Lectures on PDE

© Leon Simon

Mathematics Department
Stanford University

Contents

Preface to Tsinghua Lectures

The present rough draft originated from various introductory graduate PDE classes taught by the author, and is still in rather unfinished form, with a generous scattering of (hopefully not serious, mainly expository) errors. And there are other serious deficiencies—for example almost no proper references to the literature are presently incorporated into the text.

The notes will be regularly updated during the course of the lectures (March, April 2015). Notice that not all material in the notes will be covered explicitly in the lectures. Conversely, some material to be covered in the lectures is presently not mentioned in the notes, but hopefully this will be remedied by later updates.

The author would greatly appreciate feedback about errors and other deficiencies.

Leon Simon
Tsinghua University, March 2, 2013
lsimon@stanford.edu

Last updated 5:00pm March 5, 2015

Lecture 1

Key Examples, and Some General Remarks

We begin by presenting a list of twenty examples of P.D.E.'s. The list represents examples which are either of fundamental importance for any introductory discussion, or examples which have been the focus of important recent research efforts. The first ten examples are linear, the remainder non-linear. Quite a number of the examples on the list will be discussed in some detail during the course of the lectures.

1. The Cauchy Riemann Equations

$$\begin{cases} u_x = v_y \\ u_y = -v_x \end{cases}$$

This is perhaps the most important example of a first order system of equations; a pair of $C^1(\Omega)$ functions u, v (Ω an open subset of $\mathbb{R}^2$) are solutions if and only if the complex-valued function $u + iv$ is holomorphic in Ω.

2. Laplace's Equation

$$\Delta u = 0, \quad \text{where } \Delta = \sum_{i=1}^{n} D_i^2, \;\; D_i = \frac{\partial}{\partial x^i}.$$

The single most important example of a second-order equation. Notice that the operator Δu can be written div (grad u), where grad $u = Du$ is the gradient of u.

3. The general second order divergence-form linear elliptic equation

$$\sum_{i,j=1}^{n} D_i(a^{ij}(x)D_j u) = 0,$$

where $(a^{ij}(x))$ is a symmetric positive-definite $n \times n$ matrix with smooth dependence on x, $x = (x^1, \ldots, x^n) \in \Omega$, Ω open in $\mathbb{R}^n$. (Actually one could also consider the case when (a^{ij}) is non-symmetric.) If $n \geq 3$, in the language of Riemannian geometry this is Laplace's equation relative to the metric $g_{ij}dx^i dx^j$ where

$(g_{ij}) = ((g^{ij})^{-1})$, with $(g_{ij}) = (\det(a^{ij})^{\frac{-1}{n-2}})(a^{ij})$. If $n = 2$ the same geometric interpretation holds with $(g_{ij}) = (a^{ij})^{-1}$ provided $\det(a_{ij}) = 1$, otherwise not.

4. The Heat Equation

$$u_t - \Delta u = 0$$

This equation provides the usual model for heat flow in a homogeneous isotropic medium; in this model $u = u(x,t)$ represents the temperature at the point x at time t. Notice that equilibrium solutions (i.e. solutions with $u_t \equiv 0$) satisfy Laplace's equation $\Delta u = 0$. Notice also that we can get "separated variable" solutions $u(x,t) = e^{-\lambda t}\varphi_\lambda(x)$, where λ is any real constant and $\varphi_\lambda(x)$ is any solution of the equation $-\Delta\varphi = \lambda\varphi$.

5. Heat Equation for an inhomogeneous anisotropic medium

$$u_t - \lambda \sum_{i,j=1}^{n} D_i(a^{ij} D_j u) = 0,$$

where $\lambda > 0, (a^{ij})$ is positive definite. For a crystalline solid a^{ij} may be constant within each crystal but have jump discontinuities across the crystal boundaries.

6. Schrödinger's equation

$$iu_t - \Delta u - qu = 0$$

where q is a given function of x.

7. Time Independent Schrödinger Equation

$$\Delta u + qu = 0$$

Notice that this is the equation satisfied by equilibrium solutions (i.e. time-independent solutions) of the Schrödinger equation.

8. The Wave Equation

$$u_{tt} - \Delta u = 0$$

Provides the usual model for wave propagation in a homogeneous isotropic medium. Notice again that an equilibrium solution satisfies $\Delta u = 0$.

9. The Biharmonic Equation

$$\Delta^2 u = 0$$

This is perhaps the most important single example of a fourth order equation.

10. The General m^{th}-order Linear Equation

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

where the notation is as follows: $x = (x_1, ..., x_n) \in \mathbb{R}^n, \alpha = (\alpha_1, ..., \alpha_n)$, α_j non-negative integers, $|\alpha| = \sum_{j=1}^{n}\alpha_j$ (the α are called multi-indices), $D^\alpha = D_1^{\alpha_1} ... D_n^{\alpha_n}$.

11. The Minimal Surface Equation

$$\Delta u - \sum_{i,j=1}^{n} \frac{D_i u D_j u}{1 + |Du|^2} D_i D_j u = 0$$

This is the "Euler-Lagrange equation" for the area functional, and can alternatively be written $\sum_{i=1}^{n} D_i(\frac{D_i u}{\sqrt{1+|Du|^2}}) = 0$. (See discussion of Euler-Lagrange equations below.)

12. The Capillary Surface Equation

$$\Delta u - \sum_{i,j=1}^{n} \frac{D_i u D_j u}{1 + |Du|^2} D_i D_j u = (\kappa u + \lambda)\sqrt{1 + |Du|^2}$$

This equation is the usual model for the equilibrium free surface of a fluid in a tube with vertical side-walls. (The free surface of the fluid is represented as graph$u$, with u satisfying the above equation.)

13. The Maximal Surface Equation

$$\Delta u + \sum_{i,j=1}^{n} \frac{D_i u D_j u}{1 - |Du|^2} D_i D_j u = 0, \qquad |Du| < 1.$$

This equation arises naturally in classical general relativity theory; it is the Euler-Lagrange equation of the area functional in Minkowski Space $\{(x,t) : x \in \mathbb{R}^n, t \in \mathbb{R}\}$, equipped with the Minkowski metric $ds^2 = dx^2 - dt^2$. Notice that it is the first of the equations we have mentioned so far with a "built in" constraint concerning the class of allowed solutions (in this case we have the restriction $|Du| < 1$.)

14. The Equation of Gas Dynamics

$$\Delta u - \sum_{i,j=1}^{n} \frac{D_i u D_j u}{1 - ((\gamma-1)/2)|Du|^2} D_i D_j u = 0.$$

Notice that $\gamma = -1$ gives the minimal surface equation.

15. The Monge-Ampere Equation

$$\det(D_i D_j u) = f(x, u, Du).$$

This equation arises naturally in geometry; the left side is, up to the multiplicative factor $(1 + |Du|^2)^{-(n/2+1)}$, the Gaussian curvature of the hypersurface $x^{n+1} = u(x^1, ..., x^n)$. It is the first example mentioned which is "fully non-linear", meaning that it is non-linear in the top order derivatives. (The examples 11, 12, 13 are "quasi-linear", meaning that they are *linear* in the top order derivatives.)

16. Yamabe's Equation

$$\frac{4(n-1)}{n-2} \Delta_g u - S(x)u + \lambda u^{(n+2)/(n-2)} = 0,$$

where $g = \sum_{i,j=1}^{n} g_{ij}\, dx^i dx^j$ is an arbitrary metric for $\mathbb{R}^n$ and where the dimension $n \geq 3$, $S(x)$ is the scalar curvature of g, and λ is a constant. Geometrically the main importance of this equation lies in the fact that if u is a positive solution, then the new metric $\widetilde{g} = u^{\frac{4}{n-2}} g$ has constant curvature λ.

(The reader who is not familiar with the terminology of Riemannian Geometry used here need not be unduly concerned; the more detailed later discussion will not assume such knowledge.)

17. The sinh-Gordon equation

$$\Delta u + \sinh u = 0$$

Doubly periodic solutions of this equation played an important role in Wente's recent construction of a counterexample of the Hopf conjecture concerning constant mean-curvature immersions in $\mathbb{R}^3$.

18. The Navier Stokes Equations

$$\frac{\partial u_i}{\partial t} + \sum_{k=1}^{3} \frac{\partial u_i}{\partial x^k} u_k = -\frac{1}{\rho}\frac{\partial p}{\partial x^i} + \gamma \Delta u_i, \quad i = 1, 2, 3$$

$$\operatorname{div} u = 0$$

where $u = (u_1, u_2, u_3)$ and p are the unknowns and ρ, γ are given constants. Despite the importance of this system of equations in fluid dynamics and the intensity with which it has been studied by many mathematicians and physicists, there are many fundamental open problems concerning behaviour of solutions; for example to this day it is not known whether a solution which is smooth in a neighbourhood of the initial hyperplane $t = 0$ can develop singularities (i.e. fails to extend to a smooth solution on all of $\mathbb{R}^4$).

19. The Harmonic Map System

$$\Delta_g u^\alpha + \sum_{i,j=1}^{n} g^{ij} \sum_{\beta,\gamma=1}^{N} \Gamma^\alpha_{\beta\gamma}(u) D_i u^\beta D_j u^\gamma = 0, \quad \alpha = 1, \ldots, N,$$

where $g = g_{ij}\, dx^i dx^j$ is a metric for $\mathbb{R}^n$, (g^{ij}) is the inverse matrix of (g_{ij}), Δ_g is the Laplacian with respect to the metric g; i.e.

$$\Delta_g f = \frac{1}{\sqrt{g}} \sum_{i,j} D_j(\sqrt{g} g^{ij} D_i f),$$

where here, somewhat ambiguously, g denotes $\det(g_{ij})$) and

$$\Gamma^\alpha_{\beta\gamma}(z) = \frac{1}{2}\sum_{\mu=1}^{n} (h^{\mu\alpha}(z)(h_{\mu\gamma,\beta}(z) + h_{\mu\beta,\gamma}(z) - h_{\beta\gamma,\mu}(z)))$$

(with $(h^{\mu\nu}) = (h_{\mu\nu})^{-1}$) are the "Christoffel symbols" for a metric

$$h_{\alpha\beta}(y) dy^\alpha dy^\beta \text{ of } \mathbb{R}^N, \, y = (y^1, \ldots, y^N).$$

Geometrically $u = (u^1, \ldots, u^N)$ represents a mapping from a domain in $\mathbb{R}^n$ into $\mathbb{R}^N$, and the equation says precisely that this map is harmonic. (Again the reader not familiar with the terminology should not be unduly concerned, as such terminology will not be used without explanation in the sequel.) The harmonic map system in fact provides a prototype for an interesting class of quasi-linear systems, concerning which much interesting work has been done in recent times; some of this will be described later in these lectures.

20. The Korteweg-de Vries Equation

$$u_t + cuu_x + u_{xxx} = 0$$

This equation arises in the study of water waves. It is the only equation of odd order in the present list.

We mentioned above that the minimal surface equation is the "Euler-Lagrange" equation for the area functional. We here explain the notion of Euler-Lagrange equation in a general context as follows. Suppose $F = F(x, z, p)$ is a given smooth real-valued function of the variables $(x, z, p) \in \Omega \times \mathbb{R} \times \mathbb{R}^n$, where Ω is a bounded domain in $\mathbb{R}^n$, and consider the functional

$$\mathcal{F}(u) = \int_\Omega F(x, u, Du)\, dx, \quad u \in C^1(\overline{\Omega}).$$

(We could equally well treat more general functionals, in which the integrand F depends on second or higher derivatives, or when u is a vector function $u = (u^1, \ldots, u^N)$.)

Consider now the possibility that u is a C^2 "stationary point," relative to smooth compactly supported deformations of u, for the functional $\mathcal{F}$; that is $u \in C^2(\overline{\Omega})$

$$\nabla\mathcal{F}(u)(\varphi)(\equiv \frac{d}{ds}\mathcal{F}(u + s\varphi)|_{s=0}) = 0 \ \ \forall \varphi \in C_c^\infty(\Omega).$$

By differentiation under the integral, we see that this is equivalent to

$$\int_\Omega \frac{d}{ds} F(x, u + s\varphi, Du + sD\varphi)|_{s=0}\, dx = 0 \ \ \forall \varphi \in C_c^\infty(\Omega).$$

That is, using the chain rule and letting partial derivatives of $F(x, z, p)$ be denoted by subscripts in the appropriate variable, we see

$$(*) \qquad \int_\Omega (F_z(x, u, Du)\varphi + \sum_{j=1}^n F_{p_j}(x, u, Du) D_j\varphi)\, dx = 0.$$

Since $u \in C^2(\Omega)$, we can integrate by parts in $(*)$, giving

$$\int_\Omega (F_z(x, u, Du) - \sum_{j=1}^n D_j[F_{p_j}(x, u, Du)])\varphi\, dx = 0.$$

Since $\varphi \in C_c^\infty(\Omega)$ is arbitrary, this is equivalent to the (pointwise) equation

$$(**) \qquad -\sum_{j=1}^n D_j[F_{p_j}(x, u, Du)] + F_z(x, u, Du) = 0.$$

Notice that by using the chain rule again, we can write this as

$$-\sum_{i,j=1}^n F_{p_i p_j} D_i D_j u + F_z - \sum_{j=1}^n (F_{p_j z} D_j u + F_{p_j x_j}) = 0,$$

where all partial derivatives of F are evaluated at $(x, u(x), Du(x))$.

These are thus all examples of quasilinear second order equations (linear in the second derivatives, with coefficients depending on u and its first partial derivatives). $(**)$ is actually called the Euler-Lagrange equation for the functional $\mathcal{F}$. If $u = (u^1, \ldots, u^N)$ is vector valued, straightforward modifications in the above argument lead to a system of N equations for the N unknowns $u^1, \ldots, u^N$, and this system is called the Euler-Lagrange system.

Notice that to make sense of the equation $(**)$ we need to know that u is of class C^2, whereas $(*)$ makes sense even if u is only of class C^1. The equation $(*)$ is referred to as the weak form of the Euler-Lagrange equation $(**)$.

Whether or not an equation arises as described above (as the Euler-Lagrange equation of a functional), it may nevertheless be true that important identities are obtained by integration. We illustrate this with two simple examples involving the wave equation and Laplace's equation:

First suppose that u is a $C^2(\overline{\Omega})$ solution of the wave equation $\frac{\partial^2 u}{\partial t^2} - \Delta u = 0$ subject to the boundary condition $u(x, t) \equiv 0$ for all $\xi \in \partial\Omega$ and all $t \geq 0$. We multiply both sides of the equation by $\dot{u}$ (where $\dot{u} = \partial u/\partial t$), and integrate over Ω. This gives

$$\int_\Omega (\dot{u}\ddot{u} - \dot{u}\Delta u)\, dx = 0,$$

and using integration by parts with respect to the x variables we thus get

$$\frac{1}{2}\int_\Omega \frac{d}{dt}\left(\dot{u}^2 + |Du|^2\right) dx = 0,$$

or in other words

$$\int_\Omega (\dot{u}^2 + |Du|^2)\, dx \equiv \text{ const.},$$

independent of t. The quantity $\frac{1}{2}\int(\dot{u}^2 + |Du|^2)$ being conserved here is called the energy—in applications to physics it really does represent the energy. The above conservation result (and related results) play an important role in the study of equations like the wave equation.

Our second example involves Laplace's equation $\Delta u = 0$ on a bounded connected domain $\Omega \subset \mathbf{R}^n$. We suppose that u is a $C^2(\overline{\Omega})$ solution of this equation with $u = 0$ on $\partial\Omega$. If we multiply by u and integrate over Ω we get

$$\int_\Omega u\Delta u\, dx = 0,$$

so that, after integration by parts,

$$\int_\Omega |Du|^2 = 0,$$

which shows that $u \equiv$ const. on Ω. Since $u = 0$ on $\partial\Omega$, this shows $u \equiv 0$ in all of $\overline{\Omega}$. This is an important result in the theory of Laplace's equation. (See for example Ex. 1.8 below for one application.) We shall see in Ex. 1.6 below the same result holds if we assume at the outset only that $u \in C^0(\overline{\Omega}) \cap C^2(\Omega)$ rather than $u \in C^2(\overline{\Omega})$ as we did above.

LECTURE 1 PROBLEMS

1.1. Prove the claim made under example 1 above.

1.2. (i) Check that each of the examples 2, 3, 7, 8, 11, 12, 13, 16, 17 is the Euler-Lagrange equation of some functional.

(ii) If $(a^{ij}(x))$ is not symmetric at some point x_0, prove that the equation in Ex. 3 is not the Euler-Lagrange equation for any functional $\mathcal{F}(u)$ of the form described in Lecture 1.

(iii) Check that the Example 9 is the Euler-Lagrange equation for the second order functional $\mathcal{F}(u) = \int_\Omega (\Delta u)^2$; show that it is also the Euler-Lagrange equation for the functional $\mathcal{F}(u) = \int_\Omega \sum_{i,j=1}^n (D_i D_j u)^2$.

1.3. Suppose that the function $F(x,z,p)$ is quadratic in the variables z, p (i.e. a homogeneous degree 2 polynomial in the variables $z, p_1, \ldots, p_n$ with coefficients which are functions of x). Prove that the Euler-Lagrange operator (L say) is linear and also self-adjoint in the sense that $(Lu,\varphi)_{L^2(\Omega)} = (u, L\varphi)_{L^2(\Omega)}$ for all $u \in C^2(\Omega)$, $\varphi \in C_c^2(\Omega)$.

1.4. (i) Consider modifying the discussion of Euler-Lagrange equations in Lecture 1 above to the case $F = F(x,\mathbf{z},\mathbf{p})$, where now $\mathbf{z} = (z^1,\ldots,z^N)$, $\mathbf{p} = (p_j^\alpha)_{\alpha=1,\ldots,N,\, j=1,\ldots,n}$ and $\mathcal{F}(u)$ is defined for vector functions $u = (u^1,\ldots,u^N)$. Modify the argument given to show that stationarity of u gives a <u>system</u> of N equations (called the Euler-Lagrange system).

(ii) In case $F(x,\mathbf{z},\mathbf{p}) = \sqrt{g}\sum_{i,j=1}^n \sum_{\alpha,\beta=1}^N g^{ij}(x) h_{\alpha\beta}(\mathbf{z}) p_i^\alpha p_j^\beta$, where the matrices $(g_{ij}(x))_{i,j=1,\ldots,n}$ and $(h_{\alpha\beta}(\mathbf{z}))_{\alpha,\beta=1,\ldots,N}$ are positive definite symmetric with smooth dependence on $x = (x^1,\ldots,x^n)$ and $\mathbf{z} = (z^1,\ldots,z^N)$ respectively, and where $g = \det(g_{ij})$, then the Euler-Lagrange system is the harmonic map system of Ex. 19 above.

1.5. In case $N = n \geq 3$ and we are in the setting of 1.4(i) above with $\mathcal{F}(u) = \int_\Omega |Du|^2(1+|u|^2)^{-2}$, prove that $u(x) = \dfrac{x}{|x|}$ is a solution of the Euler-Lagrange system on $\mathbb{R}^n \sim \{0\}$.

1.6. (i) If u is a $C^3(\Omega)$ solution of the Euler-Lagrange equation for the functional $\mathcal{F}(u) = \int_\Omega F(Du)\,dx$, where $F = F(p)$ is a smooth function of $p \in \mathbb{R}^n$ and if $\ell \in \{1,\ldots,n\}$, prove that $\varphi = D_\ell u$ satisfies the equation $D_i(a_{ij} D_j \varphi) = 0$, where $a_{ij}(x) = \frac{\partial^2 F}{\partial p_i \partial p_j}(Du(x))$, $x \in \Omega$.

(ii) Show that in case of the minimal surface equation (i.e. the case $\mathcal{F}(u) =$

$\int_\Omega \sqrt{1+|Du|^2}$), the result above gives

$$\sum_{i,j=1}^n D_i\big((1+|Du|^2)^{-1/2}(\delta_{ij} - \frac{D_iuD_ju}{1+|Du|^2})D_j\varphi\big) = 0.$$

1.7. Show that if Ω is a bounded domain in $\mathbb{R}^n$, if $u \in C^0(\overline{\Omega}) \cap C^2(\Omega)$ and if $\Delta u = 0$ in Ω and $u = 0$ on $\partial\Omega$, then $u \equiv 0$. (We proved this above in case $u \in C^2(\overline{\Omega})$.)

Hint: For $\varepsilon > 0$, $(u-\varepsilon)_+^{1+\varepsilon}$ is a $C^1(\overline{\Omega})$ function which vanishes on $\partial\Omega$.

1.8 Prove that there exists at most one $C^0(\overline{\Omega}) \cap C^2(\Omega)$ solution u of the problem $\Delta u = 0$ in Ω, $u = \varphi$ on $\partial\Omega$, where φ is a given continuous function on $\partial\Omega$.

1.9 If $T > 0$ is given, and if u is a $C^2(\overline{\Omega} \times [0,T])$ solution of the heat equation $\dot u = \Delta u$ on $\Omega \times [0,T]$ with $u(x,t) = 0$ for all $(x,t) \in \partial\Omega \times [0,T)$, prove that $\int_\Omega(\frac{1}{2}u^2(x,t) + \int_0^\tau |Du(x,\tau)|^2\,d\tau)\,dx$ is constant for $t \in [0,T]$.

Lecture 2

Some Basic Remarks on Solvability

Initial Remarks on Local Solvability: Lewy's Example

Consider the first order equations

$$(1) \qquad u_x + \ xu_y = f(x,y)$$
$$(2) \qquad u_x + ixu_y = f(x,y)$$

where $u = u_1 + iu_2$, $f = f_1 + if_2$ and where $f_1,\ f_2$ are given smooth functions. These are similar looking equations with vastly different local solvability properties: Viz., (1) is an example in the class of equations of the general form $a_1(x,y)u_x + a_2(x,y)u_y = f$, where a_1, a_2 are <u>real</u>, and as such there is a general method of solution (see later remarks on first order equations) giving solutions

$$(*) \qquad u(x,y) = \phi(y - \frac{x^2}{2}) + \int_0^x f(\tau, \frac{\tau^2}{2} + y - \frac{x^2}{2})\,d\tau$$

where ϕ is an arbitrary C^1 function on $\mathbb{R}$. On the other hand we claim that there are C^∞ choices of f (in fact many choices) such that (2) has no C^1 solutions defined in a neighbourhood of the origin. Indeed we claim this is so if $f(x,y)$ is even in the x-variable (i.e. $f(-x,y) \equiv f(x,y)$) and if there is a sequence $\rho_j \downarrow 0$ with $f \equiv 0$ in a neighbourhood of C_{ρ_j} and $\int_{D_{\rho_j}} f(x,y)\,dxdy \neq 0$. Here D_ρ is the disc of radius ρ and C_ρ its boundary.

The proof of this claim is a modification (due to Nirenberg) of an argument of H. Lewy. Assume (for contradiction) that u is a C^1 solution of (2) on the disc $D_{\rho_0} = \{x^2 + y^2 < \rho_0^2\}$ where $\rho_0 > 0$. First, we may assume without loss of generality that

(i) $$u(x, y) = -u(-x, y)$$

(Otherwise replace u by $w(x, y) = \frac{1}{2}(u(x, y) - u(-x, y))$ and note that w satisfies $w_x + ixw_y = f$, because $f(-x, y) = f(x, y)$. Thus we may assume (i). Next we want to establish

$$u \equiv 0 \text{ on } C_{\rho_j}$$

(for j large enough so that $\rho_j < \rho_0$). To see this select an annulus A (centred at 0) such that $C_{\rho_j} \subset$ interior A and $f \equiv 0$ on A. Make the transformation of variables

$$(x, y) \in A \mapsto \begin{cases} (\ \frac{1}{2}x^2, y) \text{ if } x \geq 0 \\ (-\frac{1}{2}x^2, y) \text{ if } x < 0 \end{cases}$$

Notice that the transformation of variables is a homeomorphism of A onto some open subset $\widetilde{A} \subset \mathbb{R}^2$. Let

$$A_+ = \{(x, y) \in A : x > 0\}, \qquad \widetilde{A}_+ = \{(s, y) \in \widetilde{A} : s > 0\},$$

and

$$\widetilde{u}(s, y) = u(x, y), \ (x, y) \in A_+, \ s = \tfrac{1}{2}x^2.$$

Notice that then

(ii) $$\widetilde{u}_s + i\widetilde{u}_y = 0 \text{ in } \widetilde{A}_+, \quad \lim_{s \to 0, (s,y) \in \widetilde{A}_+} \widetilde{u}(s, y) = 0,$$

the latter provided $(0, y) \in A$.

Select a $\frac{1}{2}$-disc D_+ with centre on the y-axis and $D_+ \subset \widetilde{A}_+$. By (ii) we know that $\widetilde{u}$ is holomorphic in $\widetilde{A}_+$, hence in D_+, and that $\lim_{s \to 0, (s,y) \in D_+} \widetilde{u}(s, y) = 0$. It is standard that then $\widetilde{u} \equiv 0$ in D_+.

Then by unique continuation $\widetilde{u} \equiv 0$ in $\widetilde{A}_+$ and hence (using (i)) $u \equiv 0$ on C_{ρ_j}. But then using the divergence theorem we have

$$0 \neq \int_{D_{\rho_j}} f = \int_{D_{\rho_j}} (u_x + ixu_y) = \int_{C_{\rho_j}} (\frac{x}{\rho_j} + \frac{ixy}{\rho_j})u = 0,$$

a contradiction.

Actually Lewy's original example is even more striking in that, although it is in 3 independent variables x, y, z, it has the property that there are no C^1 solutions defined in any open subset of $\mathbb{R}^3$. Lewy's example is

$$u_x + iu_y - 2i(x + iy)u_z = f,$$

where again $u = u_1 + iu_2$ and $f = f_1 + if_2$. (See for example [JF, Chapter 8] for a description of suitable C^∞ functions f to put on the right side in this case.)

Existence of weak solutions

We here want to discuss a general condition for the existence of weak solutions. First it is necessary to introduce the notion of weak solution in L^p:

Let $Lu = \sum_{|\alpha| \leq m} a_\alpha(x) D^\alpha u$, $a_\alpha \in C^\infty(\Omega)$ (multi-index notation as in the introduction). The adjoint operator $L' : C^\infty(\Omega) \to C^\infty(\Omega)$ is characterized by

$$(L\phi, \psi)_{L^2(\Omega)} = (\phi, L'\psi)_{L^2(\Omega)}, \phi, \psi \in C_c^\infty(\Omega),$$

and is given explicitly by $L'u = \sum_{|\alpha| \leq m} (-1)^{|\alpha|} D^\alpha(a_\alpha u)$

Note: If we allow a_α, ϕ, ψ, u to be complex-valued (rather than real-valued) then $(\phi, \psi)_{L^2(\Omega)} = \int_\Omega \phi \overline{\psi}\, dx$ and L' is given by

$$L'\psi = \sum_{|\alpha| \leq m} (-1)^{|\alpha|} D^\alpha(\overline{a}_\alpha \psi), \ \psi \in C^\infty(\Omega),$$

where $\overline{\psi}, \overline{a}_\alpha$ denote complex conjugates of the respective quantities.

The following definition is of fundamental importance for a significant portion of the remaining lectures:

Definition: Given $f \in L^1_{\text{loc}}(\Omega)$ we say that u is an L^p weak solution of $Lu = f$ in Ω (where $p \geq 1$) if $u \in L^p_{\text{loc}}(\Omega)$ and

$$(u, L'\psi)_{L^2(\Omega)} = (f, \psi)_{L^2(\Omega)} \ \forall \psi \in C_c^\infty(\Omega).$$

Remark. In the important case when the operator L has constant coefficients, i.e.

$$Lu \equiv \sum_{|\alpha| \leq m} a_\alpha D^\alpha u, \text{ where } a_\alpha \text{ are constants},$$

there is a nice way of interpreting weak solutions in terms of classical solutions as follows:

Take a fixed $C_c^\infty(\mathbb{R}^n)$ function φ with $\varphi \ge 0$, support $\varphi \subset B_1(0) \equiv \{x : |x| < 1\}$, $\int_{\mathbb{R}^n} \varphi(x)\,dx = 1$, and let $\varphi^{(\sigma)}(x) = \sigma^{-n}\varphi(\sigma^{-1}x)$ for $\sigma > 0$. Notice that then

$$\text{support}\,\varphi^{(\sigma)} \subset B_\sigma(0) \text{ and } \int_{\mathbb{R}^n} \varphi^{(\sigma)}(x)\,dx = 1 \ \ \forall \sigma > 0,$$

and $\varphi^{(\sigma)}$ approaches the "Dirac delta function" as $\sigma \downarrow 0$, in the sense that if $u \in L^1_{\text{loc}}(\Omega)$, and if

$$(*) \qquad u_\sigma(x) \equiv \int_\Omega \varphi^{(\sigma)}(x-y)u(y)\,dy (\equiv (\varphi^{(\sigma)} * u)(x)), \ \ x \in \Omega_\sigma,$$

where $\Omega_\sigma = \{x \in \Omega : \text{dist}\,(x, \partial\Omega) > \sigma\}$ and where $f*g$ denotes convolution—$f * g(x) = \int_\Omega f(x-y)g(y)\,dy$ (valid when support of $f \subset B_\sigma(0)$ and $\text{dist}\,(x, \partial\Omega) \ge \sigma$), then

$$(**) \qquad u_\sigma(x) \to u(x) \text{ as } \sigma \downarrow 0$$

a.e. $x \in \Omega$. To see this we note simply that since $\int \varphi^{(\sigma)} = 1$, we have

$$u_\sigma(x) - u(x) = -\int_{\mathbb{R}^n} \varphi^{(\sigma)}(x-y)(u(x) - u(y))\,dy,$$

so that (since $|\varphi^{(\sigma)}| \le c\sigma^{-n}$ and support $\varphi^{(\sigma)} \subset B_\sigma(0)$)

$$|u_\sigma(x) - u(x)| \le c\sigma^{-n} \int_{B_\sigma(x)} |u(x) - u(y)|\,dy;$$

hence we in fact conclude that in particular $u_\sigma(x) \to u(x)$ at each Lebesgue point $x \in \Omega$ of the function u.

