac{r}{2^\nu}\right) \le \alpha^\nu \operatorname{osc}(r).$$

You can apply this to Bernstein's problem of minimal surfaces, but I am not going to do this.

23.2 Probability spaces

Let (X, m) be a probability space. This means that $m \geq 0$ is a measure with respect to some σ-algebra $\mathfrak{M}$, and the total measure is $m(X) = 1$. A $\mathfrak{M}$-measurable function is called a **random variable**. The toy model for this is flipping coins. You flip this many many times, and are interested in the limiting situation.

This can be done by using the Rademacher function. If you flip the coin n times, you get something in $(\mathbb{Z}/2)^n$ and we are going to identify with $[0, 1]$ by the diadic interpretation. So we identify $(\mathbb{Z}/2)^\infty \sim [0, 1]$. So given a real number $x \in [0, 1]$, we can consider the function

$$r(x) = \begin{cases} 1 & 0 \leq x < \frac{1}{2}, \\ -1 & \frac{1}{2} \leq x \leq 1. \end{cases}$$

So we can evaluate $r(x), r(2x), r(4x), \ldots$.

The main tool is the independence of measurable functions. This is much stronger than orthogonality. A family of functions $\{f_n\}$ is said to be **independent** if

$$m\bigg(\bigcap_{n=1}^{\infty} \{x : f_n(x) \in B_n\}\bigg) = \prod_{n=1}^{\infty} m(\{x : f_n(x) \in B_n\}).$$

This is a much stronger notion that just orthogonality, because if $\Phi, \Psi : \mathbb{R} \to \mathbb{R}$ are continuous functions, and f, g are independent, then $\Phi \circ f$ and $\Psi \circ g$ are independent. You can think of this as the push-forward measure to the product being equal to the product of the push-forward measure to each of the components.

On $X = (\mathbb{Z}/2)^\infty = [0, 1]$, the Rademache functions $r_n(t) = r_1(2^{n-1}t)$ form a basis for the independent functions. The advantage of this tool is that we can now imitate Fourier analysis. If we write

$$S_N(t) = \sum_{n=1}^{N} r_n(t)$$

then we have $\|S_N\|_{L^2} = \sqrt{N}$. We can also show

$$m\left(\left\{x : \alpha < \frac{S_N(x)}{\sqrt{N}} < \beta\right\}\right) = \int_\alpha^\beta \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt.$$

This follows from an application of Stirling.

24 April 19, 2018

Today we'll solve Dirichlet's problem by Kakuatni's method. Consider a bounded domain $\Omega \subseteq \mathbb{R}^n$ and suppose we are given the boundary value of g on Ω. We want to find f harmonic on Ω with boundary value g.

24.1 Theorems in probability theory

Let's first try to solve this problem from $g = \chi_E$. We are going to define

$$f_E(x) = \Pr(\text{Brownian motion from } x \text{ to exit } \Omega \text{ at a point of } E).$$

The idea for the Brownian motion is to discretize and take the limit. Here, we used $(\mathbb{Z}/2\mathbb{Z})^\infty$ to model the infinite tossing of coins. Because we have to normalize, the sample path of a Brownian motion is going to be approximately $\frac{1}{2}$-Hölder continuous.

Consider a discrete time martingale. Then we have a sequence of σ-algebras

$$\mathscr{A}_n \subseteq \mathscr{A}_{n+1} \subseteq \mathscr{A}_{n+2} \subseteq \cdots \subseteq \mathscr{M}.$$

Then for f a function that is $\mathscr{M}$-measurable, we can construct $\mathbb{E}_n(f)$ which is $\mathscr{A}_n$-measurable. This is by looking at the Radon–Nikodym derivative and writing $\mu_n = \int_{\mathscr{A}_n} f$. Then wone can show that

$$f_n = \mathbb{E}_n(f_\infty) \to f$$

almost everywhere and in L^1.

Theorem 24.1 (Kolmogorov 0-1 law). *Let (X, m) be a probability space and $f_1, f_2, \dots$ be mutually independent random variables. Then the probability that $\sum_{k=1}^\infty f_k$ converges is either 0 or 1.*

Proof. Let's say that sub σ-algebras $\{\mathscr{B}_j\}$ are mutually independent if and only if $m(\bigcap_j B_j) = \prod_j m(B_j)$ for all $B_j \in \mathscr{B}_j$. This is a generalization because f_i being independent just means that $\mathscr{B}_j = f^*(\text{Borel on } \mathbb{R})$ are mutually independent.

Kolmogorov's observation is that if $\mathscr{B}$ is independent with itself then $m(B) = m(B \cap B) = m(B)^2$. Consider $\mathscr{A}_n = \mathscr{A}_{f_n}$ and consider $\mathscr{B} = \bigcap_{l=1}^\infty (\bigvee_{k=l}^\infty \mathscr{A}_k)$. Then $\mathscr{B}_j$ are mutually independent so $\mathscr{B}_n$ is independent with $\mathscr{B}$ for all n. Then $\mathscr{B}$ is independent with itself.