We also note that, for each $\sigma > 0$,

$$(***) \qquad u_\sigma \in C^\infty(\Omega_\sigma);$$

indeed by differentiation under the integral we have $D^\alpha u_\sigma = (D^\alpha \varphi^{(\sigma)}) * u$ on Ω_σ for each multi-index α.

In view of $(**)$ and $(***)$ we call u_σ a "smoothing" or "regularization" or "mollification" of u. Here the important point we want to emphasize is the following:

Theorem. *If $u \in L^1_{\text{loc}}(\Omega)$ is a weak solution of $Lu = f$, where $f \in L^1_{\text{loc}}(\Omega)$ and the coefficients a_α of L are constants, then for each $\sigma > 0$ $Lu_\sigma = f_\sigma$ in Ω_σ. That is, u_σ is a classical solution of the equation on Ω_σ for each $\sigma > 0$.*

Proof. If $\sigma > 0$ and $x \in \Omega_\sigma$ is fixed, then the function $\psi_x(y) = \varphi^{(\sigma)}(x-y)$, $y \in \Omega$, is a $C_c^\infty(\Omega)$ function of y, hence by definition of weak solution we have $(L'\psi_x, u)_{L^2(\Omega)} = 0$, or in other words

$$\sum_{|\alpha| \le m} (-1)^{|\alpha|} a_\alpha \int_\Omega D_y^\alpha [\varphi^{(\sigma)}(x-y)] u(y)\,dy = 0.$$

Since $(-1)^{|\alpha|} D_y^\alpha[\varphi^{(\sigma)}(x-y)] = D_x^\alpha[\varphi^{(\sigma)}(x-y)]$ (by the chain rule), the differentiation under the integral formula then yields $Lu_\sigma(x) = 0$ as required.

We now want to derive a useful necessary and sufficient condition for the existence of L^2 weak solutions of $Lu = f$ in case $f \in L^2(\Omega)$ (where Ω continues to denote an arbitrary open subset of $\mathbb{R}^n$):

Theorem. *If $f \in L^2(\Omega)$, then there is an L^2 weak solution of $Lu = f$ if and only if there is a constant c such that*

$$|\langle f, \varphi \rangle_{L^2(\Omega)}| \le c\|L'\varphi\|_{L^2(\Omega)} \ \forall \varphi \in C_c^\infty(\Omega).$$

Proof: (i) "$\Rightarrow$": Given $u \in L^2(\Omega)$ with $(u, L'\varphi) = (f, \varphi) \forall \varphi \in C_c^\infty(\Omega)$. By the Schwarz inequality

$$|\langle u, L'\varphi \rangle| \le \|u\|\ \|L'\varphi\|,$$

and hence "$\Rightarrow$" holds with $c = \|u\|$.

(ii)"$\Leftarrow$": Assume

$$(*) \qquad |\langle f, \varphi \rangle| \le c\|L'\varphi\|.$$

Let $S = \{L'\varphi : \varphi \in C_c^\infty(\Omega)\}$. Clearly S is a linear subspace of $L^2(\Omega)$. We define a linear functional T on S by $T(L'\varphi) = (f, \varphi), \varphi \in C_c^\infty(\Omega)$. Notice that this is well-defined by $(*)$; also by $(*)$ we have

$$|T(L'\varphi)| \le c\|L'\varphi\| \ \forall \varphi \in C_c^\infty(\Omega)$$

so that

$$|T(w)| \le c\|w\| \ \forall w \in S.$$

Then by the Hahn-Banach theorem T extends to give a functional $\overline{T}$ on all of $L^2(\Omega)$ with

$$|\overline{T}(w)| \le c\|w\| \ \forall w \in L^2(\Omega).$$

By the Riesz representation theorem there is a $u \in L^2(\Omega)$ with

$$\overline{T}(w) \equiv (u, w)_{L^2(\Omega)} \ \forall w \in L^2(\Omega).$$

In particular, specializing to $w \in S$, we have

$$\langle u, L'\varphi \rangle = \langle f, \varphi \rangle \ \forall \varphi \in C_c^\infty(\Omega),$$

as required.

Remark: Since $|\langle f, \varphi\rangle| \le \|f\|\ \|\varphi\|$ we see that there is weak L^2 solution of $Lu = f$ for arbitrary $f \in L^2(\Omega)$ provided there is a constant c such that

$$\|\varphi\| \le c\|L'\varphi\| \ \forall \varphi \in C_c^\infty(\Omega).$$

Local solvability of the Cauchy problem for first order equations

Next we want to discuss the local solvability theory for quasilinear first order equations. Specifically, consider the equation

$$\textstyle\sum_{i=1}^n a_i(x,u) D_i u = f(x,u),$$

where a_i, f are smooth real-valued, subject to "Cauchy data"

$$u(x) = \varphi(x) \quad \text{on the hyperplane } x^n = 0,$$

where φ is a given smooth real-valued function on $\mathbb{R}^{n-1}$.

Notice that if u is a C^1 solution on a domain U and $x_0 \in U$, and if $x(t)$ is a solution of the non-linear ODE system $\dot{x} = a(x, u(x))$ on some interval of t, and if we define $z(t) = u(x(t))$, then by the chain rule we have

$$\dot{z} = Du(x(t)) \cdot \dot{x} = a(x(t), u(x(t))) \cdot Du(x(t)) \equiv f(x(t), u(x(t))).$$

Thus $(x(t), z(t))$ is a solution of the ODE system

$$(*) \qquad \begin{cases} \dot{x} = a(x,z) \\ \dot{z} = f(x,z). \end{cases}$$

This suggests that we try to find solutions of the PDE by putting together solution curves of the ODE system to give the graph of the PDE solution.

So for a given function $\psi \in C^1(\mathbb{R}^{n-1} \times \{0\})$, let $X(s,t), Z(s,t)$ be the solution of $(*)$ with initial conditions

$$X(s,0) = (s,0)$$
$$Z(s,0) = \varphi(s,0).$$

Recall (from the relevant ODE theory) that there is a neighbourhood U of $t = 0$ in (s,t)-space such that X, Z are smooth functions of s, t in U. Thus $X : U \to \mathbb{R}^n, Z : U \to \mathbb{R}$ are smooth. We can invert the equation $x = X(s,t)$ locally near the point $x_0 = (s_0, 0)$ provided the Jacobian is non-zero; checking, one finds the Jacobian matrix has e_i in the i^{th} row for $i = 1, \dots, n-1$ and n^{th} row $(a_1(x_0), \dots, a_n(x_0))$, so we can invert in a ball $B_\rho(x_0)$ if ρ is sufficiently small, provided $a_n(x_0) \ne 0$. (Notice the condition $a_n(x_0) \ne 0$ means geometrically that a is not tangent to the hyperplane $x^n = 0$ at x_0 , thus ensuring the integral curves $x_s(t)$ meet $x^n = 0$ transversely at $t = 0$ for s near s_0 .) Thus we have a local inverse $(s,t) = (S(x), T(x))$ in a neighbourhood of x_0; specifically we can locally invert in some ball $B_\rho(x_0)$ of (s,t)-space, and then $u \equiv Z(S(x), T(x))$ is defined (and smooth) in a neighbourhood W of x_0 in x-space.

Let us check now that u is a solution of the equation in W:

By construction $z_s(t) = u(x_s(t))$ and $x_s(t) \in W$ for $(s,t) \in B_{\rho_0}(x_0)$. Differentiate and use * together with the chain-rule:

$$\begin{aligned} f(x_s(t), z_s(t)) = \dot{z}_s(t) &= Du(x_s(t)) \cdot \dot{x}_s(t) \\ &= a(x_s(t), u(x_s(t))) \cdot Du(x_s(t)), \end{aligned}$$

so that

$$a(x_s(t), u(x_s(t))) \cdot Du(x_s(t)) = f(x_s(t), u(x_s(t)));$$

that is, the equation is satisfied at the point $(x_s(t), z_s(t))$ of W. Notice also that $u(x^1, \ldots, x^{n-1}, 0) \in W$, because $z_s(0) = \phi(x^1, \ldots, x^{n-1})$ for $(s, 0) \in B_{\rho_0}(x_0)$.

Finally u is the unique solution in W: if $\widetilde{u}$ also satisfies the equation in W and $\widetilde{u}(x^1, \ldots, x^{n-1}, 0) = \phi(x^1, \ldots, x^{n-1})$ for $(x^1, \ldots, x^{n-1}, 0) \in W$, then $\widetilde{u} = u$ in W. Actually we have the more general uniqueness property:

If $\widetilde{W} \subset W$ is the image under S of an open $U \subset B_{\rho_0}(x_0)$ with the property $(s_1, t_1) \in U \Rightarrow (s_1, t) \in U$ for all t between 0 and t_1 inclusive, then u is the unique C^1 solution of the equation in $\widetilde{W}$ satisfying $u(x^1, \ldots, x^{n-1}, 0) = \phi(x^1, \ldots, x^{n-1})$, for $(x^1, \ldots, x^{n-1}, 0) \in \widetilde{W}$.

Using the above uniqueness we can now piece together local solutions u of $(*)$, $(**)$ in a neighbourhood of $\{(x^1, \ldots, x^{n-1}, 0) : a_n(x^1, \ldots, x^{n-1}, 0) \neq 0\}$.

Remark: There is also a method (of which the above is a special case) for solving the Cauchy problem for non-linear first order equations with n independent variables $x^1, \ldots, x^n$. (See e.g. [JF] for discussion.) We emphasize that these methods work only for single first-order equations for the real unknown u; they do not extend to systems (for the good reason that no such general existence result holds for systems, as shown by our discussion of equation (2) above).

LECTURE 2 PROBLEMS

2.1. Check by direct computation that u as given in $(*)$ on p. 1 does satisfy the equation (1).

2.2 Show that there is a C^∞ function f such that the equation

$$\Big(\big(\frac{\partial^2}{\partial x^2} + x^2 \frac{\partial^2}{\partial y^2}\big)^2 + \frac{\partial^2}{\partial y^2}\Big) u = f$$

has no C^4 solutions in any neighbourhood of 0.

Hint: Use the non-existence result for $u_x + ixu_y = f$ established above.

2.3. Let $\Omega = (0,1) \subset \mathbb{R}$. Show that (after redefinition on a set of measure zero in each case) that:

(i) Any weak solution (in $L^1_{\mathrm{loc}}(\Omega)$) of $\frac{dy}{dx} = 0$ in Ω is also a classical solution (i.e. $y \equiv$ const.)

(ii) Any weak solution of $\frac{d^2y}{dx^2} = 1$ in Ω is also a classical solution $y \equiv \frac{1}{2}x^2 + cx + d$, $(c, d$ constants).

2.4. If $\Omega = (0,1) \times (0,1) \subset \mathbb{R}^2$, show that any weak solution in $L^1_{\mathrm{loc}}(\Omega)$ of $\frac{\partial^2 u}{\partial x \partial y} = 0$ in Ω has the form $u(x,y) = \varphi(x) + \psi(y)$ with φ, $\psi \in L^1_{\mathrm{loc}}(0,1)$.

Conversely, show that any such function is a weak solution of $\frac{\partial^2 u}{\partial x \partial y} = 0$ on Ω.

2.5. Suppose $L = L_1 + iL_2$, where $L_j = a_j(x,y)\frac{\partial}{\partial x} + b_j(x,y)\frac{\partial}{\partial y}$ with a_j, b_j real $C^\infty(\mathbb{R}^2)$ functions.

Prove that for every (complex-valued) $f \in L^2_{\mathrm{loc}}(\mathbb{R}^2)$, there is a weak solution of $Lu = f$ in some neighbourhood of $(x,y) = (0,0)$, provided that

(i), $\quad L \neq 0$ i.e. at least one of a_1, b_1, a_2, b_2 is $\neq 0$ at $(0,0)$

and

(ii) $\quad [L_1, L_2] = c(x,y)L_1 + d(x,y)L_2$

in some neighbourhood U of $(0,0)$, where $c, d \in C^\infty(U)$ and $[L_1, L_2] \equiv L_1L_2 - L_2L_1$.

Hint: We may assume $a_1(0,0) \neq 0$. Use the existence criterion proved in Lecture 2 above; also note that

$$\int_{\mathbb{R}^2} L_1'(x\varphi^2) = 0, \quad \varphi \in C_c^\infty(\mathbb{R}^2),$$

and hence $\int \varphi^2 \le C\rho^2 \int (L_1\varphi)^2$ provided support of $\varphi \subset \{(x,y) \in \mathbb{R}^2 : |x| < \rho\}$ with ρ sufficiently small.

2.6. Show that the method described in Lecture 2 above for solving first-order equations leads to the solution $(*)$ on p. 1.

2.7. Show that the problem

$$\begin{cases} xu_x - yu_y = 0 \text{ in some neighbourhood of } y = 0 \\ \quad u(x,0) = \varphi(x), \quad x \in \mathbb{R} \end{cases}$$

($\varphi \in C^1(\mathbb{R})$ given) is not solvable with $u \in C^1$ unless $\varphi \equiv$ const.

If $\varphi \equiv$ const., prove there are infinitely many solutions defined on all of $\mathbb{R}^2$.

Lecture 3

Solvability (continued): The Cauchy Kovalevski Theorem

Consider the general m^{th}-order Cauchy problem

$$(*) \qquad D_t^m u_j = F_j(x, t, \{D_x^\alpha D_t^\beta u_k\}_{|\alpha|+\beta \le m, \beta \le m-1, k=1,\dots,N}), \quad j = 1,\dots,N$$

in a neighbourhood of $t = 0$ subject to "Cauchy Data"

$$(**) \qquad \begin{aligned} u_j(x,0) &= \phi_j^0(x) \\ D_t u_j(x,0) &= \phi_j^1(x) \\ &\dots \\ D_t^{m-1} u_j(x,0) &= \phi_j^{m-1}(x) \end{aligned}$$

where F_j, ϕ_j^k are given smooth functions.

We want to show that there is a real-analytic solution of this if F_j, ϕ_j^k are real analytic functions of their arguments (this is the Cauchy-Kovalevski Theorem). Recall that a function $f(\xi)$ with $\xi = (\xi^1, \dots, \xi^N) \in \mathbb{R}^N$ is said to be real analytic near ξ_0 if it has an absolutely convergent power series expansion near ξ_0. In multi-index notation this means that there is a $\rho > 0$ such that we can express $f(\xi)$ in terms an absolutely convergent power series expansion thus:

$$f(\xi) = \sum_{|\alpha|=0}^{\infty} c_\alpha (\xi - \xi_0)^\alpha, \quad 0 \le |\xi - \xi_0| < \rho,$$

where for any multi-index $\alpha = (\alpha_1, \dots, \alpha_N)$ we take

$$(\xi - \xi_0)^\alpha = (\xi^1 - \xi_0^1)^{\alpha_1} \cdot (\xi^2 - \xi_0^2)^{\alpha_2} \dots (\xi^N - \xi_0^N)^{\alpha_N}.$$

Whenever this is so, the coefficients c_α are given by

$$c_\alpha = \frac{D^\alpha f(\xi_0)}{\alpha!},$$

where $\alpha! = (\alpha_1!)\cdot(\alpha_2!)\cdots(\alpha_N!)$. Furthermore a necessary and sufficient condition that such a power series expansion holds near ξ_0 is that $\exists$ constants $M, \rho > 0$ such that

$$\sup_{|\xi-\xi_0|<\rho} |D^\alpha f(\xi)| \le M^{|\alpha|+1}\alpha! \ \forall \text{ multi-index } \alpha.$$

The reader can check these facts by using the theory of Taylor series for functions of 1 variable on the function $g(t) = f(\xi_0 + t(\xi - \xi_0))$ (ξ fixed).

In order to prove Cauchy-Kovalevski we first have to reduce to a 1st-order system with zero Cauchy data. Reduction to an equivalent system with zero Cauchy data is easy by change:

$$\begin{aligned} &u_j \mapsto \widetilde{u}_j = u_j - \textstyle\sum_{k=0}^{m-1} \frac{t^k}{k!}\phi_j^k(x) \\ &F \mapsto \widetilde{F},\ \ \widetilde{F}(x,t,\{p_k^{\alpha\beta}\}) = F(x,t,\{p_k^{\alpha\beta} - D_x^\alpha \varphi_k^\beta(x)\}). \end{aligned}$$

(Notice $\widetilde{F}$ is still real-analytic in its arguments.)

Thus we can, and we shall, assume without loss of generality that $\varphi_j^k \equiv 0$ in $(**)$.

Next we show that we can reduce the order of the equation in two steps:

Step 1: Reduction to an equivalent order $m-1$ system (in case $m \ge 2$). To see how this step works, begin by assuming we have a solution u of $(*)$, $(**)$ (with $\varphi_j^k \equiv 0$) and define

$$v_j = D_t u_j,\ \ v_{jk} = D_{x^k} u_j,$$

and note that u_j, v_j, v_{jk} satisfy the order $m-1$ system

$$(*)' \quad \begin{cases} \text{(i)}\ D_t^{m-1}u_j = D_t^{m-2}v_j \\ \text{(ii)}\ D_t^{m-1}v_{jk} = D_t^{m-2}D_{x^k}v_j \\ \text{(iii)}\ D_t^{m-1}v_j = \hat{F}_j(x,t,\{D_x^\alpha D_t^\beta(u_j,v_j,v_{jk})\}_{|\alpha|+\beta\le m-1,\,\beta\le m-2}), \end{cases}$$

where $\hat{F}_j$ is still real analytic in its arguments,

$$(**)' \quad \begin{cases} \text{(i)}\ u_j(x,0) = D_t u_j(x,0) = \cdots = D_t^{m-2}u_j(x,0) = 0 \\ \text{(ii)}\ v_j(x,0) = D_t v_j(x,0) = \cdots = D_t^{m-2}v_j(x,0) = 0 \\ \text{(iii)}\ v_{jk}(x,0) = D_t v_{jk}(x,0) = \cdots = D_t^{m-2}v_{jk}(x,0) = 0, \end{cases}$$

i.e. zero Cauchy data still.

Now this new system is equivalent to the old, in that if u_j, v_j, v_{jk} satisfy $(*)'$, $(**)'$ then automatically $v_j = D_t u_j$, $v_{jk} = D_{x^k}u_j$, and u_j satisfies the original system $(*)$, $(**)$ (with $\varphi_j^k \equiv 0$). (We see $v_j = D_t u_j$, $v_{jk} = D_{x^k}u_j$ by integrating $m-2$ times from 0 to t in $(*)'$(i), (ii) using $(**)'$.)

Step 2: By Step 1 and induction on m we can reduce to an equivalent system of first-order equations of the form

$$\begin{aligned} D_t U_j &= f_j(x,t,U,D_xU) && (*)'' \\ U_j(x,0) &= 0, && (**)'' \end{aligned}$$

(where f_j is real-analytic in its arguments), such that if U satisfies $(*)''$, $(**)''$ then its first N components $U_1,\ldots,U_N$ satisfy the original system $(*)$, $(**)$.

We now claim that we can make a further reduction so that the above can be taken to be a quasilinear system; i.e.

$$f_j(x,t,U,D_xU) = \textstyle\sum_{\ell=1}^M \sum_{i=1}^n f_j^i(x,U)D_{x^i}U_\ell + g_j(x,U),\ \ j = 1,\ldots,M,$$

where f_j^i, g_j are real analytic in their arguments. To see this suppose for a moment that U satisfies $(*)''$, $(**)''$ and differentiate $(*)''$ with respect to t. Writing $f_j = f_j(x,t,z,p)$, $f_{jz_k} = \partial f_j(x,t,z,p)/\partial z^k$ etc., this gives

$$D_t^2U_j = f_{jt}(x,t,U,D_xU) + \textstyle\sum_k f_{jz^k}(x,t,U,D_xU)D_tU_k + \sum_{k,\ell} f_{jp_{k\ell}}(x,t,U,D_xU)D_tD_{x^k}U_\ell.$$

Thus, with

$$Q_j = D_tU_j - f_j(x,0,0,0),\ \ R_{ij} = D_{x^i}U_j,$$

we have a new system

$$\begin{cases} D_tU_j = Q_j + f_j(x,0,0,0) \\ D_tQ_j = f_{jt}(x,t,U,R) + \sum_k f_{jz^k}(x,t,U,R)Q_k + \\ \qquad\qquad \sum_{k,\ell} f_{jp_{k\ell}}(x,t,U,R)D_{x^k}Q_\ell \\ D_tR_{ij} = D_{x^i}Q_j + D_{x^i}f_j(x,0,0,0) \end{cases}$$

subject to Cauchy data

$$\begin{cases} U_j(x,0) = 0 \\ Q_j(x,0) = 0 \\ R_{ij}(x,0) = 0\,. \end{cases}$$

Furthermore it is easy to check that if U, Q, R satisfy this then automatically $Q_j = D_tU_j - f_j(x,0,0,0)$, $R_{ij} = D_{x^i}U_j$, and U satisfies $(*)''$, $(**)''$. Thus the new system is equivalent to the old. Finally we eliminate the t-dependence on the right of the equations for $D_t Q_j$ by introducing one further unknown $S = t$, with equation $\partial S/\partial t \equiv 1$ and Cauchy data $S(x,0) = 0$.

Thus we have finally reduced to an equivalent system of the form

$$\begin{cases} D_t U_j = \sum_{\ell=1}^{M} \sum_{i=1}^{n} f_j^{i\ell}(x,U) D_{x^i} U_\ell + g_j(x,U) \\ U_j(x,0) = 0, \end{cases}$$

for $j = 1, \dots, M$, where $M > N$, with the property that if U satisfies this, then the first N components $u = (U_1, \dots, U_N)$ satisfy the original problem $(*)$, $(**)$ (with $\varphi_j^k \equiv 0$).

Thus from now on we may without loss of generality consider the first-order quasilinear problem

$$\text{(i)} \qquad D_t u_j = \textstyle\sum_{i,\ell} f_j^{i\ell}(x,u) D_{x^i} u_\ell + g_\ell(x,u),$$

$$\text{(ii)} \qquad u_j(x,0) = 0,$$

$j = 1, \dots, N$, where $f_j^{i\ell}$, g_j are given real-analytic functions.

We first want to show that this problem has a real-analytic solution in a neighbourhood of $(x,t) = (0,0)$. We have (for some $\rho > 0$ and for $|x| < \rho$, $|u| < \rho$)

$$f_j^{k\ell}(x,u) = \textstyle\sum_{\alpha,\beta} a_{j\alpha\beta}^{k\ell} x^\alpha u^\beta$$
$$g_j(x,u) = \textstyle\sum_{\alpha,\beta} b_{j\alpha\beta} x^\alpha u^\beta$$

where

$$a_{j\alpha\beta}^{i\ell} = \frac{D_x^\alpha D_u^\beta f_j^{i\ell}(0,0)}{\alpha!\,\beta!}, \quad b_{j\alpha\beta} = \frac{D_x^\alpha D_u^\beta g_j(0,0)}{\alpha!\,\beta!}.$$

Notice that if u is a solution of (i), (ii) near $(x,t) = (0,0)$, then using the above expansions we can compute all derivatives $D_x^\alpha D_t^\delta u_j(0,0)$ in terms of $a_{k\alpha\beta}^{i\ell}$, $b_{k\alpha\beta}$ as follows:

First (ii) can be differentiated with respect to the x-variables to give

$$D_x^\gamma u_j(x,0) = 0 \;\; \forall \text{ multi-index } \gamma.$$

Also, by applying $D_x^\gamma D_t^{\delta-1}$ in (i) for $\delta > 0$, we can compute $D_x^\gamma D_t^\delta u_j(0,0)$ in terms of $a_{k\alpha\beta}^{i\ell}$, $b_{k\alpha\beta}$ and the functions $D_x^\gamma D_t^q u_\ell(0,0)$, $0 \le q \le \delta - 1$. Then, by induction on δ (with some fixed bound on $|\gamma|$), we have all derivatives $D_x^\gamma D_t^\delta u_j(0,0)$ in terms of the known constants $a_{j\alpha\beta}^{i\ell}$, $b_{j\alpha\beta}$. As a matter of

fact, since we only use the product and sum rule for differentiation on these computations, we actually get

$$\text{(1)} \qquad D_x^\gamma D_t^\delta u_j(0,0) = P_{\gamma\delta j}(\{(a_{m\alpha\beta}^{i\ell}, b_{m\alpha\beta})\}_{|\alpha|+|\beta| \le |\gamma|+\delta}),$$

where $P_{\gamma\delta j}$ is a polynomial with non-negative <u>integer</u> coefficients in the indicated variables. (This is formally checked as part of the above induction of δ.) Thus we have (assuming that u is a real-analytic solution)

$$u_j(x,t) = \sum \frac{1}{\gamma!\,\delta!} P_{\gamma\delta j}(\{(a_{m\alpha\beta}^{i\ell}, b_{m\alpha\beta})\}_{|\alpha|+|\beta| \le |\gamma|+\delta}) x^\gamma t^\delta,$$

where $P_{\gamma\delta j}(\{(a_{m\alpha\beta}^{i\ell}, b_{m\alpha\beta})\}_{|\alpha|+|\beta| \le |\gamma|+\delta})$ is a polynomial in the indicated variables with non-negative integer coefficients. Also if we take u to be <u>defined</u> by this identity then, provided the series on the right converges, it will give a solution of the problem (i), (ii) near $(x,t) = (0,0)$, provided the series on the right converges in some neigbourhood of $(x,t) = (0,0)$.

Thus we have only to prove that the series on the right converges near $(x,t) = (0,0)$. For this we use the "method of majorants." This begins with the following observation:

If U is a real analytic solution of a new system

$$D_t U_j = \textstyle\sum_{i,\ell} F_j^{i\ell}(x,U) D_{x^i} U_\ell + G_j(x,U), \quad j = 1, \dots, N \qquad \text{(i)}'$$
$$U(x,0) = 0, \qquad \text{(ii)}$$

where

$$F_j^{i\ell}(x,U) = \textstyle\sum_{\alpha,\beta} A_{j\alpha\beta}^{i\ell} x^\alpha U^b, \quad G_j(x,U) = \textstyle\sum_{\alpha,\beta} B_{j\alpha\beta} x^\alpha U^\beta$$

then by the same argument as above, with $A_{m\alpha\beta}^{i\ell}$, $B_{m\alpha\beta}$ in place of $a_{m\alpha\beta}^{i\ell}$, $b_{m\alpha\beta}$ respectively, we would get

$$D_x^\gamma D_t^\delta U_j(0,0) = P_{\gamma\delta j}(\{(A_{m\alpha\beta}^{i\ell}, B_{m\alpha\beta})\}_{|\alpha|+\beta \le |\gamma|+\delta}),$$

with $P_{\gamma\delta j}$ the <u>same</u> polynomial as in (1) above.

Since $P_{\gamma\delta j}$ has positive coefficients we would then evidently have

$$\text{(2)} \qquad |P_{\gamma\delta j}(\{(a_{m\alpha\beta}^{i\ell}, b_{m\alpha\beta})\})| \le P_{\gamma\delta j}(\{(A_{m\alpha\beta}^{i\ell}, B_{m\alpha\beta})\})$$

provided

$$\begin{cases} |a_{m\alpha\beta}^{i\ell}| \le A_{m\alpha\beta}^{i\ell} \\ |b_{m\alpha\beta}| \le B_{m\alpha\beta}. \end{cases}$$

Since $|b_{m\alpha\beta}|$, $|a^{i\ell}_{m\alpha\beta}| \le M^{|\alpha|+|\beta|+1}$ for suitable M, it is enough to have

$$(3)\qquad \begin{cases} A^{i\ell}_{m\alpha\beta} \ge M^{|\alpha|+|\beta|+1} \\ B_{m\alpha\beta} \ge M^{|\alpha|+|\beta|+1}. \end{cases}$$

Thus if F, G are selected so that (3) holds (then (i)′, (i)′ is called a "majorant system" for (i), (ii)) and so that there exists a real-analytic solution U of (i)′, (ii)′ near $(x,t) = (0,0)$, then by (2) we will have convergence of the series on the right of (1) as required.

Selection of $F^{i\ell}_j$, G_j:

Take

$$F^{i\ell}_j(x,U) = \frac{M}{1 - M(\sum_k x^k + \sum_k U^k)} \quad \forall i,j,\ell$$

$$G_j(x,U) = \frac{M}{1 - M(\sum_k x^k + \sum_k U_k)} \quad \forall j.$$

We show that (3) holds in this case. Indeed since

$$\frac{M}{1 - M(\sum_k x^k + \sum_k U^k)} = M(1 + \textstyle\sum_{\ell=1}^\infty M^\ell (\sum_k x^k + \sum_k U^k)^\ell)$$

and since, for any given multi-indices α, β, the coefficient of $x^\alpha U^\beta$ in this expansion is $\ge M^{1+|\alpha|+|\beta|}$, we do indeed have (3).

It remains only to show that (i)′, (ii)′ has some real-analytic solution near $(0,0)$ with this choice of F, G.

In view of the special nature of F, G it seems reasonable to look for a solution U with the special form

$$U_1(x,y) = U_2(x,t) = \cdots = U_N(x,t) = \varphi(\textstyle\sum_k x^k, t).$$

Then the system (i)′, (ii)′ reduces to

$$\begin{cases} (1 - M(y + N\varphi))\varphi_t = nMN\varphi_y + M \\ \varphi(y,0) = 0 \qquad (y = \sum_k x^k). \end{cases}$$

Using the method of Lecture 2, it is easy to obtain (explicitly) a real-analytic solution of this first order equation (see problem 3.3 below). This completes the proof of the existence of real analytic solutions of $(*)$, $(**)$ near $(x,t) = (0,0)$.

We leave the remainder of the proof as an exercise—see problem 3.4 below.

LECTURE 3 PROBLEMS

3.1 Consider the Cauchy problem

$$(*)\qquad \begin{cases} u_{xx} + u_{yy} = 0 \\ u(x,0) = x^2,\ u_y(x,0) = e^x. \end{cases}$$

Following the procedure of lecture, reduce to an equivalent first order Cauchy problem.

Either by using this first order system, or (easier!) working directly with $(*)$, write down the first few terms (up to power $x^\alpha y^\beta$, $|\alpha| + |\beta| \le 3$ say) of the Taylor series of the solution near $(0,0)$.

3.2 Suppose $p = \sum_{n=0}^\infty p_n x^n$, $q = \sum_{n=0}^\infty q_n x^n$, $f = \sum_{n=0}^\infty f_n x^n$ are given analytic functions with $|p_n|, |q_n|, |f_n| \le M/\rho^n$, where $\rho, M > 0$ are constants. Consider the ODE problem

$$(*)\qquad \begin{cases} u'' = p(x)u' + q(x)u + f(x),\ |x| < \rho \\ u(0) = u_0,\ u'(0) = u_1 \quad (u_0,\ u_1 \text{ given constants}) \end{cases}$$

Show that if $u = \sum_{j=0}^\infty a_j x^j$ is an analytic solution of $(*)$, then there is a recurrence relation giving a_{n+1} in terms of $u_0, u_1, a_0, \ldots, a_n, p_0, \ldots, p_n, q_0, \ldots, q_n, f_0, \ldots, f_n$. Show that the series $\sum_{j=0}^\infty a_j x^j$ so defined, is majorized by the power series solution of

$$(**)\qquad \begin{cases} U'' = PU' + QU + F \\ U(0) = |u(0)|,\ U'(0) = |u_1|, \end{cases}$$

where $Q = \frac{M\rho^2}{(\rho-x)^2}$, $P = \frac{M\rho}{\rho - x}$, $F = \frac{M\rho^2}{(\rho-x)^2}$.

Hence prove that $(*)$ has an analytic solution for $|x| < \rho$. (This is of course an important theorem in ODE theory.)

3.3 Show that if the first order problem

$$\begin{cases} a\varphi_x + (bx + c\varphi - 1)\varphi_y = d \\ \varphi(x,0) = 0 \end{cases}$$

(in which a, b, c, d are given constants) has the real-analytic solution $\varphi(x,y) = d(cd - ab)^{-1}\big(1 - bx - \sqrt{(1-bx)^2 + 2(cd+ab)y}\big)$ in case $cd - ab \ne 0$; find also the explicit solution formula when $ab = cd$.

3.4 Complete the proof of the Cauchy-Kovalewski theorem. (We already showed we could find a real-analytic solution in a neighbourhood of (0,0); begin by showing that this solution is unique, at least in some neighbourhood of (0,0).)

Lecture 4

Basic Theory of Harmonic Functions

Here Ω is an open, connected subset of $\mathbb{R}^n$, $n \geq 2$.

Definition. $u \in C^2(\Omega)$ is said to be harmonic if $\Delta u \equiv 0$ in Ω. (u is said to be <u>sub</u>harmonic if $\Delta u \geq 0$ in Ω, <u>super</u>harmonic if $\Delta u \leq 0$ in Ω.)

Lemma. (Mean Value Property.) *If $u \in C^2(\Omega)$ is harmonic in Ω, then for each ball $B_R(x)$ with $\overline{B}_r(x) \subset \Omega$, we have*

$$(*) \qquad u(x) = \frac{\int_{B_R(x)} u(y)\,dy}{\omega_n R^n} = \frac{\int_{\partial B_R(x)} u(y)\,d\mu_R(y)}{\sigma_{n-1} R^{n-1}}.$$

(Here ω_n = volume of the unit ball in $\mathbb{R}^n$, σ_{n-1} = $(n-1)$-dimensional surface area of the unit sphere S^{n-1}, and μ_R is $(n-1)$-dimensional surface measure on $\partial B_R(x)$.)

Proof: First recall the spherical integration formulae:

$$(1) \qquad \int_0^R \int_{\partial B_\rho(x)} f(y)\,d\mu_\rho(y) d\rho = \int_{B_R(x)} f(y)\,dy$$

$$(2) \qquad \int_{\partial B_\rho(x)} f(y)\,d\mu_\rho(y) = \rho^{n-1} \int_{S^{n-1}} f(x+\rho\omega)\,d\omega, \quad 0 < \rho \leq R,$$

where $d\omega$ = integration with respect to $(n-1)$-dimensional surface measure on S^{n-1}, $d\mu_\rho$ = integration with respect to $(n-1)$-dimensional surface measure on $\partial B_\rho(x)$, and f is an arbitrary continuous function on $\overline{B}_R(x_0)$.

Since $0 = \Delta u = \operatorname{div}(Du)$, the divergence theorem implies

$$\int_{\partial B_\rho(x)} \frac{y-x}{|y-x|} \cdot Du(y)\, d\mu_R(y) = 0,$$

so by (2)

$$\int_{S^{n-1}} \omega \cdot Du(x+\rho\omega)\, d\omega = 0.$$

That is

$$\int_{S^{n-1}} \frac{\partial}{\partial\rho}[u(x+\rho\omega)]\, d\omega = 0,$$

which by the differentiation under the integral theorem gives

$$\frac{\partial}{\partial\rho}\int_{S^{n-1}} u(x+\rho\omega)\, d\omega = 0,$$

or $\int_{S^{n-1}} u(x+\rho\omega)\, d\omega \equiv \text{const.}$, $\rho \le R$. Taking a limit as $\rho \downarrow 0$ and using continuity of u at x, we see that the constant here is $\sigma_{n-1}u(x)$, and hence

$$\int_{S^{n-1}} u(x+\rho\omega)\, d\omega = \sigma_{n-1}u(x), \quad \rho \le R.$$

Multiplying by ρ^{n-1} and using (1), (2), we then have the required relations.

Remark 1: If u is sub-(resp. super-)harmonic, then the above argument gives

$$u(x) \le (\text{resp.} \ge) \frac{\int_{\partial B_R(x)} u(y)\, d\mu_R(y)}{\sigma_{n-1}R^{n-1}}$$

and

$$u(x) \le (\text{resp.} \ge) \frac{\int_{B_R(x)} u(y)\, dy}{\omega_n R^n}.$$

Using the above mean-value result, we can derive a number of very important qualitative properties of harmonic and sub/super-harmonic functions.

Lemma. (Strict maximum principle.) *If u is C^2 and sub-harmonic in Ω (i.e. $\Delta u \ge 0$ in Ω), then u cannot attain a maximum in Ω unless it is constant.*

Proof: Suppose u attains maximum value M at some point $y \in \Omega$. Then $M-u$ is a non-negative super-harmonic function in Ω, and if $R < \operatorname{dist}\{y, \partial\Omega\}$ we have that $\int_{B_R(y)}(M-u) = 0$ by Remark 1 above, so $u \equiv M$ in $B_R(y)$. This shows that $\{x : u(x) = M\}$ is open in Ω; but by continuity of u it is also closed in Ω, so $u \equiv M$ on all of Ω, as required.

Notice that the following corollaries follow immediately from the above result, in view of the fact that a continuous function on a compact set has a maximum value somewhere in the set.

Corollary 1. *If Ω is bounded, and $u \in C^2(\Omega) \cap C^0(\overline{\Omega})$ is subharmonic in Ω, then $\sup_\Omega u = \sup_{\partial\Omega} u$; i.e.*

$$u(x) \le \sup_{\partial\Omega} u \quad \forall x \in \overline{\Omega}.$$

Corollary 2. *If $u, v \in C^2(\Omega) \cap C^0(\overline{\Omega})$ are both harmonic in Ω, if Ω is bounded, and if $u \equiv v$ on $\partial\Omega$, then $u \equiv v$ everywhere in $\overline{\Omega}$. (i.e. for given boundary values on a bounded domain, there is at most one continuous harmonic function on the closure of the domain which takes on the given boundary values.)*

Corollary 3. *If $u \in C^2(\overline{\Omega})$, $\Delta u \ge f$ on Ω (where f is a given continuous function on Ω), then the following two estimates hold:*

(i) $$u(x_0) \le \frac{\int_{B_R(x_0)} u}{\omega_n R^n} + CR^2 \sup_{B_R(x_0)} f_-,$$

for each ball $B_R(x_0) \subset \Omega$, and

(ii) $$\sup_\Omega u \le \sup_{\partial\Omega} u + C d^2 \sup_\Omega f_-,$$

where $d = \operatorname{diam}\Omega$. In each of these estimates C is a constant depending only on n, and $f_- = \max\{-f, 0\}$.

Remark 2: Notice that if $\Delta u = f$ on Ω, then we can also use the above results with $-u$, $-f$ in place of u, f respectively, thus giving

(i)^0 $$|u(x_0)| \le \frac{\left|\int_{B_R(x_0)} u\right|}{\omega_n R^n} + CR^2 \sup_{B_R(x_0)} |f|$$

and

(ii)^0 $$\sup_\Omega |u| \le \sup_{\partial\Omega} |u| + C d^2 \sup_\Omega |f|,$$

respectively.

Proof of Corollary 3: Assume without loss of generality that $0 \in \Omega$ and that, for the proof of (i), $x_0 = 0$. Notice that $\Delta|x|^2 \equiv 2n$, so that $\Delta(u + (2n)^{-1}|x|^2 \sup_\Omega f_-) = f + \sup_\Omega f_- \geq 0$, and hence we can apply the results in Remark 1 and Corollary 1 above to the function $u + (2n)^{-1}|x|^2 \sup_\Omega f_-$, giving (i) and (ii) as required.

Another important consequence of the mean-value property is given in the following lemma.

Lemma. (Harnack's Inequality.) *If $u \in C^2(\Omega)$ is a non-negative harmonic function, then for any compact $K \subset \Omega$ there is a constant $c_K > 0$ (independent of u) such that*

$$\sup_K u \leq c_K \inf_K u.$$

Proof: It is evidently enough to prove such an inequality in the special case $K \subset B_R(y)$ under the assumption that $B_{4R}(y) \subset \Omega$, because (by connectedness of Ω) we can cover K by a finite collection $\{B_{R_j}(y_j)\}$ of balls with $B_{4R_j}(y_j) \subset \Omega\ \forall j$ and with the property that for every pair of balls $B_{R_i}(y_i)$, $B_{R_j}(y_j)$ there is a sequence $\{B_{R_{i_k}}(y_{i_k})\}_{k=1,\dots,M}$ $(M \leq N)$ such that $B_{R_{i_1}}(y_{i_1}) = B_{R_i}(y_i)$ and $B_{R_{i_M}}(y_{i_M}) = B_{R_j}(y_j)$. Then the validity of the inequality on each ball $B_{R_j}(y_j)$ evidently implies its validity on K (by induction on the number M of balls needed in such a collection).

By the mean-value property $(*)$, for any two points $x_1, x_2 \in B_R(y)$ we have

$$u(x_1) = \frac{\int_{B_R(x_1)} u(x)\,dx}{\omega_n R^n} \leq \frac{\int_{B_{2R}(y)} u(x)\,dx}{\omega_n R^n}$$

since $u \geq 0$ and $B_{2R}(y) \supset B_R(x_1)$, and

$$u(x_2) = \frac{\int_{B_{3R}(x_2)} u(x)\,dx}{\omega_n (3R)^n} \geq \frac{\int_{B_{2R}(y)} u(x)\,dx}{3^n \omega_n R^n}$$

since $u \geq 0$ and $B_{2R}(y) \subset B_{3R}(x_2)$. Thus $u(x_1) \leq 3^n u(x_2)$.

In view of the arbitrariness of x_1, x_2 this gives

$$\sup_{B_R(y)} u \leq 3^n \inf_{B_R(y)} u$$

as required.

As a final application of the mean-value property we want to prove H. Weyl's famous result that all weak solutions of $\Delta u = 0$ are smooth (hence real-analytic by Exercise 4.5 (i) below).

Weyl's Lemma. *If $u \in L^1_{\mathrm{loc}}(\Omega)$ is weakly harmonic (i.e. $\int_\Omega u\Delta\psi = 0\ \forall \psi \in C_c^\infty(\Omega)$), then (after redefinition on a set of measure zero) u is a $C^\infty(\Omega)$ harmonic function.*

Proof: Evidently, since this is a local result, it is enough to prove $u \in C^\infty(B_{x_0})$ for each ball $B_R(x_0)$ with $\overline{B}_R(x_0) \subset \Omega$. By the discussion of Lecture 2, we know that $\Delta u_\sigma = 0$ on $B_{R-\sigma}(x_0)$ for each $0 < \sigma < R$, where (as in Lecture 2) u_σ denotes the mollified function associated with u; let us agree that the mollifier φ being used here is actually radially symmetric. That is

$$\varphi(r\omega) \equiv \gamma(r), \quad r > 0,$$

where $\gamma(r)$ is a non-negative smooth function of r with $\gamma(r) \equiv 0$ on $[1-\varepsilon, \infty)$ for some $\varepsilon > 0$, $\gamma \equiv$ const. in a neighborhood of $r = 0$, and

$$(1) \qquad \sigma_{n-1} \int_0^1 \gamma(r) r^{n-1}\,dr = 1.$$

Notice that the latter guarantees the condition $\int_{\mathbb{R}^n} \varphi(x)\,dx = 1$. Now let $0 < \sigma < R$ be fixed and let $v = u_\sigma$ on $B_{R-\sigma}(x_0)$ and $v \equiv 0$ on $\mathbb{R}^n \setminus B_{R-\sigma}(x_0)$, so that v is smooth harmonic on $B_{R-\sigma}(x_0)$ by the above discussion, and by the mean-value property $(*)$ we have for $0 < \tau < R - \sigma$

$$(2) \qquad \sigma_{n-1} v(x) = \int_{S^{n-1}} v(x + r\omega) d\omega \quad \forall\, x \in B_{R-\sigma-\tau}(x_0),\ r < \tau$$

On the other hand by definition of smoothing we have for x, τ as in (2),

$$\begin{aligned} v_\tau(x) &= \tau^{-n} \int \varphi\Big(\frac{x-y}{\tau}\Big) v(y)\,dy \\ &= \tau^{-n} \int \varphi\Big(\frac{y}{\tau}\Big) v(x-y)\,dy \\ &= \tau^{-n} \int \varphi\Big(\frac{-r\omega}{\tau}\Big) v(x + r\omega) r^{n-1}\,dr d\omega \ \ \text{(by change } y = -r\omega) \\ &= \int \gamma(s) v(x + s\omega) s^{n-1}\,ds d\omega (\text{ by change } s = r/t) \\ &= \sigma_{n-1} v(x) \int_0^1 \gamma(s) s^{n-1}\,ds = v(x) \ \text{(by (2) and (1))}. \end{aligned}$$

Thus we have proved that $u_\sigma \equiv (u_\sigma)_\tau (\equiv (u_\tau)_\sigma)$ in $B_{R-\sigma-\tau}(x_0)$ for each $\sigma, \tau > 0$ with $\sigma + \tau < R$. Taking limit as $\sigma \downarrow 0$ (holding τ fixed), this gives $u(x) =$

$u_\tau(x)$ a.e. $x \in B_{R-\tau}(x_0)$, so that u agrees a.e. with a smooth harmonic function on $B_{R-\tau}(x_0)$. Since τ is arbitrary, this is the required result.

Next we want to describe the Green's formulae and the notion of Green's function. Here we assume that Ω is a bounded C^1 domain, meaning that $\partial\Omega$ has an inward pointing unit normal η on $\partial\Omega$ which varies continuously on $\partial\Omega$.

Let $u, v \in C^2(\overline{\Omega})$ for the moment be arbitrary. Since $\Delta u = \operatorname{div}(Du)$, the divergence theorem implies

$$\int_\Omega v\Delta u = -\int_\Omega Du \cdot Dv + \int_{\partial\Omega} v\frac{\partial u}{\partial \eta},$$

where $\partial u/\partial\eta = \eta \cdot Du =$ directional derivative of u in the direction of η. Interchanging u and v and subtracting the result, we get

$$(1) \qquad \int_\Omega (v\Delta u - u\Delta v) = \int_{\partial\Omega}(v\frac{\partial u}{\partial\eta} - u\frac{\partial v}{\partial\eta})$$

which is called Green's second identity.

A good thing to do is to plug in the radially symmetric solutions of $\Delta u = 0$. It turns out that (up to additive and multiplicative constants) the only solution $K(x)$ of $\Delta u = 0$ which can be expressed as a function of $|x|$ is the function $K(x) \equiv \Gamma(|x|)$, where

$$\Gamma(r) = \begin{cases} \dfrac{-1}{(n-2)\sigma_{n-1}}\dfrac{1}{r^{n-2}}, & n \ge 3 \\ \dfrac{1}{2\psi}\log r, & n = 2\,. \end{cases}$$

(See Exercise 4.2 below.) Of course $K(x-y) \equiv \Gamma(|x-y|)$ is then a solution of $\Delta u = 0$ on $\mathbb{R}^n \setminus \{y\}$ for any fixed $y \in \mathbb{R}^n$. We take $y \in \Omega$, and use (1) with $\Omega \setminus B_\rho(y)$ in place of Ω and $K(x-y)$ in place of v, where $\rho < \operatorname{dist}\{y, \partial\Omega\}$.

Since $\Delta_x K(x-y) = 0$, this gives

$$\begin{aligned}\int_{\Omega\setminus B_\rho(y)} K(x-y)\Delta u(x)\,dx &= \int_{\partial\Omega}(K(x-y)\frac{\partial u}{\partial\eta} - u\frac{\partial K(x-y)}{\partial\eta}) \\ &\quad - \int_{\partial B_\rho(y)}(K(x-y)\frac{\partial u}{\partial r} - u\frac{\partial K(x-y)}{\partial r})\big|_{r=\rho},\end{aligned}$$

where $r = |x-y|$ in the last integral. Letting $\rho \downarrow 0$ we thus get

$$(*) \qquad u(y) = \textstyle\int_\Omega K(x-y)\Delta u(x)\,dx + \int_{\partial\Omega}(u\frac{\partial K(x-y)}{\partial\eta} - K(x-y)\frac{\partial u}{\partial\eta}).$$

Notice that this representation is valid for an arbitrary $C^2(\overline{\Omega})$ function.

Next, repeat the above argument with $h(x,y)$ in place of $K(x-y)$, where $h(\cdot, y)$ is $C^2(\overline{\Omega})$, and harmonic, for fixed u, with $h(x,y) = K(x-y)$ for $x \in \partial\Omega$ and $y \in \Omega$. (We shall show that such h exists in case Ω is a smooth domain later.)

Then we get an identity as in $(*)$ with zero on the left. Subtracting this new identity from $(*)$ we get

$$(**) \qquad u(y) = \int_\Omega G(x,y)\Delta u(x)\,dx + \int_{\partial\Omega} u\frac{\partial G(\cdot,y)}{\partial\eta},$$

where

$$G(x,y) = K(x-y) - h(x,y);$$

G, so defined, is called the Green's function for the domain Ω. We show later that such a Green's function exists for any smooth bounded domain. Here we show that it is possible to compute an explicit expression for the Green's function of a ball.

To do this, consider the "inversion" in the sphere $\partial B_\rho(0)$ given by $x \mapsto x^* \equiv \frac{\rho^2 x}{|x|^2}$, $x \neq 0$. Notice that $x^* = x$ for $x \in \partial B_\rho(0)$, and by direct computation,

$$|x^* - y^*| = \frac{\rho^2}{|x||y|}|x-y|,\ x, y \neq 0,$$

so in particular

$$|x-y| = \frac{|y|}{\rho}|x - y^*|,\ x \in \partial B_\rho(0).$$

With Γ as above, notice that then we have

$$\Gamma(|x-y|) = \Gamma(\frac{|y|}{\rho}|x-y^*|),\ x \in \partial B_\rho(0),\ y \in \Omega.$$

(Notice particularly that $\Gamma(\frac{|y|}{\rho}|x-y^*|) = \Gamma((\frac{|x|^2|y|^2}{\rho^2} + \rho^2 - 2x\cdot y)^{1/2})$ extends smoothly to $y = 0$.)

Then since $\Gamma(\frac{|y|}{\rho}|x-y^*|)$ is harmonic in x for fixed y except at $x = y^*$, and since $y^* \notin \overline{B}_\rho(0)$ for $y \in B_\rho(0)$, we see that the Green's function for $B_\rho(0)$ is

$$G(x,y) = \Gamma(|x-y|) - \Gamma((\frac{|x|^2|y|^2}{\rho^2} + \rho^2 - 2x\cdot y)^{1/2}).$$

According to $(**)$ above, we then have that, if u is C^2 and harmonic on $\overline{B}_\rho(0)$,

$$u(y) = \int_{\partial B_\rho(0)} u(x)\frac{\partial}{\partial r}G(r\omega, y)|_{r=\rho}\,d\mu_\rho(x).$$

By direct computation

$$\frac{\partial G}{\partial r}(r\omega, y)|_{r=\rho} = \frac{\rho^2 - |y|^2}{n\omega_n\rho}|x-y|^{-n}, \quad (x = \rho\omega)$$

so we end up with the formula

$$u(y) = \frac{\rho^2 - |y|^2}{n\omega_n\rho}\int_{\partial B_\rho(0)} \frac{u(x)}{|x-y|^n}\, d\mu_\rho(x),$$

which is known as the Poisson integral formula for harmonic u on $B_r(0)$.

Now if φ is an arbitrary continuous function on $\partial B_\rho(0)$, then we can define

$$u(x) = \frac{\rho^2 - |x|^2}{n\omega_n\rho}\int_{\partial B_\rho(0)} \frac{\varphi(y)}{|x-y|^n}\, d\mu_\rho(y), \quad x \in B_\rho(0),$$

and one can check that this gives a harmonic function on $B_\rho(0)$ which extends continuously to $\overline{B}_\rho(0)$ and which satisfies $u|_{\partial B_\rho(0)} = \varphi$. (See Gilbarg & Trudinger for example for details.)

There is one further consequence of the Green's identities which should be mentioned here. Namely, if $f \in C_c^\infty(\mathbb{R}^n)$, then the identity $(*)$ (with $u = f$ and with Ω chosen to contain the support of f) implies the identity

$$(**) \qquad \Delta\int_{\mathbb{R}^n} K(x-y)f(y)\,dy = f(x), \quad x \in \mathbb{R}^n.$$

(The reader familiar with the theory of distributions will see that this says exactly that the Laplacian of $K(x-y)$ is equal to the Dirac delta distribution with support at the point x.) Furthermore, by change of variable $z = x - y$, the integral can be written $\int_{\mathbb{R}^n} K(y)f(x-y)\,dy$ which is evidently $C^\infty(\mathbb{R}^n)$ by the differentiation under the integral sign theorem.

Notice that since $\Delta_x K(x-y) \equiv 0$, $y \neq x$, it is also merely a matter of differentiation under the integral to show that, if Ω is any open subset of $\mathbb{R}^n$, $\int_{\mathbb{R}^n\setminus\Omega} K(x-y)f(y)\,dy$, is a smooth function of x for $x \in \Omega$ and

$$\Delta_x\int_{\mathbb{R}^n\setminus\Omega} K(x-y)f(y)\,dy = 0, \quad x \in \Omega,$$

even if f is merely in $L^1(\mathbb{R}^n \setminus \Omega)$. It then follows directly that in fact if Ω is any open subset of $\mathbb{R}^n$ and if $f \in C^\infty(\Omega) \cap L^1(\Omega)$, then $\int_\Omega K(x-y)f(y)\,dy$ is a $C^\infty(\Omega)$ function of x and

$$(**) \qquad \Delta\int_\Omega K(x-y)f(y)\,dy = f(x), \quad x \in \Omega.$$

(See Exercise 4.6 below.) We show later that the last identity remains valid if f is merely Hölder continuous (rather than C^∞) on Ω.

Notice that $(**)$ makes it possible to establish the following corollary to Weyl's Lemma:

Lemma. (Corollary of Weyl's Lemma.) *If $u \in L^1_{\mathrm{loc}}(\Omega)$ is a weak solution of $\Delta u = f$ on Ω, where $f \in C^\infty(\Omega)$ is given (thus $\int_\Omega u\Delta\varphi = \int_\Omega f\varphi\ \forall\varphi \in C_c^\infty(\Omega)$), then (after redefinition on a set of measure zero) u is a C^∞ function.*

We leave the proof as an exercise (see 4.6 (ii) below).

LECTURE 4 PROBLEMS

4.1 Prove that a weak $L^1_{\text{loc}}(\Omega)$ solution u of the Cauchy-Riemann equations $u_x + iu_y = 0$ is automatically smooth.

Hint: With $u = u_1 + iu_2$, prove that u_1 is a weak solution of $\Delta u_1 = 0$ in Ω.

4.2 Suppose $u = \gamma(|x|)$ where $\gamma : (a, b) \to \mathbb{R}$ is smooth and $0 < a < b \le \infty$ are given. Prove u is harmonic on $a < |x| < b$ if and only if there are constants c_1, c_2 such that

$$\gamma(r) = \begin{cases} c_1 r^{2-n} + c_2, \; n \ge 3 \\ c_1 \log r + c_2, \; n = 2. \end{cases}$$

4.3 By reversing the appropriate steps in Lecture 4, prove that if $u \in C^2(\Omega)$ and if u has the mean-value property

$$u(x) = \frac{\int_{B_R(x)} u(y)\, dy}{\omega_n R^n}$$

whenever $\overline{B}_R(x) \subset \Omega$, then u is harmonic.

4.4 Use the mean-value property to prove that if u is a (smooth) harmonic function in Ω and if $\overline{B}_R(x) \subset \Omega$, then there exists a constant $C = C(n)$ such that for any multi-index $\alpha = (\alpha_1, \ldots, \alpha_n)$

$$\sup_{B_{R/2}(x)} |D^\alpha u| \le C^{|\alpha|} \alpha! R^{-|\alpha|} \sup_{B_R(x)} |u|.$$

Hint: First establish that for any $\rho_1 < \rho_2 < R$

$$\sup_{B_{\rho_1}(x)} |Du| \le \frac{C}{\rho_2 - \rho_1} \sup_{B_{\rho_2}(x)} |u|$$

for suitable $C = C(n)$.

4.5 Use the result of 4.4 to prove that

(i) Any harmonic function on Ω is automatically real-analytic on Ω.

(ii) If u (not identically zero) is harmonic on all of $\mathbb{R}^n$ and if $\exists$ a constant $c > 0$ and an integer $k \ge 0$ such that

$$|u(x)| \le c(1 + |x|)^k \;\; \forall x \in \mathbb{R}^n,$$

then u is a polynomial of degree k.

4.6 (i) Justify the identity $(**)$ in the above lecture for any $f \in C^\infty(\Omega) \cap L^1(\Omega)$.

Hint: For any $\Omega_1 \subset\subset \Omega$, we can find $\widetilde{f} \in C_c^\infty(\Omega)$ with $\widetilde{f} \equiv f$ on Ω_1.

(ii) Give the proof of the corollary to Weyl's lemma stated at the end of the above lecture.

4.7 Prove that the Cauchy problem

$$\begin{cases} u \in C^2 \text{ and } \Delta u = 0 \text{ in a neighbourhood of } x^n = 0 \\ u(x^1, \ldots, x^{n-1}, 0) = 0, \;\; \dfrac{\partial u}{\partial x^n}(x^1, \ldots, x^{n-1}, 0) = \varphi(x^1, \ldots, x^{n-1}) \end{cases}$$

has a C^2 solution if φ is real-analytic on $\mathbb{R}^{n-1}$, but has no C^2 solution otherwise.

4.8 (Reflection Principle.) Let $B^+_\rho(0) = \{x = (x^1, \ldots, x^n) \in B_\rho(0) : x^n > 0\}$, and suppose $u \in C^2(B^+_\rho(0)) \cap C^0(\overline{B}^+_\rho(0))$ is harmonic in $B^+_\rho(0)$ and satisfies $u = 0$ on $B_\rho(0) \cap \{x : x^n = 0\}$.

Prove that the odd extension of u (i.e. $\widetilde{u}(x^1, \ldots, -x^n) = -u(x^1, \ldots, x^n)$) is smooth and harmonic on all of $B_\rho(0)$.