Now it suffices to show that the subset of $\sum_k f_k$ converging is in $\mathscr{B}$. This can be done, by writing down the condition for convergence. $\square$

Theorem 24.2 (central limit theorem). *Let $f_1, f_2, \dots$ be mutually independent and identically distributed, i.e., $(f_j)_* m$ are all equal. Consider the mean and variance*

$$m_0 = \int_{t=-\infty}^{\infty} t d\mu_f(t) = \int_{x \in X} f(x) dm(x),$$
$$\sigma^2 = \int_{-\infty}^{\infty} (t - m_0)^2 d\mu_f(t) = \int_{x \in X} (f(x) - m_0)^2 dm(x).$$

Let $S_N = \sum_{j=1}^N f_j$ *normalized to* $(S_N - Nm_0)/\sqrt{N}$*. Then*

$$m(\{x : a < \tfrac{1}{\sqrt{N}}(S_N - Nm_0) < b\}) \to \frac{1}{\sigma\sqrt{2\pi}} \int_{t=a}^{b} e^{-\frac{t^2}{2\sigma^2}} dt.$$

Proof. We can assume that $m_0 = 0$. Let us write μ_N the distribution of $s_N/\sqrt{N}$ and μ_σ the distribution of the Gaussian distribution. We want to show that $\mu_N((a,b)) \to \mu_\sigma((a,b))$ as $N \to \infty$. To do this, we look at the Fourier transform

$$\hat{\mu}(\xi) = \int_{t=-\infty}^{\infty} e^{2\pi i \xi t} d\mu(t).$$

Mutual independence implies that

$$\hat{\mu}_N(\xi) = \prod_{n=1}^{N} \hat{\mu}_{f_n}\Big(\frac{\xi}{\sqrt{N}}\Big) = \hat{\mu}\Big(\frac{\xi}{\sqrt{N}}\Big)^N.$$

You can expand $\hat{\mu}(\xi/\sqrt{N})$ in ξ and take power and look at the limit. $\square$

Theorem 24.3 (law of large numbers)**.** *Let* $f_1, f_2, \dots$ *be identically distributed mutually independent random variables on* (X, m)*. Then* $\frac{1}{N}\sum_{j=1}^N f_j \to m_0$ *almost everywhere.*

Proof. This uses the idea of ergodicity. Recall that if $T : Y \to Y$ is a measure-preserving ergodic transformation then $\frac{1}{N}\sum_{j=1}^N T^j f \to \int f$ almost everywhere. (This was proved last semester.) Now we consider $Y = \prod_{j=0}^\infty R_j$ and define $T : Y \to Y$ as a shift $T(y)_n = y_{n+1}$. This will be measure-preserving, and ergodic. $\square$

25 April 26, 2018

The role of analysis is to rigorize intuitive arguments involving infinite processes. Kakutani's idea uses probabilistic methods to solve the Dirichlet problem. Malliavin developed this idea further and invented Malliavin calculus. Kakutani's intuition is that if you take a random walk, and look at the probability that it is going first hit some region of the boundary, this will solve the Dirichlet problem.

25.1 Wiener measure

So in the infinite setting, we need to look at the set of all continuous paths. This we can just define as

$$\mathscr{P} = \{p : [0, \infty) \to \mathbb{R}^d\}.$$

But we want this to be a probability space, so we need to give a σ-algebra, and also a measure. Giving a σ-algebra is simple, because we can just define a measure

$$d_n(p, p') = \sup_{0 \le t \le n} |p(t) - p'(t)|, \quad d = \sum_n \frac{1}{2^n} \frac{d_n}{1 + d_n}.$$

You can also define it using the subbase

$$\{p : p(t_j) \in A_j\}$$

for some $0 = t_0 \le t_1 \le t_2 \le \cdots \le t_k$.

The hard thing is to define the measure. Wiener observed that you can use the cylindrical sets defined above, and then pass to the limit. This we can also abstractly characterize because we want this measure to be the measure coming from the limiting of the random walk. A random walk can be considered as $\mathbb{Z}_{2d}^\infty$ where d is the dimension. For $x \in \mathbb{Z}_{2d}^\infty$, we consider $r_k(x) \in \{\pm e_1, \dots, \pm e_d\}$ the increment at the kth position in the random walk x. Out of this, we can define the path

$$S_t^{(N)}(x) = \frac{1}{\sqrt{N}} \sum_{1 \le k \le [Nt]} \vec{v}_k(x) + \frac{Nt - [Nt]}{\sqrt{N}} \vec{r}_{[Nt]+1}(x).$$

This normalization by $1/\sqrt{N}$ is introduced so that as we take $N \to \infty$, the variance of the endpoint doesn't blow up.