Lecture 5

Introduction to Sobolev Spaces & the Fourier Transform

To begin we need to introduce the Sobolev spaces $W^{m,p}(\Omega)$ and $H^m(\Omega)$ for a domain $\Omega \subset \mathbb{R}^n$. For this, we first need the notion of weak derivatives of a function $u \in L^1_{\text{loc}}(\Omega)$ as follows:

If $\alpha = (\alpha_1, \ldots, \alpha_n)$ is a multi-index, we say that r_α is a weak derivative of u corresponding to the multi-index α if $r_\alpha \in L^1_{\text{loc}}(\Omega)$ and if the equation $D^\alpha u = r_\alpha$ holds weakly in $L^1_{\text{loc}}(\Omega)$; i.e. if $r_\alpha \in L^1_{\text{loc}}(\Omega)$ and

$$(*) \qquad \int_\Omega r_\alpha \varphi = \int_\Omega (-1)^{|\alpha|} u D^\alpha \varphi \quad \forall \varphi \in C_c^\infty(\Omega).$$

(For those familiar with the theory of distributions this is evidently exactly the statement that the distribution derivative $D^\alpha u$ of u is in $L^1_{\text{loc}}(\Omega)$.)

If such weak derivatives r_α exist, we normally denote them by $D^\alpha u$. Notice that of course they are unique as $L^1_{\text{loc}}(\Omega)$ functions if they exist, because if $r_\alpha, \tilde{r}_\alpha$ are weak derivatives of u corresponding to the same multi-index α, then by the above definition we must have $\int_\Omega (r_\alpha - \tilde{r}_\alpha)\varphi = 0 \, \forall \varphi \in C_c^\infty(\Omega)$, and hence $r_\alpha = \tilde{r}_\alpha$ a.e. in Ω as required. Notice also that the notation $D^\alpha u$ is not in conflict with classical notation, because if $u \in C^m(\Omega)$ then the classical derivatives $D^\alpha u$ are also the weak derivatives in the above sense for $|\alpha| \leq m$. (In this case $(*)$ is just the usual integration by parts formula for $u \in C^m(\Omega)$.)

$W^{m,p}_{\text{loc}}(\Omega)$ will denote the set of all $L^p_{\text{loc}}(\Omega)$ functions u such that the weak derivatives $D^\alpha u$, $|\alpha| \le m$, exist and are in $L^p_{\text{loc}}(\Omega)$.

$W^{m,p}(\Omega)$ denotes the set of $L^p(\Omega)$ functions u with weak derivatives $D^\alpha u$ in $L^p(\Omega)$, $|\alpha| \le m$.

For $W^{m,p}(\Omega)$ we have the norm

$$\|u\|_{m,p,\Omega} = \textstyle\sum_{|\alpha|\le m} \|D^\alpha u\|_{L^p(\Omega)}.$$

In case $p = 2$ we use instead the equivalent norm $(\sum_{|\alpha|\le m}\|D^\alpha u\|^2_{L^2(\Omega)})^{1/2}$, denoted simply $\|u\|_{m,\Omega}$. Also, we write

$$H^m(\Omega) = W^{m,2}(\Omega), \;\; H^m_{\text{loc}}(\Omega) = W^{m,2}_{\text{loc}}(\Omega).$$

Notice that $W^{m,p}(\Omega)$ is a Banach space $\forall m, p \ge 1$, and $H^m(\Omega)$ is a Hilbert space with inner product

$$\langle u, v\rangle_{H^m(\Omega)} = \textstyle\sum_{|\alpha|\le m}\langle D^\alpha u, D^\alpha v\rangle_{L^2(\Omega)},$$

so that $\langle u, u\rangle_{H^m(\Omega)} = \|u\|^2_{m,\Omega}$ in the notation introduced above. (See Exercise 5.1 below.)

An important point is that mollification commutes with the operation of taking weak derivatives:

$$(*) \qquad D^\alpha u_\sigma(x) = (D^\alpha u)_\sigma(x), \;\; x \in \Omega_\sigma, \; |\alpha| \le m,$$

for $u \in W^{m,p}_{\text{loc}}(\Omega)$, where $\Omega_\sigma = \{x \in \Omega : \text{dist}\,(x, \partial\Omega) > \sigma\}$. In fact, for $x \in \Omega_\sigma$, we have

$$\begin{aligned} D^\alpha u_\sigma(x) &= \int_\Omega u(y) D^\alpha_x[\varphi^{(\sigma)}(x-y)]\,dy \\ &= \int_\Omega u(y)(-1)^{|\alpha|} D^\alpha_y[\varphi^{(\sigma)}(x-y)]\,dy \\ &= \int_\Omega D^\alpha u(y)\varphi^{(\sigma)}(x-y)\,dy \;\text{ by definition of weak derivative} \\ &= (D^\alpha u)_\sigma(x), \end{aligned}$$

as required. Notice that here we used the fact that, as a function of y, $\varphi^{(\sigma)}(x-y)$ is a C^∞_c function with support in $B_\sigma(x)(\subset \Omega$ for $x \in \Omega_\sigma)$.

As we mentioned in the previous lecture, it is true that $f_\sigma(x) \to f(x)$ a.e. $x \in \Omega$ in case $f \in L^1_{\text{loc}}(\Omega)$. It is also true (see problem 5.11 below) that $f_\sigma \to f$ locally with respect to the L^p norm in Ω in case $f \in L^p_{\text{loc}}(\Omega)$; hence using the

above fact that the operations of mollification and the taking of weak derivatives commute, we conclude

Lemma 1. *If $u \in W^{m,p}_{\text{loc}}(\Omega)$ and if $|\alpha| \le m$, then $D^\alpha u_\sigma \to D^\alpha u$ pointwise a.e. in Ω, and also locally with respect to the $\| \;\|_{m,p}$ norm in Ω. Thus $u_\sigma \to u$ in $W^{m,p}(B_\rho(x_0))$ for any ball $B_\rho(x_0)$ with $\overline{B}_\rho(x_0) \subset \Omega$.*

It is easy to prove that most of the usual rules of calculus have analogues for weak derivatives: for example

(i) $\qquad u, v \in W^{m,p}(\Omega)$ and $C \in \mathbb{R} \Rightarrow u + v, \, Cu \in W^{m,p}(\Omega)$,

and $D^\alpha(u+v) = D^\alpha u + D^\alpha v$, $D^\alpha(Cu) = CD^\alpha u$. Also

(ii) $\qquad u \in W^{m,\infty}(\Omega), \, v \in W^{m,p}(\Omega) \Rightarrow uv \in W^{m,p}(\Omega)$,

and the usual Leibniz formula is applicable for computing the derivatives of uv; thus for each multi-index α with $|\alpha| \le m$

$$D^\alpha(uv) = \textstyle\sum_{\beta+\gamma=\alpha} \dfrac{\alpha!}{\beta!\,\gamma!}(D^\beta u)(D^\gamma v).$$

Also,

(iii)

$f \in C^1(\mathbb{R})$ with f' bounded and $u \in W^{1,p}(\Omega) \Rightarrow$

$$f \circ u \in W^{1,p}(\Omega), \text{ and } D_j f \circ u = f'(u) D_j u \text{ on } \Omega.$$

(i) here is easily checked by using the definition of weak derivative; (ii), (iii) are checked by using $(*)$ above together with the fact that the stated results evidently hold with the smooth functions u_σ, v_σ on Ω_σ in place of u, v on Ω.

Recall that a function $u : \Omega \to \mathbb{R}$ is said to be locally Lipschitz on Ω (written $u \in C^{0,1}(\Omega)$) if for each ball $B_\rho(x_0)$ with $\overline{B}_\rho(x_0) \subset \Omega$ there is a constant K such that $|u(x) - u(y)| \le K|x-y|$ for all $x, y \in B_\rho(x_0)$. (Any such number K is called a Lipschitz constant for u on $B_\rho(x_0)$.) Using the above facts about mollification it is then easy to check that

$$u \in C^{0,1}(\Omega) \iff u \in W^{1,\infty}_{\text{loc}}(\Omega),$$

and that furthermore $\|Du\|_{L^\infty(B_\rho(x_0))}$ is the least Lipschitz constant for u on $B_\rho(x_0)$. For example, to prove the direction "$\Leftarrow$" here we first note that by calculus we have $u_\sigma(y) - u_\sigma(x) = \int_0^1 Du_\sigma(z_t)\cdot(y-x)$ if $x, y \in B_{\rho-\sigma}(x_0)$, where $z_t = x + t(y-x)$. (Notice that $B_{\rho-\sigma}(x_0) \subset \Omega_\sigma$ because $B_\rho(x_0) \subset \Omega$.)

By definition of mollification we have $Du_\sigma(z) = \int_{B_\sigma(z)} \varphi^{(\sigma)}(z-\zeta) Du(\zeta)\, d\zeta$, which by change of variable $\xi = \sigma^{-1}(z-\zeta)$ gives $Du_\sigma(z) = \int_{B_1(0)} \varphi(\xi) Du(z-\sigma\xi)\, d\xi$. But now if $z = z_t \in B_{\rho-\sigma}(x_0)$ as above, then, since $\int_{B_1(0)} \varphi = 1$ and since $z_t - \sigma\xi \in B_\rho(x_0)$ for $\xi \in B_1(0)$, we conclude that $|Du_\sigma(z_t)| \le \|Du\|_{L^\infty(B_\rho(x_0))}$ and hence the above identity for $u_\sigma(y) - u_\sigma(x)$ gives $|u_\sigma(y) - u_\sigma(x)| \le \|Du\|_{L^\infty(B_\rho(x_0))}|x-y|$. After letting $\sigma \downarrow 0$ we deduce that u is (almost everywhere equal to) a Lipschitz function on $B_\rho(x_0)$ and $|u(x) - u(y)| \le \|Du\|_{L^\infty(B_\rho(x_0))}|x-y|$, as claimed.

We leave the proof of "$\Rightarrow$" as an exercise.

We shall also need the notion of Lipschitz domain in the sequel: a domain is said to be Lipschitz if for each $x_0 \in \partial\Omega$ there is $\rho > 0$, a Lipschitz function $\psi : \mathbb{R}^{n-1} \to \mathbb{R}$, and coordinates $y = (y^1, \ldots, y^n) = x_0 + Q(x - x_0)$ (Q an orthogonal transformation of $\mathbb{R}^n$) such that

$$\Omega \cap B_\rho(x_0) = \{x_0 + Q^{-1}(y - x_0) : y^n > \psi(y^1, \ldots, y^{n-1})\} \cap B_\rho(x_0).$$

Notice that then in particular we have

$$\begin{aligned}\partial\Omega \cap B_\rho(x_0) &= \{x_0 + Q^{-1}(y - x_0) : y^n = \psi(y^1, \ldots, y^{n-1})\} \cap B_\rho(x_0) \\ &\equiv \{x_0 + Q^{-1}(y - x_0) : y \in \operatorname{graph}\psi \cap B_\rho(0)\}.\end{aligned}$$

Of course a rectangle (or any domain in $\mathbb{R}^2$ with boundary a finite disjoint union of piecewise C^1 Jordan curves with non-zero "corner angles") is such a Lipschitz domain.

Subsequently we also let $W_0^{m,p}(\Omega)$ (or $H_0^m(\Omega)$ in case $p = 2$) be the closure, taken in $W^{m,p}(\Omega)$ with respect to the norm $\|u\|_{m,p,\Omega}$, of the $C_c^\infty(\Omega)$ functions. Of course $W_0^{m,p}(\Omega)$ is then by definition a closed subspace of $W^{m,p}(\Omega)$. In case Ω is bounded, an important fact which distinguishes $W_0^{m,p}(\Omega)$ functions from $W^{m,p}(\Omega)$ functions is that if $u \in W_0^{m,p}\Omega$ then

$$(*) \qquad \int_\Omega \varphi D^\alpha u = \int_\Omega (-1)^{|\alpha|} u D^\alpha \varphi$$

for all $\varphi \in C^\infty(\Omega)$ with $D^\alpha\varphi$ bounded, $|\alpha| \le m$, whereas we need $\varphi \in C_c^\infty(\Omega)$ to ensure this if $u \in W^{m,p}(\Omega)$. Notice that such a formula for $u, \varphi \in C^\infty(\mathbb{R}^n)$ can be obtained directly by repeated integration by parts, only subject to the restriction that the partials of u up to order $m-1$ vanish on $\partial\Omega$; thus $(*)$ suggests that the weak derivatives of order $\le m-1$ of $W_0^{m,p}(\Omega)$ functions should vanish on $\partial\Omega$ in some weak sense; this turns out to be correct—the remark following Lemma 2 below makes this clear at least for Lipschitz domains. See also problem 5.9 below.

For the moment we make an important observation concerning an alternative norm for the Sobolev spaces $W_0^{m,p}(\Omega)$. We claim that in fact for $m, p \ge 1$ the quantity defined by taking the "top order" part of the Sobolev norm is equivalent to the whole Sobolev norm. That is,

$$\|u\|_{W_0^{m,p}(\Omega)} = \Big(\int_\Omega \textstyle\sum_{|\alpha|=m} |D^\alpha u|^p\Big)^{1/p}$$

is a norm for $W_0^{m,p}(\Omega)$ which is equivalent to the norm $\|\ \|_{m,p,\Omega}$. Thus we claim that there is a constant $C = C(n)$ such that

$$(\ddagger) \qquad \|u\|_{m,p,\Omega} \le C\|u\|_{W_0^{m,p}(\Omega)}.$$

(Of course the reverse inequality $\|u\|_{W_0^{m,p}} \le \|u\|_{m,p}$ is trivial.) To see $\ddagger$, first assume $p > 1$ and take $\varphi \in C_c^\infty(\Omega)$. Using integration by parts (noting that $|\varphi|^p \in C_c^1(\Omega)$), we obtain the identity

$$n\int_\Omega |\varphi|^p = -p\int_\Omega \operatorname{sgn}\varphi |\varphi|^{p-1}(x-y)\cdot D\varphi(x)\,dx, \quad y \in \mathbb{R}^n \text{ fixed},$$

where we used the identity $\operatorname{div} x \equiv n$ on $\mathbb{R}^n$. Taking $y \in \Omega$ and noting that $|x-y| \le d$, $d = \operatorname{diam}\Omega$, we then conclude (after an application of the Hölder inequality) that

$$(\ddagger\ddagger) \qquad \|\varphi\|_{L^p(\Omega)} \le n^{-1} pd \|D\varphi\|_{L^p(\Omega)}.$$

This inequality is known as the Poincaré inequality. By taking limits we see that it also holds for $p = 1$, and for any $\varphi \in W_0^{1,p}(\Omega)$. Using this inequality on the weak derivatives of u, it is now easy to check $\ddagger$ above.

The following lemma will be important later:

Lemma 2. *If Ω is a bounded Lipschitz domain and $U_\sigma = \{x \in \Omega : \operatorname{dist}(x, \partial\Omega) < \sigma\}$, then*

(i) $\quad \|u\|_{m-1,U_\sigma} \le C\sqrt{\sigma}\|u\|_{m,\Omega}, \quad 0 < \sigma < \sigma_0, \quad u \in H^m(\Omega),$

(ii) $\quad \|u\|_{m-1,U_\sigma} \le C\sqrt{\sigma}\|u\|_{m,U_{K\sigma}}, \quad 0 < \sigma < \sigma_0, \quad u \in H_0^m(\Omega),$

where C, K, σ_0 are positive constants depending only on Ω.

Remark: Notice that (ii) tells us that the weak derivatives up to and including order $m-1$ actually tend to zero in the L^2 sense that $\lim_{\sigma\downarrow 0} \sigma^{-1}\int_{U_\sigma} |D^\alpha u|^2 = 0$

for $|\alpha| \le m-1$; keep in mind here that σ is proportional to the Lebesgue measure of the boundary strip U_σ.

Proof of Lemma 2: Clearly it suffices to prove the case $m = 1$ of the lemma, since we can deduce the general result by applying the case $m = 1$ to the weak derivatives $D^\alpha u$, $|\alpha| \le m-1$. Let $x_0 \in \partial\Omega$, $\rho > 0$, ψ be as in the definition of Lipschitz domain given above. Since we can use the coordinates y mentioned in that definition we can without loss of generality assume that $x_0 = 0$ and that $\Omega \cap B_\rho(x_0) = \{x : x^n > \psi(x^1,\dots,x^{n-1})\} \cap B_\rho(0)$. For $0 < \tau < \sigma < \rho/4$, let $f : \mathbb{R} \to \mathbb{R}$ be a C^1 function with $f' \ge 0$ on $\mathbb{R}$, $f \equiv 0$ on $(-\infty, \frac14\tau]$, $f \equiv 1$ on $[\sigma+\tau,\infty)$ and $f' \equiv \sigma^{-1}$ on $[\tau,\sigma]$. (One easily checks that such a function exists; for example, we can mollify the piecewise linear function which is 0 on $(-\infty, \frac12\tau)$, 1 on $(\sigma + \frac12\tau, \infty)$, and linear of slope σ^{-1} on $[\frac12\tau, \sigma+\frac12\tau]$.) Also, we choose a C^1 function $\zeta : \mathbb{R}^n \to [0,1]$ with $\zeta \equiv 1$ on $B_{\rho/2}(0)$, $\zeta \equiv 0$ on $\mathbb{R}^n \setminus B_\rho(0)$ and $|D\zeta| \le 3\rho^{-1}$ on $\mathbb{R}^n$. Then it is straightforward to check that

$$f(x^n - \psi(x^1,\dots,x^{n-1}))\, u^2\zeta \in W_0^{1,1}(\Omega), \tag{1}$$

and hence by † (with $\varphi \equiv 1$ and $\alpha = (0,\dots,0,1)$) we have

$$\int_\Omega D_n\big(f(x^n - \psi(x')\, u^2\zeta\big) = 0,$$

where $x' = (x^1,\dots,x^{n-1})$. Using the calculus facts (ii), (iii) and the fact that $f' \equiv \sigma^{-1}$ on (τ,σ) (and $f' \ge 0$ everywhere) we deduce that

$$\sigma^{-1}\int_{W_{\sigma,\tau}} u^2\zeta \le -\int_\Omega f\,(x^n - \psi(x'))\,(2u\zeta D_n u + u^2 D_n\zeta),$$

where $W_{\sigma,\tau} = \{x : \psi(x') + \tau < x^n < \psi(x') + \sigma\}$, and hence

$$\sigma^{-1}\int_{W_{\sigma,\tau}\cap B_{\rho/2}(0)} u^2 \le \int_{\Omega\cap B_\rho(0)} (2|u|\,|Du| + 3\rho^{-1}u^2).$$

By using the Cauchy inequality $ab \le \frac12(a^2+b^2)$ on the right we have then that

$$\sigma^{-1}\int_{W_{\sigma,\tau}\cap B_{\rho/2}(0)} u^2 \le C\int_{\Omega\cap B_\rho(0)} (u^2 + |Du|^2),$$

where C depends only on ρ and not on σ or τ. With $\tau \downarrow 0$, this gives

$$\sigma^{-1}\int_{W_\sigma\cap B_{\rho/2}(0)} u^2 \le C\int_{\Omega\cap B_\rho(0)} (u^2 + |Du|^2),$$

where $W_\sigma = \{x : \psi(x') < x^n < \psi(x') + \sigma\}$. Now if $K \ge 1$ is any Lipschitz constant for ψ on $\{x' \in R^{n-1} : |x'| < \rho\}$, then

$$\partial\Omega \cap B_\rho(0) \equiv \operatorname{graph}\psi \cap B_\rho(0) \subset \{x : x^n \le \xi^n + K|x' - \xi'|\}, \quad \xi \in \partial\Omega\cap B_\rho(0),$$

which means that any point $\eta \in B_{\rho/2}(0) \cap \Omega \setminus W_\sigma$ has distance $\ge K^{-1}\sigma$ from $\partial\Omega$, so $W_\sigma \cap B_{\rho/2}(0) \supset U_{K^{-1}\sigma} \cap B_{\rho/2}(0)$. Therefore we have finally that for any $x_0 \in \partial\Omega$ there are ρ, $K > 0$ such that

$$\sigma^{-1}\int_{U_{K^{-1}\sigma}\cap B_{\rho/2}(x_0)} u^2 \le C\int_{\Omega\cap B_\rho(x_0)} (u^2 + |Du|^2), 0 < \sigma < \rho/4. \tag{2}$$

Notice that since $\partial\Omega$ is a compact set, we can select fixed constants K, $\rho > 0$ such that (2) is true for every $x_0 \in \partial\Omega$. Then by taking a finite cover for the compact set $\partial\Omega$ by balls $B_{\rho/2}(x_0)$ we conclude (i) of the lemma in the case $m = 1$ provided σ is sufficiently small, depending on Ω, as required.

The proof of (ii) of the lemma is similar except that when $u \in H_0^1(\Omega)$ (rather than merely in $H^1(\Omega)$) we can take the function f to satisfy $f' \le 0$, $f \equiv 1$ on (σ,∞), $f' = -\sigma^{-1}$ on (τ,σ); then we still have (1) as before because $u \in H_0^1(\Omega)$ (hence $u^2 \in W_0^{1,1}(\Omega)$), and by essentially the same argument as before we get an inequality like (2) above but with $U_\sigma \cap B_\sigma(x_0)$ as the domain of integration on the right side rather than $\Omega \cap B_\sigma(x_0)$. This completes the proof of Lemma 2.

The spaces $H^m(\Omega)$ are ideal for the study of PDE problems in that on the one hand they are Hilbert spaces (and hence Hilbert space techniques can be applied to the study of operators on them), while on the other hand for any $k \in \{1,2,\dots,\}$ and any $N > n/2 + k$ we have $H^N(\Omega) \subset C^k(\Omega)$. Another useful property of these spaces is that the inclusion $H_0^m(\Omega) \subset H_0^{m-1}(\Omega)$ is compact.

Precise versions of these latter claims are as follows:

Lemma 3. (Sobolev embedding theorem.) *If $k \in \{1,2,\dots,\}$ and if $N > n/2 + k$, then*

(i) $\quad H_0^N(\Omega) \subset C_0^k(\Omega)$ *and* $|u|_{C^k} \le C\|u\|_N$,

for all $u \in H_0^N(\Omega)$, where $C_0^k(\Omega)$ denotes the set of $C^k(\overline{\Omega})$ functions u with $D^\alpha u = 0$ on $\partial\Omega$ for all α with $|\alpha| \le k$, and

(ii) $\quad H_{\rm loc}^N(\Omega) \subset C^k(\Omega)$ *and* $|u|_{C^k(B_{\rho/2})} \le C\|u\|_{N,B_\rho}$

whenever $\overline{B}_\rho \subset \Omega$. Of course here $|u|_{C^k(A)} = \sum_{|\alpha|\le k}\sup_A |D^\alpha u|$.

Remark: Actually the inequality in (ii) holds with $|u|_{C^k(B_\rho)}$ in place of $|u|_{C^k(B_{\rho/2})}$, but this is more difficult to prove and we shall not need it here.

Lemma 4 (Rellich's Theorem.) *Suppose Ω is bounded, and $m \ge 1$. Then the inclusion $H_0^m(\Omega) \subset H_0^{m-1}(\Omega)$ is compact; that is, any bounded sequence $\{u_j\}$ in $H_0^m(\Omega)$ has a subsequence which converges in $H_0^{m-1}(\Omega)$. If Ω is a bounded Lipschitz domain then the same holds with $H^m(\Omega)$, $H^{m-1}(\Omega)$ in place of $H_0^m(\Omega)$, $H_0^{m-1}(\Omega)$ respectively.*

Remark: Analogous theorems to the above hold for the general Sobolev spaces (see e.g. Gilbarg & Trudinger Theorem 7.22 and the later discussion in these lectures), but for the moment the theorems stated above are adequate for our purposes.

Before we begin the proof, we need to recall some basic properties of the Fourier transform. We let $\mathcal{S}$ be the Schwarz class on $\mathbb{R}^n$; that is, the set of complex-valued $C^\infty(\mathbb{R}^n)$ functions u such that $x^\alpha D^\beta u$ is bounded on $\mathbb{R}^n$ for all multi-indices α, β. Of course using the inequalities

(1) $$C^{-1}(1+|x|)^{2m} \le \textstyle\sum_{|\alpha|\le m}(x^\alpha)^2 \le C(1+|x|)^{2m}, \quad x \in \mathbb{R}^n,$$

where C depends only on m, n, it then follows that $u \in \mathcal{S}$ if and only if $(1+|x|)^m D^\beta u$ is bounded for each $m \ge 0$ and each multi-index β. (Note that $x^\alpha = 1$ when $\alpha = 0$.)

For $u \in \mathcal{S}$ the Fourier transform $\hat{u}$ is defined by

(a) $$\hat{u}(\xi) = (2\pi)^{-n/2}\int_{\mathbb{R}^n} e^{-ix\cdot\xi}u(x)\,dx.$$

Notice that (since we can differentiate under the integral in (a)), then $\hat{u} \in C^\infty(\mathbb{R}^n)$ and

(b) $$D^\alpha\hat{u}(\xi) = (-i)^{|\alpha|}\widehat{x^\alpha u}(\xi),$$

and also (by using (a) with $D^\alpha u$ in place of u and integrating by parts) we have

(c) $$\widehat{D^\alpha u} = i^{|\alpha|}\xi^\alpha\hat{u}, \quad u \in \mathcal{S}.$$

Notice that the validity of (c) for $u \in \mathcal{S}$ follows from the fact that the integration by parts formula $\int \psi D^\alpha\varphi = (-1)^{|\alpha|}\int\varphi D^\alpha\psi$ is valid for any $\psi \in \mathcal{S}$ and any $\varphi \in C^\infty(\mathbb{R}^n)$ with $|D^\beta\varphi(x)| \le C(1+|x|)^m\ \forall\beta$, where the constants C, m may depend on β. (See Exercise 5.2(i).)

It of course follows directly from (b), (c) that $u \in \mathcal{S} \Rightarrow \hat{u} \in \mathcal{S}$, and by Fubini's Theorem it is straightforward to check the identity

(d) $$\int_{\mathbb{R}^n} u(y)\hat{v}(y)\,dy = \int_{\mathbb{R}^n}\hat{u}(y)v(y)\,dy, \quad u, v \in \mathcal{S}.$$

Using this with the special choice $v(y) = e^{-\varepsilon|y|^2}$ (then by direct computation we see $\hat{v}(\xi) = (2\varepsilon)^{-n/2}e^{-|\xi|^2/(4\varepsilon)}$—see Exercise 5.4 below) and letting $\varepsilon \downarrow 0$ we obtain $u(0) = (2\pi)^{-n/2}\int\hat{u}(\xi)\,d\xi$. Applying this with $u_x(y) \equiv u(x+y)$ and noting that $\widehat{u_x}(\xi) = e^{ix\cdot\xi}\hat{u}(\xi)$, we obtain the Fourier inversion formula

(e) $$u(x) = (2\pi)^{-n/2}\int_{\mathbb{R}^n} e^{ix\cdot\xi}\hat{u}(\xi)\,d\xi, \quad u \in \mathcal{S}.$$

Notice that this says $\hat{\hat{u}}(x) = u(-x)$, so that, up to composition with the map $u \mapsto \check{u}$, $\check{u}(x) \equiv u(-x)$, the Fourier transform is its own inverse on $\mathcal{S}$; in particular it is a bijection of $\mathcal{S}$ onto itself. By replacing v in (d) by the complex conjugate of $\hat{v}$ (which by formulae (a) and (e) is the same as the inverse transform of the complex conjugate $\overline{v}$ of v), we deduce the Parseval formula

(f) $$\langle u, v\rangle_{L^2(\mathbb{R}^n)} = \langle\hat{u},\hat{v}\rangle_{L^2(\mathbb{R}^n)}, \quad u, v \in \mathcal{S}.$$

Thus the Fourier transform provides an isometric isomorphism (with respect to the $L^2(\mathbb{R}^n)$ inner product) of $\mathcal{S}$ onto itself. Of course since $C_c^\infty(\mathbb{R}^n)$ (and hence $\mathcal{S}$) is dense in $L^2(\mathbb{R}^n)$, we then see that the Fourier transform extends uniquely as an isometry (i.e. $\|u\|_{L^2} = \|\hat{u}\|_{L^2}$) of $L^2(\mathbb{R}^n)$ onto $L^2(\mathbb{R}^n)$. (This is Plancherel's Theorem.)

From (b), (c) and (f) we have

(g) $$\|u\|^2_{m,\Omega} = \int_\Omega \textstyle\sum_{|\alpha|\le m}(\xi^\alpha)^2|\hat{u}(\xi)|^2\,d\xi, \quad u \in C_c^\infty(\Omega),$$

and hence by (1)

(h)

$\|u\|_{m,\Omega}$ and $\left(\int_\Omega (1+|\xi|)^{2m}|\hat{u}(\xi)|^2\,d\xi\right)^{1/2}$ are equivalent norms on $C_c^\infty(\Omega)$.

Then since $\mathcal{S}$ is dense in $L^2(\mathbb{R}^n)$, and since $H_0^m(\Omega)$ is just the closure in $L^2(\Omega)$ of $C_c^\infty(\Omega)$ with respect to the norm $\|\ \|_{m,\Omega}$, the following alternative characterization of the space $H_0^m(\Omega)$ follows immediately.

Lemma 5. *For any domain $\Omega \subset \mathbb{R}^n$ (bounded or not), we have*

$$H_0^m(\Omega) = \overline{C_c^\infty(\Omega)} \subset \{u \in L^2(\Omega) : (1+|\xi|)^m\hat{u}(\xi) \in L^2(\mathbb{R}^n)\},$$

where the closure is taken with respect to the norm $\|(1+|\xi|)^m\hat{u}\|$.

Remarks: (1) Notice that of course in computing $\hat{u}$ for $u \in L^2(\Omega)$, we first extend u to all of $\mathbb{R}^n$ by defining $u = 0$ on $\mathbb{R}^n \setminus \Omega$; then $u \in L^2(\mathbb{R}^n)$ and

$\hat{u} \in L^2(\mathbb{R}^n)$ makes sense by the above discussion of Fourier transform on $L^2(\mathbb{R}^n)$.

(2) In case Ω is a bounded Lipschitz domain, the inclusion of the above lemma can be replaced by equality. Since we shall not need this fact we omit the details—but the proof is straightforward. See problem 5.12 below.

Proof of Sobolev Embedding Theorem: We first prove statement (i) of the theorem. Since $H_0^N(\Omega)$ is the closure of $C_c^\infty(\Omega)$ with respect to the norm $\|u\|_{N,\Omega}$, and since the space $C_0^k(\Omega)$ is complete with respect to the norm $|\ |_{C^k}$ in the statement of the theorem, it is evidently enough to establish the inequality in (i) for $u \in C_c^\infty(\Omega)$.

If $u \in C_c^\infty(\Omega)$, first use the formulae (c), (e) above to give

$$D^\alpha u(x) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{ix\cdot\xi} i^{|\alpha|} \xi^\alpha \hat{u}(\xi)\, d\xi$$

for any multi-index α. Then if $|\alpha| \le k$ we get

$$\begin{aligned}|D^\alpha u(x)| &\le C \int_{\mathbb{R}^n} (1+|\xi|)^k |\hat{u}(\xi)|\, d\xi \\ &= C \int_{\mathbb{R}^n} ((1+|\xi|)^{k-N}(1+|\xi|)^N |\hat{u}(\xi)|)\, d\xi \\ &\le C \Big(\int_{\mathbb{R}^n} (1+|\xi|)^{2k-2N}\, d\xi\Big)^{1/2} \|u\|_N \text{ (by (h) above).}\end{aligned}$$

Since $2N - 2k > n$, the first integral on the right is finite, so

$$|D^\alpha u(x)| \le C\|u\|_{N,\Omega},\ |\alpha| \le k,\ x \in \Omega,$$

which evidently implies the inequality in (i).

To prove (ii), first take any ball $B_\rho(y)$ with $\overline{B}_\rho(y) \subset \Omega$, and let $\psi(x) \equiv \varphi(\rho^{-1}(x-y))$, where φ is any non-negative $C_c^\infty(\mathbb{R}^n)$ function with support in $B_1(0)$ and with $\varphi \equiv 1$ in $B_{1/2}(0)$. Then

$$(*) \qquad \text{support } \psi \subset B_\rho(y) \text{ and } \sup |D^\alpha \psi| \le c_\alpha \rho^{-|\alpha|}$$

for each multi-index α. In view of Lemma 1, we know that ψu_σ converges in $H^m(\Omega)$ to ψu as $\sigma \downarrow 0$. Since ψu_σ is in $C_c^\infty(\Omega)$, we thus have $\psi u \in H_0^m(\Omega)$, and hence part (i) above is applicable to ψu. Also $D^\alpha(\psi u)$ $(= \lim D^\alpha(\psi u_\sigma))$ can be computed according to the Leibniz formula (see (ii) above) as

$$D^\alpha(\psi u) = \textstyle\sum_{\gamma+\delta=\alpha} \frac{\alpha!}{\gamma!\,\delta!} (D^\gamma \psi)(D^\delta u),$$

and hence the required result follows from $(*)$.

Proof of Rellich's Theorem: Evidently it is enough to prove the case $m = 1$, because the general case then follows by induction on m, applying the case $m = 1$ to the sequences $D^\alpha u_j$, where $|\alpha| \le m-1$.

Thus let $\{u_j\}$ be any bounded sequence in $H_0^1(\Omega)$, and note that by definition of $H_0^1(\Omega)$ we can find $v_j \in C_c^\infty(\Omega)$ such that

$$(1) \qquad \|u_j - v_j\|_{1,\Omega} \le 1/j$$

for each $j \ge 1$. We can of course think of v_j as a $C_c^\infty(\mathbb{R}^n)$ function with support $\subset \Omega$. Notice that $\|v_j\|_{1,\mathbb{R}^n}$ is bounded independent of j, and hence for any fixed $\sigma > 0$ we have, by the Schwarz inequality and the fact that $|\varphi^{(\sigma)}| \le C\sigma^{-n}$,

$$|(v_j)_\sigma(x)| \le \Big|\int_{\mathbb{R}^n} \varphi^{(\sigma)}(x-y) v_j(y)\, dy\Big| \le C(\sigma) \quad \forall x \in \mathbb{R}^n,$$

where $C(\sigma)$ is a constant depending on σ but not on j. We similarly have

$$|D(v_j)_\sigma(x)| = |(Dv_j)_\sigma(x)| \le \Big|\int_{\mathbb{R}^n} (D\varphi^{(\sigma)}(x-y)) v_j(y)\, dy\Big| \le C(\sigma), \forall x \in \mathbb{R}^n.$$

Thus for fixed σ the sequence $\{(v_j)_\sigma\}_{j=1,2,\ldots}$ is uniformly bounded and equicontinuous, and hence by the Arzela theorem there is a subsequence $\{j_k\}$ of $\{j\}$ with

$$(2) \qquad (v_{j_k})_\sigma \to w$$

uniformly on $\mathbb{R}^n$, where w depends on σ.

Notice that we also have

$$\begin{aligned}(v_j)_\sigma(x) - v_j(x) &= \int_{\mathbb{R}^n} \varphi^{(\sigma)}(x-y)\big(v_j(y) - v_j(x)\big)\, dy \\ &= \int_{B_1(0)} \varphi(y)\big(v_j(x-\sigma y) - v_j(x)\big)\, dy.\end{aligned}$$

Since by elementary calculus we have $v_j(x-\sigma y) - v_j(x) = -\sigma \int_0^1 y \cdot Dv_j(x - t\sigma y)\, dt$, this gives, after squaring and integrating with respect to x (and using the Cauchy-Schwarz inequality on the right),

$$\begin{aligned}\|(v_j)_\sigma - v_j\|_{0,\Omega} &\le C\sigma \|Dv_j\|_{L^2(\mathbb{R}^n)} \le C\sigma\|v_j\|_{1,\mathbb{R}^n} \qquad (3) \\ &\le C_1\sigma,\end{aligned}$$

where C_1 is independent of j and σ.

Now let $\varepsilon > 0$ be given, and choose $\sigma = (1 + C_1)^{-1}\varepsilon$ in the above. According to (1), (2), (3) and the triangle inequality we then have $N = N(\varepsilon)$ such that

$$\|u_{j_k} - u_{j_\ell}\|_{0,\Omega} \le \varepsilon \quad \forall k, \ell \ge N.$$

Notice that the subsequence j_k also depends on ε; to show that we can get a subsequence independent of ε, we use the above with $\varepsilon = 1/q$, $q = 1, 2, \ldots$, and inductively select for each $q \ge 1$ a subsequence such that for $q \ge 2$ the subsequence contained in the previous choice (corresponding to $q - 1$). This then gives a convergent subsequence of u_j as required by taking a diagonal process; i.e. if $\{u_{q,j}\}$ is the q^{th} subsequence, then we take $\{u_{q,q}\}$ as the final subsequence.

This completes the proof of Rellich's Theorem for the case $u \in H_0^m(\Omega)$; the proof for the case $u \in H^m(\Omega)$ when Ω is a bounded Lipschitz domain uses this case together with Lemma 2 above, and is left as an exercise (see problem 5.10 below).

The following interpolation inequality will be used frequently in later discussions:

Lemma 6. *If $\varepsilon > 0$, there is a constant $C = C(n, \varepsilon)$ such that*

$$\|u\|_{m-1,\Omega} \le \varepsilon \|u\|_{m,\Omega} + C\|u\|_{0,\Omega}, \quad u \in H_0^m(\Omega),$$

where $\| \|_{0,\Omega}$ denotes the L^2 norm on Ω. If Ω is a bounded Lipschitz domain, then the same holds for $u \in H^m(\Omega)$ and with $C = C(\Omega, \varepsilon)$.

Remark: Analogous inequalities apply to the more general spaces $W_0^{m,p}(\Omega)$, but we shall not need these.

Proof of Lemma 6: For $u \in H_0^m(\Omega)$ the inequality is a direct consequence of (h) above, together with the fact that, for any $\varepsilon > 0$, $(1 + |\xi|)^{m-1} \le \varepsilon(1 + |\xi|)^m + C\varepsilon^{-m}$ for suitable $C = C(m)$, as one easily checks.

To prove the case when $u \in H^m(\Omega)$ with Ω a bounded Lipschitz domain, we use the above case (i.e. the $H_0^m(\Omega)$ inequality) together with Lemma 2. (See problem 5.9 below.)

We shall also need the following lemma concerning the convergence of difference quotients to the corresponding weak derivatives. The notation is as follows:

If $j \in \{1, 2, \ldots, n\}$ and $h \in \mathbb{R} \setminus \{0\}$, define

$$\Delta_h^{(j)} u(x) = \frac{u(x + he_j) - u(x)}{h}.$$

Notice that this is well-defined on

$$\Omega_{|h|} \equiv \{x \in \Omega : \operatorname{dist}(x, \partial\Omega) > |h|\},$$

provided u is well-defined on Ω.

Lemma 7. *If $u \in L^2_{\text{loc}}(\Omega)$, and if*

$$(*) \qquad \limsup_{h \downarrow 0} \|\Delta_h^{(j)} u\|_{L^2(K)} \le c_K \ (< \infty)$$

for each compact $K \subset \Omega$. Then u has a weak derivative $D_j u \in L^2_{\text{loc}}(\Omega)$, and $\Delta_h^{(j)} u \rightharpoonup D_j u$ in the weak sense that $\langle \Delta_h^{(j)} u, \varphi \rangle_{L^2} \to \langle D_j u, \varphi \rangle_{L^2}$ $\forall \varphi \in C_c^\infty(\Omega)$. Further, if $(*)$ holds with $c_K = c$, c independent of K, then $D_j u \in L^2(\Omega)$.

The proof is left as an exercise, based on the definition of weak derivative and the general fact that a bounded sequence in a Hilbert space has a weakly convergent subsequence. (See Exercise 5.3 below.) Note again that there is a version of this in an L^p setting for any $p \ge 1$.

LECTURE 5 PROBLEMS

5.1 (i) Prove the claim made in lecture that $W^{m,p}(\Omega)$ is a Banach space and $H^m(\Omega)$ is a Hilbert space. (The main point is completeness.)

(ii) Prove that $W^{m,p}(\Omega)$ is separable and reflexive if $p > 1$, and separable if $p = 1$. Hint for (ii): There is a natural map from $W^{m,p}(\Omega)$ to the product $\times_{|\alpha|\le m} L^p(\Omega)$, defined by $u \mapsto \{D^\alpha u\}_{|\alpha|\le m}$. Begin by showing that the range of this map is a closed subspace of the product space, and that the map is an isometric isomorphism (assuming the product space is equipped with the norm $\|\{\varphi_\alpha\}_{|\alpha|\le m}\| = \sum_{|\alpha|\le m}\|\varphi_\alpha\|_{L^p(\Omega)}$).

5.2 (i) Prove that if $\psi \in \mathcal{S}$, and if $\varphi \in C^\infty(\mathbb{R}^n)$ with $|D^\beta\varphi(x)| \le C(\beta)(1+|x|)^{m(\beta)}\ \forall x,\ \beta$, then the integration by parts formula

$$\int_{\mathbb{R}^n} \psi D^\alpha \varphi = (-1)^{|\alpha|}\int_{\mathbb{R}^n} \varphi D^\alpha \psi$$

holds for each multi-index α.

(ii) Prove (as claimed) that if Ω is bounded, if $u \in W_0^{m,p}(\Omega)$, and if $|\alpha| \le m$, then $\int_\Omega \varphi D^\alpha u = (-1)^{|\alpha|}\int_\Omega u D^\alpha\varphi\ \ \forall \varphi \in C^\infty(\Omega)$ with $D^\beta\varphi$ bounded on Ω, $|\beta| \le m$.

5.3 Prove Lemma 7.

5.4 Check the claim made above that

$$v(x) = e^{-\varepsilon|x|^2} \Rightarrow \hat{v}(\xi) = (2\varepsilon)^{-n/2}e^{-|\xi|^2/4\varepsilon}.$$

Hint: $(2\pi)^{n/2}\hat{v} = \int e^{-\varepsilon|x|^2 - ix\cdot\xi}\,dx$ which (by completing the square in the exponent) is $e^{-|\xi|^2/4\varepsilon}\prod_{j=1}^n \int e^{-\varepsilon(x^j - i\xi^j/2\varepsilon)^2}\,dx^j$; keep in mind the standard integral $\int_{\mathbb{R}} e^{-t^2}\,dt = \sqrt{\pi}$.

5.5 If $N > n/2 + k$, then improve the Sobolev embedding theorem proved in lecture by showing that $H_0^N(\Omega) \subset C_0^{k,\beta}(\Omega)$ and that $|u|_{C^{k,\beta}(\Omega)} \le c\|u\|_{N,\Omega}$ for any $\beta < \min\{1, N-k-n/2\}$.

(Here $|u|_{C^{k,\beta}}(\Omega) = |u|_{C^k(\Omega)} + \sum_{|\alpha|=k}\sup_{x,y\in\Omega, x\ne y}|x-y|^{-\beta}|D^\alpha u(x) - D^\alpha u(y)|$.)

Hence conclude that the embedding $H_0^N(\Omega) \subset C_0^k(\Omega)$ is compact in the sense that if $\{u_j\}$ is bounded in $H_0^N(\Omega)$, then some subsequence converges in $C^k(\Omega)$.

5.6 (Absolute continuity properties of Sobolev functions.) Suppose $v \in W^{1,1}(\Omega)$, where $\Omega = (a_1,b_1)\times\cdots\times(a_n,b_n) \subset \mathbb{R}^n$ for some given $a_j < b_j,\ j = 1,\ldots,n$.

(i) Show that there is a sequence $\sigma_j \downarrow 0$ such that $D_1 v_{\sigma_j}(\cdot, x')$ converges in $L^1(a_1,b_1)$ and pointwise a.e. on (a_1,b_1) for almost all $x' = (x^2,\ldots,x^n) \in (a_2,b_2)\times\cdots\times(a_n,b_n)$.

(ii) If representatives $\overline{v},\ \overline{D_1 v}$ are defined for the L^1 classes of $v,\ D_1 v$ by setting $\overline{f}(x) = \lim_{j\to\infty} f_{\sigma_j}(x)$ at all points $x \in \Omega$ where the limit exists, and $\overline{f}(x) = 0$ otherwise, where $f = v,\ D_1 v$ respectively and where σ_j is as in (i), prove that for Lebesgue almost all $x' = (x^2,\ldots,x^n) \in (a_2,b_2)\times\cdots\times(a_n,b_n)$ we have the identity

$$\overline{v}(t_2,x') - \overline{v}(t_1,x') = \int_{t_1}^{t_2}\overline{D_1 v}(t,x')\,dt, \text{ for all } t_1,\ t_2 \in (a_1,b_1)$$

5.7 Prove that if $\Omega \subset \mathbb{R}^n$ is bounded, if $\mathbb{R}^n\setminus\Omega = \overline{\mathbb{R}^n\setminus\overline{\Omega}}$, and if $u \in C^{m-1}(\overline{\Omega}) \cap H_0^m(\Omega)$, then $D^\alpha u \equiv 0$ on $\partial\Omega$ for $|\alpha| \le m-1$.

Hint: Show first that if u is extended to give a function v on all of $\mathbb{R}^n$ by taking $v \equiv 0$ outside Ω, then $v \in H^m(\mathbb{R}^n)$. Then use 5.6 above, noting also that if $v_j \in C_c^\infty(\mathbb{R}^n)$ with support in Ω, and if $v_j \to v$ in $H^m(\mathbb{R}^n)$, then $Dv_j(\cdot,x')$ converges in $L^1(a_1,b_1)$ to $\overline{Dv}(\cdot,x')$ for almost all $x' = (x^2,\ldots,x^n) \in (a_2,b_2)\times\cdots\times(a_n,b_n)$. (Here the notation is as in 5.6 with $(a_1,b_1)\times\cdots\times(a_n,b_n)$ an arbitrary rectangle in $\mathbb{R}^n$.) Note: The condition $\mathbb{R}^n\setminus\Omega = \overline{\mathbb{R}^n\setminus\overline{\Omega}}$ is obviously satisfied if Ω is a Lipschitz domain in the sense of the above lecture.

5.8 (i) Prove that $f \in H^k(\Omega) \Rightarrow \|\Delta_j^{(h)} f\|_{k-1,\Omega_{|h|}} \le \|f\|_{k,\Omega}$ for any $h \in \mathbb{R}\setminus\{0\}$. Here $\Delta_j^{(h)}$ are the difference quotient operators, given by $\Delta_j^{(h)} g(x) = h^{-1}(g(x+he_j) - g(x))$ for $h \ne 0$.

(ii) Prove that $f \in W^{k,\infty}(\Omega) \Rightarrow |\Delta_j^{(h)} D^\gamma f(x)| \le \|f\|_{k,\infty,\Omega}$ for any $x \in \Omega_{|h|}$, $|\gamma| \le k-1$, and $h \in \mathbb{R}\setminus\{0\}$.

5.9 Give the proof of the interpolation inequality of Lemma 6 for the case when $u \in H^m(\Omega)$ and Ω is a bounded Lipschitz domain. (Use the result for the case $H_0^m(\Omega)$ applied to φu where $\varphi \in C_c^\infty(\Omega)$ with $\varphi \equiv 1$ on Ω_σ. Then use Lemma 2.)

5.10 Give the proof of the second part of Lemma 4 (i.e., Rellich's theorem for $H^m(\Omega)$ in case Ω is a bounded Lipschitz domain). Hint: For a given bounded sequence $u_j \in H^m(\Omega)$, apply the first part of Lemma 4 (which we already proved) to the sequence φu_j, where $\varphi \in C_c^\infty(\Omega)$ with $\varphi \equiv 1$ on Ω_σ. Then use Lemma 2.

5.11 Prove that if $f \in L^p_{\rm loc}(\Omega)$, then $f_\sigma \to f$ locally with respect to the L^p norm in Ω. Hint: Begin by using the fact that the continuous functions are dense in L^p to prove that $\lim_{s\downarrow 0}\sup_{|y|\le s}\int_{B_\rho(x_0)}|f(x+y)-f(x)|^p = 0$ for any ball $B_\rho(x_0)$ with $\overline{B}_\rho(x_0)\subset\Omega$.

5.12 (i) Prove that $H_0^m(\Omega)=\{u|\Omega : u\in H^m(\mathbb{R}^n) \text{ and } u=0 \text{ a.e. on } \mathbb{R}^n\setminus\Omega\}$ if Ω is a bounded Lipschitz domain.

(ii) Use the result of (i) and the definition of weak derivatives to prove that equality holds in the inclusion of Lemma 5 in case Ω is a bounded Lipschitz domain. Hint for (ii): Show that if $u\in L^2(\mathbb{R}^n)$ with compact support and if $(1+|\xi|)^m\hat{u}\in L^2(\mathbb{R}^n)$, then $r_\alpha(x)=\int_{\mathbb{R}^n} i^{|\alpha|}e^{ix\cdot\xi}\xi^\alpha\hat{u}(\xi)\,d\xi$ is the weak derivative $D^\alpha u$ of u for $|\alpha|\le m$.

Lecture 6

Weak Solutions in Sobolev Space and Local Regularity

Here we consider weak solutions u in Sobolev space $H^m_{\text{loc}}(\Omega)$ of equations of the form

$$(*)\qquad \textstyle\sum_{|\alpha|,|\beta|\le m} D^\beta(a_{\alpha\beta}D^\alpha u)=\sum_{|\beta|\le m}D^\beta f_\beta,$$

where $\{f_\beta\}_{|\beta|\le m}$ are given $L^2_{\text{loc}}(\Omega)$ functions and $a_{\alpha\beta}$ are given $L^\infty_{\text{loc}}(\Omega)$ (i.e. locally bounded) functions.

Our reason for picking this class of equations is that on the one hand it is sufficiently general to include the linearization of the Euler-Lagrange operator of a general m-th order functional (see problem 6.1 below), but at the same time is a rather convenient class to work with. Certainly it is a sufficiently general class for our later purposes in these lectures.

We say $u\in H^m_{\text{loc}}(\Omega)$ is a weak solution of $*$ if

$$(**)\qquad \int_\Omega \textstyle\sum_{|\alpha|,|\beta|\le m}(-1)^{|\beta|}a_{\alpha\beta}D^\alpha u\,D^\beta\zeta=\int_\Omega\sum_{|\beta|\le m}(-1)^{|\beta|}f_\beta D^\beta\zeta$$

for each $\zeta\in C_c^\infty(\Omega)$. Of course this definition is reasonable in the sense that if $u\in C^{2m}(\Omega)$, $a_{\alpha\beta},\ f_\beta\in C^m(\Omega)$, and if $(**)$ holds, then u satisfies $*$ in the classical (i.e. pointwise) sense, as one checks by integration by parts. Notice also that if $\alpha_{\alpha\beta}\in C^m(\Omega)$, then this definition of weak solution in $H^m_{\text{loc}}(\Omega)$ is equivalent to the requirement that $u\in H^m_{\text{loc}}(\Omega)$ and that u is an $L^2_{\text{loc}}(\Omega)$ weak solution in the sense of our previous discussion in Lecture 2.

Notice that $(**)$ must hold for any ζ of the form $\zeta = \varphi h$, where $\varphi \in C_c^\infty(\Omega)$ and $h \in H^m_{\text{loc}}(\Omega)$, because we can write such ζ as $\lim_{\sigma\downarrow 0} \varphi h_\sigma$ with respect to the $H^m(\Omega)$ norm; since $\varphi h_\sigma \in C_c^\infty(\Omega)$ for σ sufficiently small, this gives $(**)$ for $\zeta = \varphi h$ as claimed. Then, with this choice of ζ, keeping in mind the Leibniz formula $D^\beta(\varphi u) = \sum_{\gamma+\delta=\beta} \frac{\beta!}{\gamma!\,\delta!} D^\gamma u D^\delta \varphi$, we see that $(**)$ implies

$$\int_\Omega \textstyle\sum_{|\alpha|,|\beta|\le m}((-1)^{|\beta|} a_{\alpha\beta} D^\alpha u(\sum_{\gamma+\delta=\beta} \dfrac{\beta!}{\gamma!\,\delta!} D^\gamma u D^\delta \varphi)) = \int_\Omega \sum_{|\beta|\le m}((-1)^{|\beta|} f_\beta \sum_{\gamma+\delta=\beta}(\dfrac{\beta!}{\gamma!\,\delta!} D^\gamma u D^\delta \varphi))$$

for each $\varphi \in C_c^\infty(\Omega)$.

To get further, we need to impose an <u>ellipticity</u> <u>condition</u> on the coefficients $a_{\alpha\beta}$ corresponding to $|\alpha|, |\beta| = m$. We shall consider more refined conditions below, but for the moment we impose the strongest ellipticity condition, which is the following:

(E) $$\textstyle\sum_{|\alpha|,|\beta|=m} a_{\alpha\beta}(x)\lambda_\alpha\lambda_\beta \ge \mu \sum_{|\alpha|=m}(\lambda_\alpha)^2$$

for every $x \in \Omega$ and every collection $\{\lambda_\alpha\}_{|\alpha|=m}$ of real numbers, where $\mu > 0$ is fixed independent of x, λ_α.

From now on we work in a ball $B_R(x_0)$ with $\overline{B}_R(x_0) \subset \Omega$, and we assume the explicit boundedness restriction on the coefficients $a_{\alpha\beta}$ as follows:

(B) $$|a_{\alpha\beta}(x)| \le M, \quad \forall x \in B_R(x_0),\ |\alpha|,\ |\beta| \le m.$$

Then using (E), (B) in $*$, we conclude that

(1) $$\mu \int_\Omega \textstyle\sum_{|\alpha|=m}(D^\alpha u)^2\varphi \le M m! \int_\Omega (\sum_{|\alpha|\le m}|D^\alpha u|)(\sum_{|\delta|\le m, |\gamma|\le m-1}(|D^\gamma u||D^\delta\varphi|)) + m! \int_\Omega \sum_{|\beta|\le m}(|f_\beta| \sum_{\gamma+\delta=\beta}(|D^\gamma u||D^\delta \varphi|)).$$

Let $\theta \in (0,1)$, and choose φ to be a "cut-off function" for the balls $B_{\theta R}(x_0)$ and $B_R(x_0)$ in the usual sense. That is, $\varphi(x) \equiv \psi(R^{-1}(x-y))$, where $\psi \in C_c^\infty(B_1(0))$, $\psi \equiv 1$ in $B_\theta(0)$ and $\psi \ge 0$ everywhere, and $|D^\alpha \psi| \le C(\alpha)(1-\theta)^{-|\alpha|}$. Then

(2) $$\begin{aligned} &\varphi \in C_c^\infty(\Omega),\ 1 \ge \varphi \ge 0,\ \varphi \equiv 0 \text{ outside } B_R(x_0) \\ &\varphi \equiv 1 \text{ on } B_{\theta R}(x_0),\ \sup|D^\delta\varphi| \le C((1-\theta)R)^{-|\delta|} \end{aligned}$$

for any multi-index δ with $|\delta| \le m$, where $C = C(\delta)$.

With such a φ we have $\sup|D^\delta(\varphi^{2m})| \le C((1-\theta)R)^{-|\delta|}\varphi^m$ for any multi-index δ with $|\delta| \le m$, as one easily checks by induction on m; here C depends only on m. Then, if we substitute φ^{2m} in place of φ in (1), we obtain (in view of (2))

(3) $$\int_{B_R(x_0)} \textstyle\sum_{|\alpha|=m}|D^\alpha u|^2\varphi^{2m} \le C \int_{B_R(x_0)} (\sum_{|\alpha|\le m}|D^\alpha u|\varphi^m)(\sum_{|\gamma|\le m-1}|D^\gamma u| + \sum_{|\beta|\le m}|f_\beta|),$$

where C depends only on M, μ, θ, R, m, and n.

By virtue of the Cauchy inequality $ab \le \varepsilon a^2/2 + b^2/(2\varepsilon)$ (which just says $(\sqrt{\varepsilon}a - b/\sqrt{\varepsilon})^2 \ge 0$, valid for any real numbers a, b and any $\varepsilon > 0$), we deduce directly from (3) that

$$\int_{B_R(x_0)} \textstyle\sum_{|\alpha|\le m}|D^\alpha u|^2\varphi^{2m} \le C \int_{B_R(x_0)} (\sum_{|\alpha|\le m-1}|D^\alpha u|^2 + \sum_{|\beta|\le m}|f_\beta|^2),$$

where C is a constant depending only on M, μ, m, θ, R, and n.

Since $\varphi \equiv 1$ in $B_{\theta R}(x_0)$ and $\varphi \equiv 0$ outside $B_R(x_0)$, we have thus established the following:

Lemma 1. *If $u \in H^m_{\text{loc}}(\Omega)$ is a weak solution of the equation $*$, if (E), (B) hold and if $\overline{B}_R(x_0) \subset \Omega$, then for each $\theta \in (0,1)$ we have*

$$\|u\|_{m,B_{\theta R}(x_0)} \le C(\|u\|_{m-1,B_R(x_0)} + \textstyle\sum_{|\beta|\le m}\|f_\beta\|_{0,B_R(x_0)}),$$

where $C = C(R, M, \mu, m, n, \theta)$.

We now want to show that the inequality of the Lemma 1 implies good regularity results for weak solutions u, provided that we are willing to assume suitable differentiability properties concerning the coefficients $a_{\alpha\beta}$ and the given functions f_β.

We in fact henceforth assume that $k \ge 1$ and that, in addition to (E), (B), we have $a_{\alpha\beta} \in W^{k,\infty}(\Omega)$, $f_\beta \in H^k(\Omega)$ and that

(B_k) $$|D^\gamma a_{\alpha\beta}(x)| \le M \text{ a.e. } x \in B_R(x_0),\ |\gamma| \le k,$$

subject to which we prove the following:

Theorem 1. *If $u \in H^m(B_R(x_0))$ is a weak solution of $*$, if $k \geq 0$, and if (E), (B_k) hold, then $u \in H^{m+k}_{\text{loc}}(B_R(x_0))$ and*

$$\|u\|_{m+k, B_{\theta R}(x_0)} \leq C\left(\|u\|_{m-1, B_R(x_0)} + \textstyle\sum_{|\beta|\leq m} \|f_\beta\|_{k, B_R(x_0)}\right)$$

for any $\theta \in (0,1)$, where C is a constant depending only on n, m, k, θ, M, μ.

Remark: Actually we will show at the end of this lecture, using an additional interpolation inequality argument, that the inequalities in Theorem 1 (and also Corollary 1 below) can be improved to the extent that we can replace $\|u\|_{m-1,B_R(x_0)}$ on the right with the L^2 norm $\|u\|_{0,B_R(x_0)}$; notice that the result is unchanged in the special case $m = 1$.

Notice that by virtue of the Sobolev embedding theorem, this gives us the following corollary:

Corollary 1. *If the hypotheses are as in Theorem 1 with $m + k > n/2 + \ell$ (where $\ell \in \{0, 1, \ldots\}$) then $u \in C^\ell(\Omega)$ and*

$$|u|_{C^\ell(B_{\theta R}(x_0))} \leq C\left(\|u\|_{m-1, B_R(x_0)} + \textstyle\sum_{|\beta|\leq m} \|f_\beta\|_{k, B_R(x_0)}\right),$$

with $C = C(m, n, k, R, \theta, M, \mu)$. In particular if $a_{\alpha\beta}$, $f_\alpha \in C^\infty(\Omega)$ then $u \in C^\infty(\Omega)$.