Now we have this map

$$i_N : \mathbb{Z}_{2d}^\infty \to \mathscr{P}; \quad x \mapsto S_t^{(N)}.$$

Let μ_N be the measure given by pushing forward. We want to show that $\mu_N \to \mu$ in some sense. This can be made rigorous by looking at finite time slices. For $0 \le t_1 \le \cdots \le t_k$ in $\mathbb{R}$, we have an evaluation map $\mathscr{P} \to (\mathbb{R}^d)^k$ and look at the measure in $(\mathbb{R}^d)^k$ given by pushing forward. Then we get $\mu_N^{(t_1,\dots,t_k)}$ on $(\mathbb{R}^d)^k$ for each N, and then we can discuss convergence by evaluating at Borel sets. There is also a notion of weak convergence, which is that if you integrate a bounded measurable function, you get pointwise convergence.

The way you do this is to use Ascoli–Arzela.

Definition 25.1. A sequence of probability measures is called **tight** if for every $\epsilon > 0$ there exists a compact subset K_ϵ of $\mathscr{P}$ such that for all N we have $\mu_N(K_\epsilon^c) < \epsilon$. (In otherwise, it is almost supported on compact subsets.)

In $\mathscr{P}$, the notion of compactness is the following. We say that a closed set $K \subseteq \mathscr{P}$ is compact if for each $T > 0$, there exists a positive bounded function w_T such that $w_T(h) \to 0$ as $h \to 0$ and

$$\sup_{p \in K} \sup_{0 \le t < T} |p(t+h) - p(t)| \le w_P(h).$$

for all $0 < h \le 1$.

Theorem 25.2 (Ascoli–Arzela)**.** *If $\{\mu_N\}$ is tight, there exists a weakly converging subsequence.*

So what we can do is show that $\{\mu_N\}$ is tight. Then there exists a weakly converging subsequence, converging to μ. Now we can show that $\mu_N^{(t_1,\dots,t_k)}$ weakly converges to some measure, and then $\mu_N^{(t_1,\dots,t_k})) \to \mu^{(t_1,\dots,t_k)}$ weakly because it converges on a subsequence.

To check that $\mu_N^{(t_1,\dots,t_N})$ are weakly convergent, you can use martingales. Because $\mathbb{Z}_2^\infty$ is like $[0,1]$, we can test only on Rademacher functions.

25.2 Kakutani's solution to the Dirichlet problem

Let $(\Omega, W) = (\mathscr{P}, \mu)$, so that for $\omega \in \Omega$ we have $\omega(t)$. This is also written as $B_t(\omega)$, where B_t is thought of as the Brownian motion.

It can be formulated abstractly. Take an abstract probability space (Ω, P) and for each $t \ge 0$ consider a $\mathbb{R}^d$-valued random variable $B_t(\omega)$ for $\omega \in \Omega$. We say that this is a **Brownian motion** if the following are satisfied.

(B1) (independence of increment) The variables $B_{t_1}, B_{t_2} - B_{t_1}, \dots, B_{t_k} - B_{t_{k-1}}$ are mutually independent for $0 \le t_1 < t_2 < \cdots < t_k$.

(B2) (increment in Gaussian) For $0 \le t < \infty$ and $h > 0$, the variable $B_{t+h} - B_h$ is a Gaussian with mean zero and covariance matrix hI_d.

(B3) (almost everywhere continuity) For almost all $\omega \in \Omega$, the path $t \mapsto B_t(\omega)$ is continuous.

Stopping time is something that can be used for restarting. For $0 \le t < \infty$, let $\mathscr{A}_t$ be the smallest σ-algebra such that all functions B_s for $0 \le s \le t$ are $\mathscr{A}_t$-measurable.

Definition 25.3. A nonnegative function $\tau : \Omega \to [0, \infty)$ is a **stopping time** if

$$\{\omega : \tau(\omega) \le t\} \in \mathscr{A}_t$$

for all $t \ge 0$.

Brownian motions have a strong Markov property.

Theorem 25.4. *If B_t is a Brownian motion and $\sigma \geq 0$ is a stopping time, then*

$$B_t^*(\omega) = B_{t+\sigma(\omega)}(\omega) - B_{\sigma(\omega)}(\omega)$$

is also a Brownian motion.

Let $\mathcal{R}$ be a bounded domain in $\mathbb{R}^d$. For $x \in \mathcal{R}$, let $B_t^x(\omega) = x + B_t(\omega)$ and define the exit time as

$$\tau^x(\omega) = \inf\{t \geq 0 : B_t^x(\omega) \in \mathcal{R}\}.$$

For problems that occur at boundary points, we also define the strict exit time as

$$\tau_*^x(\omega) = \inf\{t > 0 : B_t^x(\omega) \in \mathcal{R}\}.$$

If $x \in \partial\mathcal{R}$ satisfies $\tau_*^x(\omega) = \tau^x(\omega)$ for almost every ω, we way that $x \in \partial\mathcal{R}$ is regular. In this case, we can solve the Dirichlet problem by just looking at where it first hits the boundary.

This is used in the uniformization problem. Given M a negatively curved simply connected manifold, what can we say about it? By Hadamard, it is going to be diffeomorphic to $\mathbb{R}^n$. But if it is negatively curved, you can find a bounded harmonic function. The idea is that if you look at a Brownian motion, anything moving in the direction transverse to the radial direction, it is sort of trapped because the non-radial direction is exponentially long.

Index