Before beginning the proof of Theorem 1 , we need to make some observations about the difference quotient operators $\Delta_j^{(h)}$ introduced in Lecture 5. We note that for $h \neq 0$ the following properties hold for $f, g \in L^1(B_R(x_0))$:

(a) $$\Delta_j^{(h)}(f+g) = \Delta_j^{(h)} f + \Delta_j^{(h)} g \text{ on } B_{R-|h|}(x_0)$$

(b) $$\Delta_j^{(h)}(fg) = g\Delta_j^{(h)} f + \tilde{f}\Delta_j^{(h)} g \text{ on } B_{R-|h|}(x_0),\ \ f, g \in L^1(B_R(x_0))$$

where $\tilde{f}(x) \equiv f(x + he_j)$,

(c) $$D^\alpha \Delta_j^{(h)} f = \Delta_j^{(h)} D^\alpha f, \text{ on } B_{R-|h|}(x_0),\ |\alpha| \leq m,$$

provided $f \in H^m(B_R(x_0))$.

We also have the "integration by parts" formula

(d) $$\int_{B_R(x_0)} f\Delta_j^{(h)} g\, dx = -\int_{B_R(x_0)} g\Delta_j^{(-h)} f\, dx,$$

provided $f, g \in L^1(B_R(x_0))$ and the product f vanishes outside $B_{R-|h|}(x_0)$; this is proved simply by making the change of variable $y = x + he_j$ in the integral $\int_{B_R(x_0)} f(x)g(x + he_j)\, dx$ (which is one part of the expression on the left of (d)).

We note that by problem 5.8(ii) (B_k) implies that

(e) $$|\Delta_j^{(h)} D^\gamma a_{\alpha\beta}(x)| \leq M \text{ a.e. } x \in \Omega_{|h|},\ |\gamma| \leq k-1;$$

also, (see problem 5.8(i) above) we have the useful general fact that if $\ell \geq 1$, if $v \in H^\ell(B_R(x_0))$, then

(f) $$\|\Delta_j^{(h)} v\|_{\ell-1, B_{R-|h|}(x_0)} \leq \|v\|_{\ell, B_{R+|h|}(x_0)}.$$

Proof of Theorem 1: We first describe the proof in the case $k = 1$. (The general proof will go by induction on k.) Take any $h \neq 0$ with $|h|$ small enough so that $\overline{B}_R(x_0) \subset \Omega_{|h|}$ (i.e. $|h| <$ dist $(\overline{B}_R(x_0), \partial\Omega)$), and choose $\zeta \in C_c^\infty(\Omega_{|h|})$. Then $\Delta_j^{(h)}\zeta \in C_c^\infty(\Omega)$ and we can replace ζ by $\Delta_j^{(-h)}\zeta$ in ($**$). After a little rearrangement using properties (a)–(d) above, we can write the result as

(1) $$\int_\Omega \textstyle\sum_{|\alpha|,|\beta|\leq m} (-1)^{|\beta|}\tilde{a}_{\alpha\beta} D^\alpha v_h D^\beta \zeta = \int_\Omega \textstyle\sum_{|\beta|\leq m} (-1)^{|\beta|} F_{\beta,h} D^\beta \zeta,$$

where $\tilde{a}_{\alpha\beta}(x) \equiv a_{\alpha\beta}(x + he_j)$ for $x \in \Omega_{|h|}$, where

$$F_{\beta,h} = \left(\Delta_j^{(h)} f_\beta - \textstyle\sum_{|\alpha|\leq m} (\Delta_j^{(h)} a_{\alpha\beta}) D^\alpha u\right),$$

and where

$$v_h = \Delta_j^{(h)} u.$$

Then by Lemma 1 (with v_h in place of u, $F_{\beta,h}$ in place of f_β, and $R - |h|$ in place of R) we get, for any $\theta \in (0,1)$ and $0 < |h| < R/2$,

$$\|v_h\|_{m, B_{\theta R}(x_0)} \leq C\left(\|v_h\|_{m-1, B_R(x_0)} + \textstyle\sum_{|\beta|\leq m} \|F_{\beta,h}\|_{0, B_R(x_0)}\right),$$

with $C = C(m, n, R, M, \mu)$. In view of (e), (f) above, we have

$$\sup_{B_{R-|h|}(x_0)} |\Delta_j^{(h)} a_{\alpha\beta}| \leq M, \quad \|\Delta_j^{(h)} f_\beta\|_{0, B_{R-|h|}(x_0)} \leq \|f_\beta\|_{1, B_R(x_0)},$$

and $\|v_h\|_{m-1, B_{R-|h|}(x_0)} \leq \|u\|_{m, B_R(x_0)}$, so after an application of Cauchy's inequality we then conclude that

(2) $$\|v_h\|_{m, B_{\theta(R-|h|)}(x_0)} \leq C\left(\|u\|_{m, B_R(x_0)} + \textstyle\sum_{|\beta|\leq m} \|f_\beta\|_{1, B_R(x_0)}\right),$$

with constant C independent of h. Thus $\|D^\gamma v_h\|_{0, B_{\theta(R-|h|)}(x_0)}$, $|\gamma| \leq m$, and hence v_h is weakly convergent in $H^m(B_{\theta R}(x_0))$, as $h \to 0$, to some v, and (by Lemma 7 of Lecture 5) $v = D_j u$. At the same time $\Delta_j^{(h)} a_{\alpha\beta}$, $\Delta_j^{(h)} f_\beta$ converge weakly to $D_j a_{\alpha\beta}$ and $D_j f_\beta$ respectively in $L^2(B_{\theta R}(x_0))$. Thus $u \in$

$H^{m+1}_{\rm loc}(B_R(x_0))$ and we can pass to the limit in (1) to deduce that $v = D_j u$ satisfies

$$(3)\qquad \int_{B_R(x_0)} \sum_{|\alpha|,|\beta|\le m}(-1)^{|\beta|}a_{\alpha\beta}D^\alpha v D^\beta\zeta = \int_{B_R(x_0)}\sum_{|\beta|\le m}(-1)^{|\beta|}F_\beta D^\beta\zeta,\ \zeta\in C^\infty_c(B_R(x_0))$$

where

$$F_\beta = D_j f_\beta - \sum_{|\alpha|\le m}(D_j a_{\alpha\beta})D^\alpha u,$$

and also that (2) holds with $h=0$ on the right and with $v = D_j u$ in place of v_h on the left, thus giving

$$\|u\|_{m+1,B_{\theta R}(x_0)} \le C(\|u\|_{m,B_R(x_0)} + \sum_{|\beta|\le m}\|f_\beta\|_{1,B_R(x_0)})$$

for each $\theta\in(0,1)$ with $C = C(m,n,M,\mu,\theta,R)$.

We can of course repeat the above difference quotient procedure (starting with the equation (3) on $B_{\theta R}(x_0)$ for $v = D_j u$ instead of the original equation for u on B_R), until, after k steps, keeping in mind that $\theta\in(0,1)$ is arbitrary, we obtain $u\in H^{m+k}_{\rm loc}(B_R(x_0))$, and

$$\|u\|_{m+k,B_{\theta R}(x_0)} \le C(\|u\|_{m,B_R(x_0)} + \sum_{|\beta|\le m}\|f_\beta\|_{k,B_R(x_0)}),$$

where C is a constant depending only on n,m,θ,k,R,M and μ. After replacing $B_R(x_0)$ by $B_{\theta R}(x_0)$ and applying Lemma 1, we then have

$$\|u\|_{m+k,B_{\theta^2 R}(x_0)} \le C(\|u\|_{m-1,B_R(x_0)} + \sum_{|\beta|\le m}\|f_\beta\|_{k,B_R(x_0)}),$$

which of course implies, in view of the arbitrariness of θ, that

$$\|u\|_{m+k,B_{\theta R}(x_0)} \le C(\|u\|_{m-1,B_R(x_0)} + \sum_{|\beta|\le m}\|f_\beta\|_{k,B_R(x_0)}),$$

with a new constant C still depending only on n,m,θ,k,R,M and μ, so Theorem 1 is proved.

As promised in the remark following Theorem 1, we can actually replace the quantity $\|u\|_{m-1,B_R(x_0)}$ on the right by the L^2 norm $\|u\|_{0,B_R(x_0)}$. To see this we first observe that the proof of Theorem 1 applies equally well with any ball $B_\rho(y)\subset B_R(x_0)$ in place of $B_R(x_0)$, giving the inequality

$$\|u\|_{m+k,B_{\theta\rho}(y)} \le C(\|u\|_{m-1,B_\rho(y)} + \sum_{|\beta|\le m}\|f_\beta\|_{k,B_\rho(y)}).$$

Using the interpolation inequality (Lemma 6 of Lecture 5), we then have

(‡)
$$\begin{aligned}\|u\|_{m+k,B_{\theta\rho}(y)} &\le \varepsilon\|u\|_{m,B_\rho(y)} + C(\|u\|_{0,B_\rho(y)} + \sum_{|\beta|\le m}\|f_\beta\|_{k,B_\rho(y)})\\ &\le \varepsilon\|u\|_{m+k,B_\rho(y)} + C(\|u\|_{0,B_R(x_0)} + \sum_{|\beta|\le m}\|f_\beta\|_{k,B_R(x_0)}).\end{aligned}$$

Now a simple exercise in scaling (see problem 6.7 below), shows that this is true with $C = \rho^{-k-m}C_0$, where $C_0 = C_0(R,m,k,n,\theta,M,\mu)$. To complete the proof we are going to apply the following simple abstract lemma (which we shall use with the choices $S(A) = \|u\|_{m+k,A}$, $\mu = 1$, $\ell = m+k$).

Lemma 2. *Let S be a real-valued monotone sub-additive function on the class of all convex subsets of $B_R(x_0)$ (i.e. $S(A)\le\sum_{j=1}^N S(A_j)$ whenever $A, A_1,\ldots,A_N$ are convex sets with $A\subset\cup_{j=1}^N A_j\subset B_R(x_0)$), and suppose that $\theta\in(0,1)$, $\mu\in(0,1]$, $\gamma\ge 0$, and $\ell\ge 0$ are given constants. There is $\varepsilon_0 = \varepsilon_0(\ell,\theta,n) > 0$ such that if*

$$\rho^\ell S(B_{\theta\rho}(y)) \le \varepsilon_0\rho^\ell S(B_\rho(y)) + \gamma$$

whenever $B_\rho(y)\subset B_R(x_0)$ and $\rho\le\mu R$, then

$$R^\ell S(B_{\theta R}(x_0)) \le C\gamma,$$

where $C = C(\theta,\mu,\ell,n)$.

Note: Notice that the choice $S(A) = \|u\|_{m+k,A}$ does satisfy the stated conditions for S, and by virtue on ‡ the hypotheses are then all satisfied with $\mu = 1$ and $\ell = m+k$, so (since we can take any $\varepsilon > 0$ in ‡) the proof of Theorem 1 will thus be complete as soon as we establish Lemma 2.

Proof of Lemma 2: Let

$$Q = \sup_{B_\rho(y)\subset B_R(x_0),\,\rho\le\mu R}\rho^\ell S(B_{\theta\rho}(y)).$$

Then in view of the given inequality we have

$$(*)\qquad (\theta\rho)^\ell S(B_{\theta^2\rho}(y)) \le \varepsilon_0 Q + \gamma$$

for each ball $B_\rho(y)\subset B_R(x_0)$ with $\rho\le\mu R$. Take any ball $B_\rho(y)\subset B_R(x_0)$ with $\rho\le\mu R$. Then we can select balls $\{B_{(1-\theta)\rho}(y_j)\}_{j=1,\ldots,N}$ with centers $y_j\in B_{\theta\rho}(y)$ such that $B_{\theta\rho}(y)\subset\cup_{j=1}^N B_{\theta^2(1-\theta)\rho}(y_j)$ and with $N\le C$, where C is a constant depending only on θ, n. Since each $B_{(1-\theta)\rho}(y_j)\subset B_R(x_0)$, we can then apply $(*)$ with $B_{(1-\theta)\rho}(y_j)$ in place of $B_\rho(y)$ and use the given subadditivity of S; this gives

$$\rho^\ell S(B_{\theta\rho}(y)) \le C(\varepsilon_0 Q + \gamma),\quad C = C(n,\ell,\theta).$$

In view of the arbitrariness of $B_\rho(y)$, this gives

$$Q \le C\varepsilon_0 Q + C\gamma, \quad C = C(n, \ell, \theta),$$

and hence if $\varepsilon_0 \le \frac{1}{2}C^{-1}$, we get

$$Q \le 2C\gamma,$$

so that

$$\rho^\ell S(B_{\theta\rho}(y)) \le 2C\gamma$$

for every ball $B_\rho(y) \subset B_R(x_0)$ with $\rho \le \mu R$. Since we can cover $B_{\theta R}(x_0)$ by at most $C = C(\theta, \mu, n)$ balls $B_{\theta\mu R}(y_j)$ with $B_{\mu R}(y_j) \subset B_R(x_0)$, we can again use the given subadditivity of S to conclude the stated inequality.

LECTURE 6 PROBLEMS

6.1 Let $F = F(x, \{p_\alpha\}_{|\alpha| \le m})$ be a smooth function of the variables $x \in \overline{\Omega}$, and $p_\alpha \in \mathbb{R}$, $|\alpha| \le m$. Find the Euler-Lagrange operator $\mathcal{M}$ of the functional $\mathcal{F}(u) = \int_\Omega F(x, \{D^\alpha u\})\, dx$, and show that its linearization at a smooth function u_0 is the operator $D^\beta(a_{\alpha\beta} D^\alpha u)$, where

$$a_{\alpha\beta}(x) = F_{p_\alpha p_\beta}(x, \{D^\delta u_0(x_0)\}_{|\delta| \le m}).$$

(The subscripts p_α, p_β here denote partial derivatives; also, recall that the linearization of $\mathcal{M}(u)$ at u_0 is given by $\mathcal{L}(v) = \frac{d}{ds}\mathcal{M}(u_0 + sv)|_{s=0}$.)

6.2 Suppose $u \in L^2_{\text{loc}}(\Omega)$ is a solution of the equation

$$\Delta u = \textstyle\sum_{j=1}^n b_j D_j u + cu + f$$

in the appropriate weak sense (which you should formulate), where $b_j, c, f \in C^\infty(\Omega)$. Prove that $u \in C^\infty(\Omega)$. (Notice that this would be included in Corollary 1 above if we already knew $u \in H^1_{\text{loc}}(\Omega)$.)

Hint: Show u_σ satisfies an equation of the form $\Delta u_\sigma = \sum D_j(b_j u)_\sigma + (\tilde{c}u)_\sigma + f_\sigma$, and then use the estimates of Theorem 1 and the Sobolev embedding theorem.

6.3 Note that in case $m = 1$ in the above, it may be convenient to use standard index notation: for example $\sum_{|\alpha|,|\beta|=1} D^\beta(a_{\alpha\beta} D^\alpha u)$ would be better written $\sum_{i,j=1}^n D_i(a_{ij} D_j u)$ etc. Write out the proof of Theorems 1,2 in such notation, paying particular attention to the ellipticity condition in this case.

6.4 The usual convention is that we write

$$\textstyle\sum_{|\alpha|,|\beta| \le m} D^\beta(a_{\alpha\beta} D^\alpha) = \sum_{|\alpha|,|\beta| \le m} D^\beta(b_{\alpha\beta} D^\alpha)$$

on Ω, provided $a_{\alpha\beta}, b_{\alpha\beta} \in L^1_{\text{loc}}(\Omega)$ and

$$\int_\Omega \textstyle\sum_{|\alpha|,|\beta| \le m} (-1)^{|\beta|} a_{\alpha\beta} D^\alpha\varphi D^\beta\psi = \int_\Omega \sum_{|\alpha|,|\beta| \le m} (-1)^{|\beta|} b_{\alpha\beta} D^\alpha\varphi D^\beta\psi$$

for each $\varphi, \psi \in C^\infty_c(\Omega)$.

(i) Discuss the fact that in case $m = 1$ (when we write a_{ij} rather than $a_{\alpha\beta}$ if $\alpha = e_i, \beta = e_j$), there can be many choices of $\{a_{ij}\}_{i,j=1,\dots,n}$ such that $\Delta = \sum_{i,j=1}^n D_j(a_{ij} D_i)$. (One choice is of course $a_{ij} = \delta_{ij}$.) Discuss the ellipticity condition (E) in these various cases.

(ii) Same question with $m = 2$ and with Δ^2 in place of Δ.

6.5 Suppose that in place of condition (E) we have the condition that there exists constants c_1, c_2 such that

$$\text{(C)} \qquad \int_\Omega \sum_{|\alpha|=m} |D^\alpha \varphi|^2 \le c_1 \int_\Omega \sum_{|\alpha|,|\beta|=m} a_{\alpha\beta} D^\alpha \varphi D^\beta \varphi + c_2 \|\varphi\|_{0,\Omega}^2$$

for each $\varphi \in C_c^\infty(\Omega)$. Show that the proofs of Lemma 1 and Theorem 1 can be modified in such a way that it suffices to know (C) in place of (E).

6.6 (i) Prove that condition (C) in the previous problem is equivalent to condition (E) in case $m = 1$.

Hint: Take $\cos(\xi \cdot x)\varphi$, $\sin(\xi \cdot x)\varphi$ in place of φ in (C), where $\xi \in \mathbb{R}^n$ is arbitrary.

(ii) In general, prove that condition (C) implies

$$\sum_{|\alpha|,|\beta|=m} a_{\alpha\beta} \xi^\alpha \xi^\beta \ge \mu |\xi|^{2m} \ \forall \xi = (\xi^1, \dots, \xi^n) \in \mathbb{R}^n$$

for a fixed constant $\mu > 0$.

6.7 (An exercise on scaling.) (i) In the above lecture if we apply Lemma 1 to an arbitrary ball $B_\rho(y) \subset B_R(x_0)$, then we get

$$\|u\|_{m,B_{\theta\rho}(y)} \le \varepsilon \|u\|_{m,B_\rho(y)} + C(\|u\|_{0,B_\rho(y)} + \sum_{|\beta|\le m} \|f_\beta\|_{0,B_\rho}(y)),$$

where $C = C(\varepsilon, m, n, \theta, \rho, M, \mu)$. Prove that $C \le C_0 \rho^{-m}$, where C_0 does not depend on ρ; i.e. $C_0 = C_0(\varepsilon, m, n, R, \theta, M, \mu)$.

Hint: Let $\xi = \rho^{-1}(x-y)$, $\tilde{u}(\xi) = u(x)$, and notice that $D_\xi^\alpha \tilde{u}(\xi) = \rho^{|\alpha|} D_x^\alpha u(x)$. Show that $\tilde{u}$ satisfies an equation of the same form as the equation for u with the ball $B_1(0)$ in place of $B_\rho(y)$ and with coefficients satisfying (E) and also (B) but with $C(R)M$ in place of M.

(ii) Similarly show that $C \le C_0 \rho^{-m-k}$, where $C_0 = C_0(\varepsilon, m, n, R, k, \theta, M, \mu)$, in the inequality ‡ of the above lecture.

6.8 (Another exercise on scaling.)

If we replace (B_k) by the "scale-invariant" version

$$R^{|\gamma|+2m-|\alpha|-|\beta|} \sup_{B_R(x_0)} |D^\gamma a_{\alpha\beta}| \le M, \quad |\gamma| \le k,$$

prove that the result of Theorem 1 can be written in the scale-invariant form

$$\sum_{|\alpha|\le m+k} R^{|\alpha|} \|D^\alpha u\|_{0,B_{\theta R}(x_0)} \le C\Big(\sum_{|\alpha|\le m-1} R^{|\alpha|} \|D^\alpha u\|_{0,B_R(x_0)} + \sum_{|\beta|\le m, |\gamma|\le k} R^{2m-|\beta|+|\gamma|} \|D^\gamma f_\beta\|_{0,B_R(x_0)}\Big),$$

where now $C = C(m, n, k, \theta, M, \mu)$; i.e., C is independent of R.

Hint: Apply the result of Theorem 1 on the unit ball $B_1(0)$ to the function $\tilde{u}(x)$, $\tilde{u}(x) = u(x_0 + Rx)$, $x \in B_1(0)$; i.e., $u(x) = \tilde{u}(R^{-1}(x - x_0))$, $x \in B_R(x_0)$.

Lecture 7

Solvability of the Dirichlet Problem in Sobolev Space, Coercivity and Gårding's Inequality

We begin by describing the Dirichlet problem and its weak formulation, in the context of the operators considered in Lecture 6, i.e. the operators

$$Lu = (-1)^m \textstyle\sum_{|\alpha|,|\beta|\le m} D^\beta \big(a_{\alpha\beta}(x) D^\alpha u\big),$$

where $m \ge 1$ and where the factor $(-1)^m$ in front is purely for later notational convenience, and where we continue to assume the ellipticity and boundedness conditions

$$\begin{array}{ll} \text{(E)} & \sum_{|\alpha|,|\beta|=m} a_{\alpha\beta}\lambda^\alpha\lambda^\beta \ge \mu \sum_{|\alpha|=m}(\lambda^\alpha)^2 \quad \forall \text{ set of real constants } \{\lambda^\alpha\}_{|\alpha|=m} \\ \text{(B)} & \qquad\qquad |a_{\alpha\beta}| \le M \quad |\alpha|, |\beta| \le m, \end{array}$$

where $\mu,\ M > 0$ are given fixed constants. (In case $m \ge 2$, we consider below a relaxation of the ellipticity condition (E).)

As usual let Ω be a bounded domain in $\mathbb{R}^n$. The classical Dirichlet problem for the operators of the above type is the problem of determining u on Ω such that

$$\text{(D)}\qquad \begin{cases} Lu = f \text{ on } \Omega \\ D^\alpha u = 0 \text{ on } \partial\Omega, \quad 0 \le |\alpha| \le m-1, \end{cases}$$

where $f = \sum_{|\beta|\le m} D^\beta f_\beta$ with $\{f_\beta\}_{|\beta|\le m}$ given. Notice that if $u \in C^{m-1}(\overline{\Omega})$, and if the domain has C^1 boundary with inward pointing unit normal η, then the Dirichlet boundary conditions $D^\alpha u = 0$, $|\alpha| \le m-1$ are equivalent to the m conditions $D^j_\eta u = 0$, $j = 0, \ldots, m-1$, where $D^j_\eta u = \frac{d^j}{ds^j} u(x + s\eta(x))\,|_{s=0}$, $x \in \partial\Omega$, as one easily checks by induction on $m \ge 1$.

Actually, the above description of the Dirichlet problem is not complete, because we haven't specified the class to which the solution is required to belong, nor have we said anything about assumptions on the function f_β appearing on the right. We ultimately (at least often) are interested in classical solutions; i.e. smooth functions u which satisfy the equation and the boundary conditions in the classical sense, and of course we generally would only expect such classical solutions to exist if the given coefficients $a_{\alpha\beta}$ and function f are sufficiently smooth, not to mention smoothness assumptions which might be appropriate for the boundary $\partial\Omega$. We shall discuss such assumptions in due course, but for the moment, as a first step, we are interested in formulating the Dirichlet problem weakly in Sobolev space, and in developing a general existence theory in that setting.

The weak formulation is actually very simple: we say that u is a weak solution of the Dirichlet problem (D) provided

$$\text{(P)}\qquad u \in H^m_0(\Omega), \text{ and } \int_\Omega (-1)^{m+|\beta|} a_{\alpha\beta} D^\alpha u D^\beta \varphi = \int_\Omega (-1)^{|\beta|} f_\beta D^\beta \varphi$$

for all $\varphi \in H^m_0(\Omega)$. Notice that this simply requires that u be an $H^m_{\mathrm{loc}}(\Omega)$ weak solution of the equation in the sense of the previous lecture, together with the additional condition that $u \in H^m_0(\Omega)$; this latter condition is just a weak formulation of the classical Dirichlet boundary conditions $D^\alpha u = 0$ on $\partial\Omega$ for $|\alpha| \le m-1$ in (D) above. Indeed, by virtue of Exercise 5.7, if $u \in C^{m-1}(\overline{\Omega}) \cap H^m(\Omega)$, then (P) is equivalent to finding u which satisfies the equation $Lu = \sum_{|\beta|\le m} D^\beta f_\beta$ (in the weak Sobolev space sense of Lecture 6) with $D^\alpha u = 0$ on $\partial\Omega$, $|\alpha| \le m-1$ in the classical sense. Indeed we shall show that the regularity theory in Corollary 1 of Lecture 6 implies that any solution of the above weak formulation (P) of (D) automatically provides a solution of the classical problem $*$ if $a_{\alpha\beta},\ f_\beta \in C^\infty(\overline{\Omega})$.

Actually, as we shall see in a moment, it is better to consider the following more general problem in place of (D), in which $\lambda \in \mathbb{R}$ is a free parameter:

$$(\text{D}_\lambda)\qquad \begin{cases} Lu = \lambda u + f \text{ on } \Omega \\ D^\alpha u = 0 \text{ on } \partial\Omega, \quad 0 \le |\alpha| \le m-1. \end{cases}$$

The weak formulation of this modified problem is

$$(\text{P}_\lambda)\qquad \text{Find } u \in H^m_0(\Omega) \text{ such that } A_\lambda(u,\zeta) = F(\zeta), \quad \forall\, \zeta \in H^m_0(\Omega)$$

where $A_\lambda(u,\zeta) = A_0(u,\zeta) - \lambda\langle u,\zeta\rangle_0$, with

$$A_0(u,\zeta) = \int_\Omega \big(\textstyle\sum_{|\alpha|,|\beta|\le m} (-1)^{m+|\beta|} a_{\alpha\beta} D^\alpha u D^\beta \zeta\big)$$
$$F(\zeta) = \int_\Omega \textstyle\sum_{|\beta|\le m} (-1)^{|\beta|} f_\beta D^\beta \zeta.$$

To begin our discussion of solvability of (P_λ) we first need a simple result valid in abstract Hilbert space:

Lemma. (Lax-Milgram Lemma.) *Let H be any real Hilbert space, and let B be a bounded strictly coercive bilinear form; that is*

(i) $\qquad B : H \times H \to \mathbb{R},\ \ B(u,v)$ is linear in u, v

(ii) $\qquad |B(u,v)| \le \Lambda \|u\|_H \|v\|_H$

(iii) $\qquad B(u,u) \ge \lambda \|u\|^2_H$

where Λ, λ are fixed positive constants.

Then there exists an isomorphism T of H onto H with

$$\lambda\|u\| \le \|Tu\| \le \Lambda\|u\|, \quad B(u,v) \equiv \langle Tu, v\rangle_H \quad \forall u, v \in H.$$

Proof: The proof is based on the Riesz representation theorem in Hilbert space as follows. For fixed $u \in H$, $B(u,v)$ is a bounded linear functional on H, hence by the Riesz representation theorem $\exists z \in H$ such that $B(u,v) \equiv \langle z, v\rangle_H\ \forall v \in H$. Such z is trivially unique for a given u, so we denote it by Tu, where $T : H \to H$. So defined, T is obviously linear because $B(u,v)$ is linear in u. By (iii) and the Cauchy-Schwarz inequality we have

$$\lambda\|u\|^2 \le \langle T(u), u\rangle_H \le \|Tu\|\,\|u\|,$$

so that in particular $\lambda\|u\| \le \|Tu\|$, which evidently implies that T has <u>closed range</u>. By (ii) with $v = T(u)$ we also get $\|T(u)\| \le \lambda\|u\|$, so that T is a bounded operator with closed range. Then if the range of T is not all of H we could choose $v \neq 0$ orthogonal to the range of T, so $B(v,v) = \langle T(v), v\rangle_H = 0$, contradicting (iii).

Definition: for any given $\lambda \in \mathbb{R}$, we say the problem (P_λ) (or the original problem (D)=(P_0)) is coercive if there are constants $\mu > 0$ and $\gamma \in \mathbb{R}$ such that

$$\text{(C)} \qquad A_\lambda(u,u) \geq \mu \|u\|^2_{m,\Omega} - \gamma \|u\|^2_0 \;\; \forall\, u \in H^m_0(\Omega).$$

The problem (P_λ) is said to be strictly coercive (for a given λ) if this holds with $\gamma = 0$. Notice that of course if the problem is strictly coercive for a given λ, say λ_0, then it is also strictly coercive for any $\lambda \leq \lambda_0$.

Remarks: (1) Notice that (P_λ) is coercive for any λ and strictly coercive for $\lambda < -C$, $C = C(n,m,M,\mu)$, provided (E), (B) above hold. This is proved as follows: If $u \in H^m_0(\Omega)$ then by (E) we have

$$\mu \int_\Omega \textstyle\sum_{|\alpha|=m} (D^\alpha u)^2 \leq \int_\Omega \sum_{|\alpha|=|\beta|=m} a_{\alpha\beta} D^\alpha u D^\beta u =$$
$$A_0(u,u) - \int_\Omega \textstyle\sum_{|\alpha|,|\beta|\leq m, |\alpha|+|\beta|<2m} a_{\alpha\beta} D^\alpha u D^\beta u,$$

and hence (since $\|u\|^2_{m,\Omega} = \int_\Omega \sum_{|\alpha|=m} (D^\alpha u)^2 + \|u\|^2_{m-1,\Omega}$)

$$\|u\|^2_{m,\Omega} \leq \mu^{-1} A_0(u,u) - \mu^{-1} \int_\Omega \textstyle\sum_{|\alpha|,|\beta|\leq m, |\alpha|+|\beta|<2m} a_{\alpha\beta} D^\alpha u D^\beta u + \|u\|^2_{m-1,\Omega},$$

and hence

$$\|u\|^2_{m,\Omega} \leq \mu^{-1} A_0(u,u) + C \int_\Omega \textstyle\sum_{|\alpha|\leq m, |\beta|\leq m-1} |D^\alpha u||D^\beta u| + \|u\|^2_{m-1,\Omega}$$

with $C = C(m,n,M,\mu)$, which by the Cauchy-Schwarz inequality $|ab| \leq \varepsilon a^2 + \varepsilon^{-1} b^2$ and the interpolation inequality

$$\|u\|^2_{m-1,\Omega} \leq \varepsilon \|u\|^2_{m,\Omega} + C\varepsilon^{1-m} \|u\|^2_{0,\Omega}$$

(see Lemma 6 of Lecture 5) gives

$$\|u\|^2_{m,\Omega} \leq \mu^{-1} A_0(u,u) + 2\varepsilon \|u\|^2_{m,\Omega} + C\|u\|^2_{0,\Omega},$$

with $C = C(n,m,M,\mu,\varepsilon)$. Taking $\varepsilon = 1/4$ then gives

$$A_{\lambda_0}(u,u) \geq \frac{1}{2}\mu \|u\|^2_{m,\Omega} - (C + \lambda_0)\|u\|^2_{0,\Omega}$$

with $C = C(n,m,M,\mu)$. Thus $A_{\lambda_0}(u,u) \geq \frac{1}{2}\mu\|u\|^2_{m,\Omega}$ provided $\lambda_0 \leq -C$, and indeed we have strict coercivity of $A_{\lambda_0}(u,u)$ provided $\lambda_0 \leq -C$ for suitable constant $C = C(n,m,M,\mu)$.

(2) If the $a_{\alpha\beta}$ corresponding to $|\alpha| = |\beta| = m$ are at least continuous on $\overline{\Omega}$ then we shall describe below a weaker condition (E)′ which (together with (B)) guarantees coercivity. See the discussion of Gårding's inequality below.

We are now ready to begin our discussion of the existence of weak solutions of the problem (P_λ); we assume that the problem (P_{λ_0}) is strictly coercive for some $\lambda_0 \in \mathbb{R}$. (Recall that this is always guaranteed by (E), (B) by Remark (1) above.)

We look first at the case when $f_\beta \in H^{|\beta|}(\Omega)$ for each $|\beta| \leq m$ (the general case is discussed in Exercises 7.1, 7.2 below), so that (P_λ) can be written

$$A_\lambda(u,\zeta) = \langle f, \zeta\rangle_0, \quad f = \textstyle\sum_{|\beta|\leq m} D^\beta f_\beta, \;\; \zeta \in \mathcal{H},$$

where, here and subsequently, we write $\mathcal{H} = H^m_0(\Omega)$ and $\langle\,\rangle_0$ denotes the inner product on $L^2(\Omega)$. Notice that of course that the above identity is the same as

$$A_0(u,\zeta) - \lambda\langle u,\zeta\rangle_0 = \langle f,\zeta\rangle_0, \quad \zeta \in \mathcal{H}.$$

The strict coercivity of A_{λ_0} tells us that $\exists\, \beta > 0$ such that

$$A_{\lambda_0}(u,u) \geq \beta \|u\|^2_{m,\Omega}, \quad u \in \mathcal{H},$$

while the boundedness condition (B) and the Schwarz inequality in $\mathcal{H}$ tells us that

$$|A_{\lambda_0}(u,\zeta)| \leq C\|u\|_{m,\Omega}\|\zeta\|_{m,\Omega}, \quad u,\zeta \in \mathcal{H},$$

and hence we can apply the Lax-Milgram lemma with $B = A_{\lambda_0}$. We in fact claim:

Lemma. *There is a solution operator $S : L^2(\Omega) \to \mathcal{H}$ for problem (P_{λ_0}) in the sense that there is a bounded linear injective operator $S : L^2(\Omega) \to \mathcal{H}$ with*

$$A_{\lambda_0}(S(f),\zeta) \equiv \langle f,\zeta\rangle_0, \quad \zeta \in \mathcal{H},\; f \in L^2(\Omega).$$

Here λ_0 is any real constant such that A_{λ_0} is strictly coercive as in the above discussion.

Proof: By the Lax-Milgram lemma $\exists$ an isomorphism T of $\mathcal{H}$ with

$$A_{\lambda_0}(u,\zeta) \equiv \langle Tu,\zeta\rangle_{m,\Omega}, \quad u,\zeta \in \mathcal{H}.$$

Next notice that $|\langle f,\zeta\rangle_0| \leq \|f\|_0\|\zeta\|_0 \leq \|f\|_0\|\zeta\|_m$, so $\psi \mapsto \langle f, zeta\rangle_0$ is a continuous linear functional on $\mathcal{H}$. Thus by the Riesz representation theorem there is a $w \in \mathcal{H}$ with $\langle w,\zeta\rangle_{H^m(\Omega)} \equiv \langle f,\zeta\rangle_0\; \forall \zeta \in \mathcal{H}$.

Clearly such w is unique for given f, so write $w = S_0(f)$, where $S_0 : L^2(\Omega) \to \mathcal{H}$. S_0 is clearly bounded linear, and injective (the latter since $\mathcal{H}$ is dense in $L^2(\Omega)$). Using the definition of T, S_0 we see that if $S = T^{-1} \circ S_0$, then $A_{\lambda_0}(S(f), \zeta) \equiv \langle f, \zeta\rangle_0$ as required.

Thus we have our solution operator for (P_{λ_0}) (with λ_0 chosen to ensure strict coercivity of A_{λ_0}). Similarly we have a solution operator $\widetilde{S}$ for the adjoint problem, corresponding to the adjoint operator

$$L'u \equiv \textstyle\sum_{|\alpha|,|\beta|\le m}(-1)^{m+|\alpha|+|\beta|}D^\alpha(a_{\alpha\beta}D^\beta u)$$

with corresponding bilinear form $\widetilde{A}_{\lambda_0}(u, v) = A_{\lambda_0}(v, u)$, so that

$$A_{\lambda_0}(\zeta, \widetilde{S}(f)) = \langle f, \zeta\rangle_0 = A_{\lambda_0}(S(f), \zeta)\ \forall \zeta \in \mathcal{H},$$

so (replacing f by g and ζ by $S(f)$ in the first identity and ζ by $\widetilde{S}(g)$ in the second) we have

$$A_{\lambda_0}(S(f), \widetilde{S}(g)) = \langle f, \widetilde{S}(g)\rangle_0 = \langle S(f), g\rangle_0, \quad f, g \in L^2(\Omega),$$

Thus, with ι the inclusion map $\mathcal{H} \subset L^2(\Omega)$, we have $\iota \circ \widetilde{S} : L^2(\Omega) \to L^2(\Omega)$ is the adjoint $(\iota \circ S)'$ of $\iota \circ S : L^2(\Omega) \to L^2(\Omega)$.

Since the inclusion map $\iota : \mathcal{H} \to L^2(\Omega)$ is compact by Rellich's theorem, we see that $\iota \circ S : L^2(\Omega) \to L^2(\Omega)$ and $\iota \circ \widetilde{S} = (\iota \circ S)' : L^2(\Omega) \to L^2(\Omega)$ are also compact. Thus we can apply the Fredholm alternative for Hilbert space to give the following important facts, in which I is the identity map on $L^2(\Omega)$:

$\exists$ a discrete set $\Lambda \subset (\lambda_0, \infty)$ such that

(i) For $\lambda \notin \Lambda$, $I - (\lambda - \lambda_0)\iota \circ S$, $I - (\lambda - \lambda_0)(\iota \circ S)'$ are isomorphisms of $L^2(\Omega)$ onto itself;

(ii) For $\lambda \in \Lambda$: $\ker(I - \lambda\iota \circ S)$, $\ker(I - (\lambda - \lambda_0)(\iota \circ S)')$ are finite dimensional subspaces of $\mathcal{H}$ of the same (positive) dimension, and

$$\begin{cases} \text{range}\,(I - (\lambda - \lambda_0)(\iota \circ S)) = (\ker(I - (\lambda - \lambda_0)(\iota \circ S)'))^\perp \\ \text{range}\,(I - (\lambda - \lambda_0)(\iota \circ S)') = (\ker(I - (\lambda - \lambda_0)(\iota \circ S)))^\perp, \end{cases}$$

where $(\,)^\perp$ means orthogonal complement in $L^2(\Omega)$.

We now want to check what the above Fredholm facts tell us about weak solutions of (P_λ). Let us consider the following problems:

($\mathrm{P}_\lambda(f)$) The problem of finding $u \in \mathcal{H}$, with $A_\lambda(u, \zeta) \equiv \langle f, \zeta\rangle_0, \quad \zeta \in \mathcal{H}$.

($\mathrm{Q}_\lambda(f)$) The problem of finding $u \in \mathcal{H}$, with $A_\lambda(\zeta, u) \equiv \langle f, \zeta\rangle_0, \quad \zeta \in \mathcal{H}$.

Recall also (by definition of the solution operator S) that we have

$$A_{\lambda_0}(S(f), \zeta) \equiv \langle f, \zeta\rangle_0, \quad f \in L^2(\Omega),\ \zeta \in \mathcal{H},$$

and

$$A_\lambda(u, \zeta) = A_{\lambda_0}(u, \zeta) - (\lambda - \lambda_0)\langle u, \zeta\rangle_0$$

and hence

(∗)

$$\begin{aligned} A_\lambda(u, \zeta) \equiv \langle f, \zeta\rangle_0,\ \forall \zeta \in \mathcal{H} &\iff A_{\lambda_0}(u, \zeta) \equiv \langle(\lambda - \lambda_0)u + f, \zeta\rangle_0,\ \forall \zeta \in \mathcal{H} \\ &\iff u = S((\lambda - \lambda_0)u + f) \\ &\iff u = S(w) \text{ with } w \in L^2(\Omega) \text{ and} \\ &\qquad (I - (\lambda - \lambda_0)\iota \circ S)w = f, \end{aligned}$$

Similarly, by the same reasoning applied to $\widetilde{A}_\lambda(u, \zeta) = A_\lambda(\zeta, u)$, we have

$$\begin{aligned} A_\lambda(\zeta, u) \equiv \langle f, \zeta\rangle_0, \quad \zeta \in L^2(\Omega) &\iff u = (\iota \circ S)'(w) \text{ and} \\ &\qquad (I - (\lambda - \lambda_0)(\iota \circ S)')w = f. \end{aligned}$$

In light of these equivalences, the general Fredholm facts above enable us to conclude the following about weak solutions (i.e. solutions $u \in \mathcal{H}$ of ($\mathrm{P}_\lambda(f)$), ($\mathrm{Q}_\lambda(f)$)):

Theorem 1. *There is a discrete set $\Lambda \subset (\lambda_0, \infty)$ (where λ_0 is any real number such that A_{λ_0} is strictly coercive) so that the following alternatives hold:*

(i) *If $\lambda \notin \Lambda$, then for each $f \in L^2(\Omega)$ both the problems $\mathrm{P}_\lambda(f)$, $\mathrm{Q}_\lambda(f)$ have a unique solution $u \in \mathcal{H}$, and*

(ii) *If $\lambda \in \Lambda$, then* (a) $\mathrm{P}_\lambda(0)$ *(resp. $\mathrm{Q}_\lambda(0)$) has a set of weak solutions forming a finite dimensional subspace N_λ of $\mathcal{H}$ (resp. M_λ of $\mathcal{H}$), and* $\dim M_\lambda = \dim N_\lambda \ge 1$, *and* (b) *For arbitrary $f \in L^2(\Omega)$ the problem $\mathrm{P}_\lambda(f)$ (resp. $\mathrm{Q}_\lambda(f)$) has a weak solution if and only if $f \in (M_\lambda)^\perp$ (resp. $f \in (N_\lambda)^\perp$), where the orthogonal complements are taken in $L^2(\Omega)$.*

Remark: Notice that in checking (ii)(a) above we used the fact that, for $\lambda \ne \lambda_0$, $w \in \ker(I - (\lambda - \lambda_0)\iota \circ S) \iff w = (\lambda - \lambda_0)S(w) \iff A_\lambda(u, \zeta) = 0\ \forall \zeta \in \mathcal{H}$, with $u = S(w) = (\lambda - \lambda_0)^{-1}w$, and similarly $w \in \ker(I - (\lambda - \lambda_0)(\iota \circ S)') \iff A_\lambda(\zeta, u) = 0\ \forall \zeta \in \mathcal{H}$, with $u = (\iota \circ S)'(w) = (\lambda - \lambda_0)^{-1}w$.

We later (in Lecture 10) discuss further details in the self-adjoint case when $A_0(u, v) = A_0(v, u)$.

If we use the regularity results of the previous lecture then we obtain the following important additional information:

Corollary 1. *If $a_{\alpha\beta}$, $f \in C^\infty(\Omega)$, then the weak solutions in $P_\lambda(f)$, $Q_\lambda(f)$ are all $C^\infty(\Omega)$.*

Remarks: (1) Notice this in particular applies to solutions of the homogeneous problems $P_\lambda(0)$, $Q_\lambda(0)$ (which can be non-trivial in the case $\lambda \in \Lambda$).

(2) The above corollary says nothing about regularity, or even continuity, up to the boundary; this will be discussed in the next lecture.

We want to conclude with a discussion of Gårding's inequality, which, as we shall show, makes it possible to relax the ellipticity condition (E) in case $m \ge 2$, provided we have that the coefficients $a_{\alpha\beta}$ with $|\alpha| = |\beta| = m$ are continuous; indeed, we want to show that in this case it suffices that

$$\text{(E)}' \qquad \sum_{|\alpha|,|\beta|=m} a_{\alpha\beta}\xi^\alpha\xi^\beta \ge \mu|\xi|^{2m} \quad \forall \xi \in \mathbb{R}^n.$$

Lemma. (Gårding's inequality.) *Suppose* (B), (E)$'$ *hold. If $\mu > \varepsilon > 0$, and if there exists $\rho > 0$ such that $|a_{\alpha\beta}(x) - a_{\alpha\beta}(y)| \le n^{-m}\varepsilon$ whenever $x, y \in \Omega$ with $|x - y| \le \rho$ then*

$$\int_\Omega \sum_{|\alpha|,|\beta|=m} a_{\alpha\beta} D^\alpha\psi D^\beta\psi \ge \beta\|\psi\|^2_{m,\Omega} - C\|\psi\|^2_{0,\Omega}, \quad \psi \in H_0^m(\Omega),$$

where $\beta = \mu - \varepsilon' > 0$, $\varepsilon < \varepsilon' < \mu$ arbitrary, and where C depends only $\mu, m, n, \rho, M, \varepsilon'$, and Ω.

Remark: An important case is when the $a_{\alpha\beta}$ corresponding to $|\alpha| = |\beta| = m$ are continuous on $\overline{\Omega}$, in which case there of course exists $\rho = \rho(\varepsilon, n, m) > 0$ as in the statement for any given $\varepsilon > 0$, by virtue of the uniform continuity of $a_{\alpha\beta}$ on the compact set $\overline{\Omega}$.

Proof: We first establish the inequality assuming that $\psi \in C_c^\infty(\Omega \cap B_\rho(y))$, where $y \in \Omega$ and ρ is as in the statement of the lemma.

We have, using the properties of the Fourier transform discussed in Lecture 5, that

$$\begin{aligned}\int_{\Omega\cap B_\rho(x_0)} \sum_{|\alpha|,|\beta|=m} a^0_{\alpha\beta} D^\alpha\psi D^\beta\psi &= \sum_{|\alpha|,|\beta|=m} a^0_{\alpha\beta} (\widehat{D^\alpha\psi}, \widehat{D^\beta\psi})_{L^2} \\ &= \int_{\mathbb{R}^n} \sum_{|\alpha|,|\beta|=m} a^0_{\alpha\beta}\xi^\alpha\xi^\beta|\hat\psi|^2\,d\xi \\ &\ge \mu\int_{\mathbb{R}^n}\sum_{|\alpha|=m}(\xi^\alpha)^2|\hat\psi|^2\,d\xi \\ &\equiv \mu\int_{\Omega\cap B_\rho(x_0)}\sum_{|\alpha|=m}|D^\alpha\psi|^2.\end{aligned}$$

(Notice that we used the inequality $|\xi|^{2m} \ge \sum_{|\alpha|=m}(\xi^\alpha)^2$ here.) Since

$$\sum_{|\alpha|=|\beta|=m}|D^\alpha\psi|\cdot|D^\beta\psi| \le n^m\sum_{|\alpha|=m}|D^\alpha\psi|^2$$

by the Cauchy inequality (since there are $\le n^m$ multi-indices α with $|\alpha| = m$), we then have

$$(1) \qquad \int_{\Omega\cap B_\rho(x_0)}\sum_{|\alpha|,|\beta|=m} a_{\alpha\beta} D^\alpha\psi D^\beta\psi \ge (\mu-\varepsilon)\int_{\Omega\cap B_\rho(x_0)}\sum_{|\alpha|=m}|D^\alpha\psi|^2.$$

The rest of the proof now follows follows direct from this by making appropriate use of partition of unity and the Cauchy inequality. The details are as follows:

Let $\{\zeta_j^2\}$ be a partition of unity for $\overline{\Omega}$ subordinate to the cover given by the balls $\{B_\rho(y)\}_{y\in\overline{\Omega}}$ such that $\zeta_j \in C_c^\infty(\mathbb{R}^n)$, fix j for the moment and write $\zeta = \zeta_j$. From (1)

$$\int_\Omega \sum_{|\alpha|,|\beta|=m} a_{\alpha\beta} D^\alpha(\psi\zeta)D^\beta(\psi\zeta) \ge (\mu-\varepsilon)\int_\Omega\sum_{|\alpha|=m}|D^\alpha(\psi\zeta)|^2.$$

By virtue of the Leibniz formula and Cauchy's inequality, this gives

$$\begin{aligned}(\mu-\varepsilon')\int_\Omega \zeta^2\sum_{|\alpha|=m}|D^\alpha\psi|^2 \le{}& \int_\Omega\zeta^2\sum_{|\alpha|,|\beta|=m} a_{\alpha\beta}D^\alpha\psi D^\beta\psi \\ &+ C\int_\Omega\sum_{|\gamma|\le m-1}|D^\gamma\psi|^2\sum_{|\delta|\le m}|D^\delta\zeta|^2,\end{aligned}$$

for any $\varepsilon' > \varepsilon$, where $C = C(\varepsilon', n, m, M, \mu, \zeta)$.

Keeping in mind that $\zeta = \zeta_j$ and that $\sum_j\zeta_j^2 \equiv 1$ in Ω, we then get

$$(\mu-\varepsilon')\int_\Omega\sum_{|\alpha|=m}|D^\alpha\psi|^2 \le \int_\Omega\sum_{|\alpha|,|\beta|=m} a_{\alpha\beta}D^\alpha\psi D^\beta\psi + C\|\psi\|^2_{m-1,\Omega}.$$

The required inequality now follows by using the interpolation inequality for $H_0^m(\Omega)$ space, which (see Lecture 5) tells us that, for any $\delta > 0$, $\|u\|_{m-1,\Omega} \le \delta\|u\|_{m,\Omega} + C(m,\delta)\|u\|_{0,\Omega}$.

LECTURE 7 PROBLEMS

7.1 Discuss the existence and regularity theory for the weak problem $P_\lambda(f)$ in case f is replaced by $\sum_{|\beta|\le m}(-1)^{|\beta|}D^\beta f_\beta$, where $f_\beta \in L^2(\Omega)$; notice that then the expression $\langle f, \zeta\rangle$ of the right of $P_\lambda(f)$ must be replaced by $\int_\Omega \sum_{|\beta|\le m} f_\beta D^\beta \zeta$. Hint: Use the fact that there are $H^m(\Omega)$ functions $\{f_\beta^{(j)}\}$ converging to f_β in $L^2(\Omega)$, and begin by applying the theory developed in the above lecture with $f^{(j)} \equiv \sum_{|\beta|\le m} D^\beta f_\beta^{(j)}$ in place of f.

7.2 (i) Prove that there is always a weak solution $u \in H_0^1(\Omega)$ for the problem

$$\begin{cases} -\Delta u + qu = f \text{ on } \Omega \\ \qquad\quad u = 0 \text{ on } \partial\Omega, \end{cases}$$

given that q is a non-negative bounded function on Ω and $f \in L^2(\Omega)$.

Hint: You have to show that $0 \notin \lambda$.

(ii) Extend to the case when f on the right is replaced by $f + \sum_{j=1}^n D_j f_j$.

(iii) Discuss regularity properties of the solution under suitable conditions on $q,\ f,\ f_j$.

7.3 If (P_λ) is as above, and (P_λ) is strictly coercive, prove that the solution $u \in H_0^m(\Omega)$ of (P_λ) is unique.

7.4 Discuss the problem

$$\begin{cases} \Delta^2 u = \Delta^2 f + \Delta g \text{ on } \Omega \\ u = 0,\ Du = 0 \text{ in } \partial\Omega. \end{cases}$$

(Include a discussion of weak formulation, existence, uniqueness, and regularity under appropriate conditions on the given functions $f,\ g$.)

Lecture 8

Another Approach to Interior Regularity, and a General Half-space Boundary Regularity Lemma

Our aim here is to lay the groundwork for our later discussion (in Lecture 9) of boundary regularity for solutions of the Dirichlet problem and other boundary value problems.

We begin with a general interior regularity lemma for constant coefficient elliptic operators of the form $Lu = \sum_{|\alpha|\le m} a_\alpha D^\alpha u$. (Later we consider variable coefficients also.)

Let $m \ge 1$ be an integer, $M,\ \mu > 0$, and $\{a_\alpha\}_{|\alpha|\le m}$ be given constants. We suppose

(i) $$|a_\alpha| \le M, \qquad |\alpha| \le 2m,$$

and with the $\{a_\alpha\}_{|\alpha|=2m}$ satisfying the ellipticity condition

(ii) $$\textstyle\sum_{|\alpha|=2m} a_\alpha \xi^\alpha \ge \mu|\xi|^{2m}, \qquad \xi \in \mathbb{R}^n.$$

We consider weak solutions (in $L^2_{\mathrm{loc}}(\Omega)$) of equations of the form $Lu = f$. We in fact have the following interior regularity theorem, which generalizes the regularity results for constant coefficient operators proved in Lecture 6.

Theorem 1. *Suppose (1), (2) hold, $k \in \{0, \dots, 2m\}$, $\ell \geq 0$, and $u \in L^2_{\mathrm{loc}}(B_R(x_0))$ is a weak solution of the equation*

$$Lu = \textstyle\sum_{|\beta| \leq 2m-k} D^\beta f_\beta \text{ on } B_R(x_0), \text{ where } \|f_\beta\|_{\ell, B_R(x_0)} < \infty.$$

Then $u \in H^{k+\ell}_{\mathrm{loc}}(B_R(x_0))$, and in fact

$$\|u\|_{k+\ell, B_{\theta R}(x_0)} \leq C(\|u\|_{0, B_R(x_0)} + \textstyle\sum_{|\beta| \leq 2m-k} \|f_\beta\|_{\ell, B_R(x_0)}),$$

for each $\theta \in (0,1)$, where C depends only on n, M, μ, θ, R. (In particular $u \in C^\infty(B_R(x_0))$ if all the $f_\beta \in C^\infty(B_R(x_0))$.)

Proof: Recall from Lecture 2 that the mollified function u_σ satisfies the classical equation

$$(1) \qquad Lu_\sigma = \textstyle\sum_{|\beta| \leq 2m-k} D^\beta (f_\beta)_\sigma \text{ on } B_{R-\sigma}(x_0).$$

So assume $\theta \in (0,1)$, $\sigma < (1-\theta)R/2$, and let ψ be an arbitrary $C^\infty_c(B_R(x_0))$ function. Notice that by repeatedly using the Leibniz formula

$$D^\alpha(\psi v) = \textstyle\sum_{\gamma+\delta=\alpha} \frac{\alpha!}{\gamma!\delta!} (D^\gamma \psi)(D^\delta v),$$

we can write the equation (1) in the form

$$\begin{aligned}\textstyle\sum_{|\alpha|=2m} D^\alpha(a_\alpha \psi u_\sigma) &= \textstyle\sum_{|\gamma|+|\delta| \leq 2m, |\delta| \leq 2m-1} D^\delta(b_{\gamma\delta}(D^\gamma\psi)u_\sigma) \\ &+ \textstyle\sum_{|\delta|+|\gamma| \leq 2m-k, |\delta| \leq 2m-k-1} D^\delta(d_{\beta\gamma\delta}(D^\gamma\psi)(f_\beta)_\sigma + \textstyle\sum_{|\beta|=2m-k} D^\beta(\psi f_\beta)\end{aligned}$$

for suitable constant $b_{\gamma\delta}, d_{\beta\gamma\delta}$ with $|b_{\gamma\delta}| \leq CM$ and $|d_{\beta\gamma\delta}| \leq C$, where C depends only on n, m and k. Now taking Fourier transforms on each side we obtain

$$\begin{aligned}(2) \quad \textstyle\sum_{|\alpha|=2m} a_\alpha (-i\xi)^\alpha \widehat{\psi u_\sigma} &= \textstyle\sum_{|\gamma|+|\delta| \leq 2m, |\delta| \leq 2m-1} (i)^{|\delta|} \xi^\delta b_{\gamma\delta} \widehat{(D^\gamma\psi)u_\sigma)} \\ &+ \textstyle\sum_{|\delta|+|\gamma| \leq 2m-k, |\delta| \leq 2m-k-1} (i)^{|\delta|} \xi^\delta d_{\beta\gamma\delta} \widehat{(D^\gamma\psi)(f_\beta)_\sigma} \\ &+ \textstyle\sum_{|\beta|=2m-k} (i)^{2m-k} \xi^\beta \widehat{\psi f_\beta}\end{aligned}$$

Then using the ellipticity condition (2) we deduce that

$$\begin{aligned}(1+|\xi|)^{2m} |\widehat{\psi u_\sigma}| &\leq C(1+|\xi|)^{2m-1} \textstyle\sum_{|\gamma| \leq 2m} |\widehat{(D^\gamma\psi)u_\sigma}| \\ &+ C(1+|\xi|)^{2m-k-1} \textstyle\sum_{|\gamma| \leq 2m-k} |\widehat{(D^\gamma\psi)(f_\beta)_\sigma}| + C(1+|\xi|)^{2m-k} |\widehat{\psi f_\beta}|\end{aligned}$$

for all $\xi \in \mathbb{R}^n$, so, recalling the fact that the H^ℓ-norm of a function v with compact support in a ball $B_R(x_0)$ is equivalent to the L^2 norm of $(1+|\xi|)^\ell \hat{v}$, we see that by multiplying through by $(1+|\xi|)^{k+\ell-2m}$, where $k+\ell \geq 1$ (ℓ may be negative at this stage of the argument), that

$$\begin{aligned}(3) \quad \|\psi u_\sigma\|_{k+\ell, B_R(x_0)} &\leq \textstyle\sum_{|\gamma| \leq 2m} \|(D^\gamma\psi)u_\sigma\|_{k+\ell-1, B_R(x_0)} + \\ &\textstyle\sum_{|\gamma| \leq 2m-k} \|(D^\gamma\psi)(f_\beta)_\sigma\|_{(\ell-1)_+, B_R(x_0)} + C\|\psi(f_\beta)_\sigma\|_{\ell_+, B_R(x_0)},\end{aligned}$$

where j_+ means $\max(j, 0)$. Since ψ was an arbitrary $C^\infty_c(B_R(x_0))$ function (so that if $k+\ell \geq 2$ the argument can be repeated with $D^\gamma\psi$ in place of ψ), we conclude by induction on ℓ that

$$\begin{aligned}(4) \quad \|\psi u_\sigma\|_{k+\ell, B_R(x_0)} &\leq \textstyle\sum_{|\gamma| \leq 2(k+\ell)m} \|(D^\gamma\psi)u_\sigma\|_{0, B_R(x_0)} + \\ &\textstyle\sum_{|\gamma| \leq 2(k+\ell)m} \|(D^\gamma\psi)(f_\beta)_\sigma\|_{\ell, B_R(x_0)}.\end{aligned}$$

Now we select the function ψ such that $\psi \equiv 1$ in $B_{\theta R}(x_0)$, $\psi \equiv 0$ outside $B_{(1+\theta)R/2}(x_0)$ and

$$|D^\alpha \psi| \leq C(1-\theta)^{-|\alpha|} R^{-|\alpha|}$$

for each multi-index α, with $C = C(n, \alpha)$. Then (4) yields the required inequality.

Next we want to give a boundary regularity lemma which will be useful in the next lecture for discussing boundary regularity for solutions of boundary value problems.

Here we adopt the following notation: points in $\mathbb{R}^n$ will be denoted (x, y), where $x = (x^1, \dots, x^{n-1}) \in \mathbb{R}^{n-1}$, and we let

$$\mathbb{R}^n_+ = \{(x,y) \in \mathbb{R}^n : y > 0\}, \quad B^+_R = B_R(0) \cap \mathbb{R}^n_+.$$

For $x_0 \in \mathbb{R}^{n-1}$, we also let

$$B^+_R(x_0, 0) = B_R(x_0, 0) \cap \mathbb{R}^n_+.$$

For $|h| \neq 0$, we let δ^j_h denote the tangential difference quotient operators

$$\delta^j_h v(x) = h^{-1}(v(x+he_j) - v(x)), \quad j = 1, \dots, n-1,$$

valid whenever the values $v(x+he_j)$, $v(x)$ of v are defined. For multi-indices $\gamma = (\gamma_1, \dots, \gamma_{n-1})$ we let

$$\delta^\gamma_h = (\delta^1_h)^{\gamma_1} \circ \cdots \circ (\delta^{n-1}_h)^{\gamma_{n-1}}.$$

Using one dimensional calculus, one can check that if $k \geq 1$, $u \in H^k(B^+_R)$, and if $\theta \in (0,1)$, then

$$\text{(iii)} \qquad \|\delta^j_h u\|_{k-1, B^+_{\theta\mathbb{R}}} \leq \|u\|_{k, B^+_R}$$

provided $|h| \le (1-\theta)R$. (The formal proof of this is elementary, using a slight modification of problem 5.8).

The main lemma here is the following; the reader should keep in mind that the case $k_0 = m$ is the important case for the applications to boundary problems discussed in the next lecture.

Lemma. *Suppose m, k_0 are positive integers, $\ell \in \{0,1,\dots\}$, $\mu,\ M > 0$, suppose $u \in H^m(B_R^+)$ satisfies weakly in B_R^+ an equation of the form*

$$(*) \qquad \textstyle\sum_{|\beta|\le k_0} D^\beta(\sum_{|\alpha|\le m} a_{\alpha\beta} D^\alpha u) = \sum_{|\beta|\le k_0} D^\beta f_\beta,$$

where $a_{\alpha\beta},\ f_\beta$ are given functions of x with $a_{(0,k_0)(0,m)} \ge \mu$ and with $a_{\alpha\beta}$ having weak derivatives up to order ℓ satisfying $|\sum_{|\gamma|\le\ell} D^\gamma a_{\alpha\beta}| \le M$. Suppose further that we already have estimates of the form

$$(**) \qquad \textstyle\sum_{|\gamma|\le\ell} \|\delta_h^\gamma u\|_{m,B^+_{\theta^2 R}} \le K_\theta(\|u\|_{m-1,B^+_{\theta R}} + \sum_{|\beta|\le k_0,\,|\gamma|\le\ell} \|\delta_h^\gamma f_\beta\|_{0,B_{\theta R}})$$

for each $\theta \in (0,1)$ and all sufficiently small $|h| \ne 0$, where K_θ does not depend on h. Then

$$\|u\|_{m+\ell,B^+_{\theta R}} \le C(\|u\|_{m-1,B^+_R} + \textstyle\sum_{|\beta|\le k_0} \|f_\beta\|_{\ell,B^+_R})$$

for each $\theta \in (0,1)$, where C depends only on $n,\ \mu,\ R,\ \theta,\ M$, and the constants K_θ of the given inequality.

Remarks: (1) In practice the checking of $(**)$ for $\ell \ge 1$ is often an automatic consequence of the fact that it holds for $\ell = 0$ (in cases when we are working with u in a subspace which is invariant under application of the tangential difference quotient operators—see for example the discussion of the Dirichlet boundary value problem in the next lecture).

(2) The content of the lemma is roughly that if we have appropriate control of "tangential" derivatives in the half-space (i.e. derivatives with respect to the variables $x^1,\dots,x^{n-1}$), then we can also control the derivatives with respect to the y variables and the mixed derivatives, provided u satisfies a linear equation like $*$. Notice that the equation is not explicitly required to be elliptic, although in practice this would typically be needed to check $(**)$ anyway.

Before we begin the proof of the lemma, we recall the following important facts about extension of functions from the half-ball B_R^+ into the full ball $B_R(0)$:

Let $q \ge 1$ be arbitrary, and let $c_0,\dots,c_q$ be any real constants which satisfy the system of $(q+1)$-linear equations

$$\textstyle\sum_{j=0}^q c_j\left(-\frac{1}{j+1}\right)^k = 1, \quad k = 0,\dots,q.$$

Of course such c_j exist because the $(q+1)\times(q+1)$ matrix with k^{th} row $(-\frac{1}{1})^k \cdots (-\frac{1}{q+1})^k$, $k = 0,\dots,q$, has non-zero determinant. (Indeed it is an example of a "Vandermonde" determinant.)

Now we define an extension operators $T,\ T_p$ taking functions defined on B_R^+ to functions defined on $B_R(0)$ as follows:

$$Tu(x,y) = \begin{cases} u(x,y) \text{ if } y > 0 \\ \sum_{j=0}^q c_j u(x, -\frac{y}{j+1}) \text{ if } y < 0 \end{cases}$$

and for $p \in \{0,\dots,q-1\}$

$$T_p u(x,y) = \begin{cases} u(x,y) \text{ if } y > 0 \\ \sum_{j=0}^q c_j\left(-\frac{1}{j+1}\right)^p u(x, -\frac{y}{j+1}) \text{ if } y < 0 \end{cases}$$

(Notice that $T_0 \equiv T$.) If $k \le q-p$, one readily checks by direct computation that T_p is a bounded operator from $C^k(\overline{B}_R^+)$ to $C^k(\overline{B}_R(0))$ and from $H^k(B_R^+)$ to $H^k(B_R(0))$. Thus we have a fixed constant C (depending on q, R, and n only) such that

$$(\ddagger) \qquad \begin{aligned} &\|T_p u\|_{C^k(B_R(0))} \le C\|u\|_{C^k(B_R^+)}, \quad u \in C^k(\overline{B}_R^+) \\ &\|T_p u\|_{H^k(B_R(0))} \le C\|u\|_{H^k(B_R^+)}, \quad u \in H^k(B_R^+) \end{aligned}$$

Proof of Lemma: The first step is to note that $(**)$ implies that all the tangential derivatives $D_x^\gamma u$, $|\gamma| \le \ell$, exist and satisfy the estimates

$$(1) \qquad \textstyle\sum_{|\gamma|\le\ell} \|D_x^\gamma u\|_{m,B^+_{\theta^2 R}} \le K_\theta(\|u\|_{m-1,B^+_{\theta R}} + \sum_{|\beta|\le k_0} \|f_\beta\|_{\ell,B^+_{\theta R}}).$$

To see this note that if $\sum_{|\beta|\le k_0} \|f_\beta\|_{\ell,B^+_{\theta R}} = \infty$, then there is nothing to prove, hence assume the right side of (1) is finite. Then we have simply to observe first, by repeatedly applying (iii), that

$$\textstyle\sum_{|\gamma|\le\ell} \|\delta_h^\gamma f_\beta\|^2_{0,B^+_{\theta R}} \le \|f_\beta\|^2_{\ell,B^+_{\theta R+|h|}},$$

so that by $(**)$ the quantity $\sum_{|\gamma|\le\ell} \|\delta_h^\gamma u\|^2_{m,B^+_{\theta^2 R}}$ is bounded by the quantity on the right of (1) as $|h| \downarrow 0$. Using the definition of weak derivative (see in particular Lemma 7 of Lecture 5) we then have the required existence of $D_x^\gamma u$ and the bound (1).

Next we note that we can rewrite the identity $*$ in the form

(2) $$D_y^{k_0}(b_0 D_y^m u + \textstyle\sum_{|\gamma|+j\le m,\, j\le m-1} b_{\gamma j} D_x^\gamma D_y^j)$$
$$= -\textstyle\sum_{|\gamma|+j\le k_0,\, j\le k_0-1} D_y^j D_x^\gamma (\sum_{|\delta|\le m} b_{\gamma\delta j} D^\delta u) + \sum_{|\beta|\le k_0} D^\beta f_\beta,$$

where now $b_0 = a_{(0,k_0)(0,m)} \ge \mu$ and each $b_{\gamma j},\, b_{\gamma\delta j}$ is one of the $a_{\alpha\beta}$. In this case we define

$$v = b_0 D_y^m u + \textstyle\sum_{|\gamma|+j\le m,\, j\le m-1} b_{\gamma j} D_x^\gamma D_y^j$$
$$u_{\gamma j} = -\textstyle\sum_{|\delta|\le m} b_{\gamma\delta j} D^\delta u,$$

and note that (2) can be written in the form

(3) $$D_y^{k_0} v = \textstyle\sum_{|\gamma|+j\le k_0, j\le k_0-1} D_y^j D_x^\gamma u_{\gamma j} + \sum_{|\beta|\le k_0} D^\beta f_\beta.$$

Now notice that for each $\gamma,\, k,\, j$ with $k \le k_0$ and $|\gamma| + j \le k_0,\ j \le k_0 - 1$ we can write

$$D_y^j D_x^\gamma u_{\gamma j} = \textstyle\sum_{j\ge k} D_x^\gamma D_y^{j-k}(D_y^{k-1} u_{\gamma j}) + \sum_{j\le k-1} D_x^\gamma D_y^j u_{\gamma j},$$

and hence (3) can be written

$$D_y^{k_0} v = \textstyle\sum_{|\beta|\le k_0-k} D^\beta(\sum_{|\delta|+r\le k, r\le k-1} D_x^\delta D_y^r u_{\beta\delta r})$$
$$+ \textstyle\sum_{|\gamma|\le k_0-k} D^\gamma(\sum_{|\delta|\le k} D^\delta f_{\gamma\delta}),$$

where each $u_{\beta\delta r}$ is either identically zero or one of the $u_{\gamma j}$ and each $f_{\gamma\delta}$ is either identically zero or one of the f_β.

Now let $0 < \sigma < \tau < \theta^2 R$ be arbitrary, as usual let v_σ denote the mollification of v on $\{(x,y) \in B_R^+ : y > \sigma\}$, and let $v_\sigma^\tau(x,y) = v_\sigma(x, y+\tau)$ for $(x,y) \in \Omega_\tau^+ \equiv \{(x,y) : y > 0,\ (x, y+\tau) \in B_R^+\}$. Similarly define $(u_{\gamma j})_\sigma^\tau$ and $(f_\beta)_\sigma^\tau$ on Ω_τ^+.

Now let $q \ge 1 + \ell + m$, and let $T,\, T_p$ be the extension operators in ‡ above. Then notice that by direct computation we have

$$D_y^{k_0} T v_\sigma^\tau \equiv$$
$$T_{k_0}(D_y^{k_0} v_\sigma^\tau) = T_{k_0}(\textstyle\sum_{j\le k_0-1,\,|\gamma|+j\le k_0} D_x^\gamma D_y^j (u_{\gamma j})_\sigma^\tau + \sum_{|\beta|\le k_0} D^\beta (f_\beta)_\sigma^\tau)$$
$$= \textstyle\sum_{j\le k_0-1,\,|\gamma|+j\le k_0} D_x^\gamma D_y^j T_{k_0-j}(u_{\gamma j})_\sigma^\tau + \sum_{|\beta|\le k_0} D^\beta T_{k_0-\beta_n}(f_\beta)_\sigma^\tau.$$

Now if k_0 is even we can add $\sum_{j=1}^{n-1} D_{x^j}^{k_0} T v_\sigma^\tau$ to each side of the equation, thus giving

(4)
$$L T v_\sigma^\tau \equiv T_{k_0}(D_y^{k_0} v_\sigma^\tau) = \textstyle\sum_{j=1}^{n-1} D_{x^j}^{k_0} T v_\sigma^\tau +$$
$$\textstyle\sum_{j\le k_0-1,\,|\gamma|+j\le k_0} D_x^\gamma D_y^j T_{k_0-j}(u_{\gamma j})_\sigma^\tau + \sum_{|\beta|\le k_0} D^\beta T_{k_0-\beta_n}(f_\beta)_\sigma^\tau,$$

where L is the elliptic operator given by $Lv = (D_y^{k_0} + \sum_{j=1}^{n-1} D_{x^j}^{k_0})v$, and so we can apply the interior estimates of Theorem 1 above with k_0 in place of $2m$ and with the expression on the right of (4) in place of the expression $D^\beta f_\beta$ of Theorem 1. This gives for each $k \in \{1, \ldots, k_0\}$ that

(5) $$\|T v_\sigma^\tau\|_{k+\ell, B_{\theta^2 R}} \le C(\|T v_\sigma^\tau\|_{0,B_{\theta R}} + \textstyle\sum_{j\le k-1,\,|\gamma|+j\le k} \|D_x^\gamma D_y^j T v_\sigma^\tau\|_{\ell, B_{\theta R}}$$
$$+ \textstyle\sum_{j\le k-1,\,|\gamma|+j\le k} \|D_x^\gamma D_y^j T_{k_0-j}(u_{\gamma j})_\sigma^\tau\|_{\ell, B_{\theta R}}$$
$$+ \textstyle\sum_{|\beta|\le k_0,\,|\gamma|\le k} \|D^\gamma T_{k_0-\beta_n}(f_\beta)_\sigma^\tau\|_{\ell, B_{\theta R}}).$$

Now we use $T v_\sigma^\tau = v_\sigma^\tau$ on $B_{\theta R}^+$ on the left together with the inequalities ‡ on the right and let $\sigma \downarrow 0$ and $\tau \downarrow 0$ (we emphasize that the constant C in (5) is independent of $\sigma,\, \tau$), thus giving

(6) $$\|v\|_{k+\ell, B_{\theta^2 R}^+} \le C(\|v\|_{0, B_{\theta R}^+} + \textstyle\sum_{s\le k-1,\,|\gamma|+s\le k} \|D_x^\gamma D_y^s v\|_{\ell, B_{\theta R}^+}$$
$$+ \textstyle\sum_{j\le k-1,\,|\gamma|+j\le k} \|D_x^\gamma D_y^j u_{\gamma j}\|_{\ell, B_{\theta R}^+} + \sum_{|\beta|\le k_0} \|f_\beta\|_{k+\ell, B_{\theta R}^+}).$$

Notice that at this stage we must consider the possibility that some of the terms on the right (and hence possibly also the quantity on the left) might be infinite, but in any case the inequality (6) is valid. Now take first $k = 1,\ \ell = 0$ and keeping in mind the definitions (3), we then deduce directly that

$$\|D_y u\|_{m, B_{\theta^2 R}^+} \le C(\|u\|_{m, B_{\theta R}^+} + \textstyle\sum_{|\gamma|=1} \|D_x^\gamma u\|_m + \sum_{|\beta|\le k_0,\,|\gamma|\le 1} \|D^\gamma f_\beta\|_{0, B_{\theta R}^+}).$$

where we used the definition of v together with $b_0 \ge \mu$ and $\sum_{|\gamma|\le 1} |D^\gamma a_{\alpha\beta}| \le M$. In view of the bounds (1) this ensures that

$$\|u\|_{m+1, B_{\theta^3 R}^+} \le C(\|u\|_{m-1, B_{\theta R}^+} + \textstyle\sum_{|\beta|\le k_0} \|f_\beta\|_{1, B_{\theta R}^+}).$$

Proceeding inductively, the required higher estimates also now follow directly from (6).

Thus the proof is complete in case k_0 is even. If k_0 is odd we can (weakly) differentiate each side of the equation $*$ to give an identity of similar form with $\widetilde{k}_0 = k_0 + 1$ in place of k_0. That is, weakly, we have an identity of the form

$$D_y^{\widetilde{k}_0}(b_0 D_y^m u + \textstyle\sum_{|\beta|+j\le m,\, j\le m-1} b_{\gamma j} D_x^\gamma D_y^j u)$$
$$= -\textstyle\sum_{|\gamma|+j\le \widetilde{k}_0,\, j\le \widetilde{k}_0-1} D_y^j D_x^\gamma (\sum_{|\delta|\le m} b_{\gamma\delta j} D^\delta u) + \sum_{|\beta|\le k_0} D_y D^\beta f_\beta,$$

where again all $b_{\gamma j},\, b_{\gamma\delta j}$ are either identically zero or equal to one of the $a_{\alpha\beta}$. Thus we can repeat the above argument with $\widetilde{k}_0$ in place of k_0 to again deduce the required estimates.

Lecture 9

Boundary Regularity for the Dirichlet Problem and Other Elliptic Boundary-value Problems

Here we want to establish local regularity results near the boundary for the solutions of the Dirichlet problem discussed in the previous lecture. The results here are entirely analogous to the interior regularity results of Corollary 1 of Lecture 6, provided the appropriate smoothness is assumed for the boundary of the domain.

To make such assumptions precise, we need the following definition of C^k domain; notice the definition here is essentially identical to our definition of Lipschitz domain given in Lecture 5, except that here we require the function ψ to be in class C^k rather than in $C^{0,1}$ as before.

Definition 1: Let $x_0 \in \partial\Omega$ and $R > 0$. We say that $\partial\Omega \cap B_R(x_0)$ is C^k if there is a C^k function $\psi : \mathbb{R}^{n-1} \to \mathbb{R}$ and coordinates $y = x_0 + Q(x - x_0)$, with Q an orthogonal transformation of $\mathbb{R}^n$, such that

$$\Omega \cap B_R(x_0) = \{x_0 + Q^{-1}(y - x_0) : y^n > \psi(y^1, \dots, y^{n-1})\} \cap B_R(x_0).$$

Notice that then in particular we have

$$\begin{aligned}\partial\Omega \cap B_R(x_0) &= \{x_0 + Q^{-1}(y - x_0) : y^n = \psi(y^1, \dots, y^{n-1})\} \cap B_R(x_0)\\ &\equiv \{x_0 + Q^{-1}(y - x_0) : y \in \operatorname{graph}\psi \cap B_R(0)\}.\end{aligned}$$

We shall not explicitly assume the ellipticity condition (E) here, but we shall assume at least that we have the coercivity condition

$$\text{(C)} \qquad \int_\Omega \textstyle\sum_{|\alpha|=|\beta|=m} a_{\alpha\beta} D^\alpha\varphi D^\beta\varphi \geq \mu \int_\Omega \sum_{|\alpha|=m} |D^\alpha\varphi|^2 - C\|\varphi\|^2_{0,\Omega}$$

$\forall\, \varphi \in H_0^m(\Omega)$.

Recall that (as discussed in the Lecture 7) the Gårding inequality guarantees that (C) is satisfied if the $a_{\alpha\beta}$ are continuous and if the ellipticity condition (E)$'$ of Lecture 7 holds.

So let us assume that $u \in H_0^m(\Omega)$ satisfies

$$(-1)^m \textstyle\sum_{|\alpha|,|\beta|\leq m} D^\beta(a_{\alpha\beta} D^\alpha u) = \sum_{|\beta|\leq m} D^\beta f_\beta,$$

in the weak sense that

$$(**) \qquad \int_\Omega \textstyle\sum_{|\alpha|,|\beta|\leq m} (-1)^{m+|\beta|} a_{\alpha\beta} D^\alpha u D^\beta\zeta = \int_\Omega \sum_{|\beta|\leq m} (-1)^{|\beta|} f_\beta D^\beta\zeta$$

for all $\zeta \in H_0^m(\Omega)$, where $f_\beta \in H^k(\Omega)$. We here also assume that (as in Lecture 6) the coefficients $a_{\alpha\beta}$ have bounded (weak) derivatives of order up to and including k:

$$(\mathrm{B}_k) \qquad \sup_{\Omega\cap B_R(x_0)} |D^\gamma a_{\alpha\beta}| \leq M, \quad |\alpha| \leq k.$$

Then we have the following boundary regularity theorem:

Theorem 1. *Suppose $x_0 \in \partial\Omega$, $\partial\Omega \cap B_R(x_0)$ is C^{m+k} in the sense of the Definition 1 with $m+k$ in place of k and with $|\psi|_{C^{m+k}} \leq M$. Suppose also that (C), (B_k) hold, that $u \in H_0^m(\Omega)$ and that $k \geq 0$. Then for any $\theta \in (0,1)$*

$$\|u\|_{m+k,\Omega\cap B_{\theta R}} \leq C\big(\|u\|_{0,\Omega\cap B_R(x_0)} + \textstyle\sum_{|\beta|\leq m}\|f_\beta\|_{k,\Omega\cap B_R(x_0)}\big),$$

where C depends only on m, k, θ, M, μ, R. If $\ell \in \{0,1,\dots\}$ and $k > \frac{n}{2} + \ell$, then

$$|u|_{C^\ell(\Omega\cap B_{\theta R}(x_0))} \leq C\big(\|u\|_{0,\Omega\cap B_R(x_0)} + \textstyle\sum_{|\beta|\leq m}\|f_\beta\|_{k,\Omega\cap B_R(x_0)}\big).$$

Remarks: (1) The last part of the above theorem follows directly from the first by virtue of the Sobolev embedding theorem (Lemma 3 of Lecture 5).

(2) The assumption that $u \in H_0^m(\Omega)$ can be replaced by the weaker assumption that $\zeta u \in H_0^m(\Omega \cap B_R(x_0))$ for any $\zeta \in C_c^\infty(B_R(x_0))$ without necessitating any change in the proof to be given below.

Proof of Theorem 1: The proof is based on the estimates in $\frac{1}{2}$-balls given in the Lemma of Lecture 8. But to apply this we first have to make a coordinate transformation which "flattens" $\partial\Omega$ near x_0. So let ψ be the function on $\mathbb{R}^{n-1}$ as in the definition of C^k boundary given above, but with $m+k$ in place of k, and assume, as in the statement of the theorem that $|\psi|_{C^{m+k}(\mathbb{R}^{n-1})} \leq M$. Noting that we may without loss of generality assume that $x_0 = 0$, we consider the coordinate transformation

$$x = (x^1, \dots, x^n) \mapsto \xi = (y^1, \dots, y^{n-1}, y^n - \psi(y^1, \dots, y^{n-1})),$$

where the y are as in Definition 1 above. Notice that this transformation is invertable; indeed the inverse is given explicitly by

$$\xi \mapsto Q^{-1}((\xi^1, \dots, \xi^{n-1}, \xi^n + \psi(\xi^1, \dots, \xi^{n-1})),$$

with Q as in Definition 1) and takes $B_R(x_0)$ onto a neighbourhood U of $0 \in \mathbb{R}^n$ in such a way that $\Omega \cap B_R(x_0)$ is mapped to $U_+ \equiv U \cap \mathbb{R}^n_+$, where $\mathbb{R}^n_+ = \{x \in \mathbb{R}^n : x^n > 0\}$, and $\partial\Omega \cap B_R(x_0)$ is mapped to $U \cap \mathbb{R}^{n-1} \times \{0\}$. Thus the coordinate transformation does indeed flatten the boundary. Further, if C is chosen large enough (depending on θ and M), it is easy to show that

$$(1) \qquad B^+_{\theta R} \subset U^+_{\theta' R} \text{ and } U^+_{\theta R} \subset B^+_{\theta' R}.$$

for suitable $\theta' \in (0,1)$ depending only on M and θ, where we use the notation that U^+_ρ is the image of $B_\rho \cap \Omega$ under the coordinate transformation.

Now, with x, ξ always related according to the above coordinate transformation, we define $\tilde u$ on U_+ by $\tilde u(\xi) \equiv u(x)$. Using the chain rule repeatedly and the transformation formula for the integral, it is easy to check that $\tilde u$ satisfies an equation of the same form as $(**)$ but with $\tilde u$ in place of u, with U_+ in place of Ω, and with new coefficients $\tilde a_{\alpha\beta}$ in place of $a_{\alpha\beta}$. Thus

$$(\ddagger) \qquad \int_{U_+} \textstyle\sum_{|\alpha|,|\beta|\leq m} (-1)^{m+|\beta|} \tilde a_{\alpha\beta} D^\alpha \tilde u D^\beta\zeta = \int_{U_+} \sum_{|\beta|\leq m} (-1)^{|\beta|} \tilde f_\beta D^\beta\zeta$$

for $\zeta \in H_0^m(U_+)$.

Notice also that (B_k) and (C) also continue to hold with new constants $\widetilde M$, $\tilde\mu = C\mu$ in place of M, μ, where $\widetilde M$ and C are determined by M and n. Finally we note that (again by the chain rule) that the Sobolev norms are related by

$$(2) \qquad C^{-1}\|u\|_{j,A} \leq \|\tilde u\|_{j,\tilde A} \leq C\|u\|_{j,A}$$

for any open $A \subset \Omega \cap B_R(x_0)$ and any $j \le m+k$ such that the norm $\|\widetilde{u}\|_{j,\widetilde{A}} < \infty$; here $\widetilde{A}$ means the image of A under the coordinate transformation and C depends only on M, n.

Now, to complete the proof, we note that the argument of Lemma 1 of of Lecture 6 carries over to the present setting almost without change, giving

$$\|\widetilde{u}\|_{m,B^+_{\theta R}} \le C\,(\|\widetilde{u}\|_{m-1,B^+_R} + \textstyle\sum_{|\beta|\le m}\|\widetilde{f}_\beta\|_{0,B^+_R}), \tag{3}$$

where C depends only on n, M, μ, and R. But of course now, if δ_h^j are the tangential operators of the previous lecture, then since we have the product rule $\delta_h^j(fg)(x) = (\delta_h^j f(x))g(x+he_j) + f(x)\delta_h^j g$ $\delta_h^j\widetilde{u}$ satisfies (weakly) the equation

$$\textstyle\sum_{|\alpha|,|\beta|\le m} D^\beta(\widetilde{a}_{\alpha\beta}(x+he_j)D^\alpha(\delta_h^j\widetilde{u})) = -\sum_{|\alpha|,|\beta|\le m} D^\beta((\delta_h^j\widetilde{a}_{\alpha\beta})D^\alpha\widetilde{u} + D^\beta f_\beta,$$

so we can apply the same estimate (3) to $\delta_h^j\widetilde{u}$ with $\delta_h^\gamma\widetilde{f}_\beta + \sum_{|\alpha|\le m}((\delta_h^j\widetilde{a}_{\alpha\beta})D^\alpha\widetilde{u}$ in place of $\widetilde{f}_\beta$. This gives (since we have $|\delta_h\widetilde{a}_{\alpha\beta}| \le C$ for fixed C independent of h)

$$\|\delta_h^j\widetilde{u}\|_{m,B^+_{\theta^2 R}} \le C\,(\|\delta_h^j\widetilde{u}\|_{m-1,B^+_{\theta R}} + \textstyle\sum_{|\beta|\le m}\|\delta_h^j\widetilde{f}_\beta\|_{0,B^+_{\theta R}}), \tag{3}$$

provided $|h|$ is small enough, where C depends only on M, μ, R, θ. Now (see Lecture 5) we have $\|\delta_h^j\widetilde{u}\|_{m-1,B^+_{\theta R}} \le C\|\widetilde{u}\|_{m,B^+_R}$ and $\|\delta_h^j\widetilde{f}_\beta\|_{0,B^+_{\theta R}} \le \|\widetilde{f}_\beta\|_{1,B^+_R}$ for $|h|$ small enough, so proceeding inductively this shows that

$$\textstyle\sum_{|\gamma|\le\ell}\|\delta_h^\gamma\widetilde{u}\|_{m,B^+_{\theta^2 R}} \le C\,(\|\widetilde{u}\|_{m-1,B^+_{\theta R}} + \sum_{|\beta|\le m}\|\widetilde{f}_\beta\|_{\ell,B^+_{\theta R}}), \tag{4}$$

provided $\sum_{|\gamma|\le\ell}|D^\gamma\widetilde{a}_{\alpha\beta}| \le M$.

In view of (4) we have precisely the hypotheses needed to apply the lemma of the previous lecture with $k_0 = m$, hence we deduce the required estimates, and Theorem 1 is proved.

Next we want to extend our discussion of the Dirichlet problem to other boundary value problems. We restrict ourselves to those problems which can be handled by only trivial modification of the methods we used for the Dirichlet problem in Lecture 7. Viz., we treat the so-called partially free boundary conditions. (The terminology is explained below.)

While these boundary conditions are fairly special when compared to the the most general conditions which can be handled, they are sufficient for most

practical purposes. In particular, as we show in the discussion below, the partially free boundary conditions considered here include the Neumann boundary condition and the oblique derivative boundary condition in the second order case (i.e. the case $m = 1$), as well as many of the important boundary conditions in the higher order case.

For a discussion of more general elliptic boundary value problems the reader is referred to [Morrey], [ADN].

We in fact consider here the boundary-value problems corresponding to formally replacing the space $H_0^m(\Omega)$ in the weak formulation of the Dirichlet problem (see (P_λ) of the previous lecture) by any of the subspaces H_k given by

$$H_k = H^m(\Omega) \cap H_0^k(\Omega), \tag{$*$}$$

where $k \in \{0, 1, \ldots, m\}$. (Here we continue to use the convention that $H^0(\Omega) \equiv L^2(\Omega)$, so that if $k = 0$, H_k is simply the entire Sobolev space $H^m(\Omega)$; notice also that $H_m = H_0^m(\Omega)$, which corresponds exactly to the Dirichlet problem considered in the last lecture.)

The reader should keep in mind that the Gårding inequality proved in the last lecture does not carry over automatically to this more general setting, although of course (C) is still guaranteed if the stronger ellipticity condition (E) of the previous lecture holds.

We can also copy the proof of the existence theory of Lecture 7 for the problem (P^k_λ), where (P^k_λ) is the same problem as $(\mathrm{P})_\lambda$ of Lecture 7 except that now we have H_k in place of $H_0^m(\Omega)$. The proofs go through without any change at all, provided again that the domain Ω is Lipschitz so as to ensure that the interpolation inequality and Rellich's theorem (Lemmas 6 and 4 of Lecture 5) are available. Also, if ‡ holds for $\widetilde{u} \in H^m(U_+)$ and $\zeta \in H$, with H any subspace of $H^m(U_+)$ which includes ψu for any $\psi \in C_c^\infty(B_{\theta R})$ and which is closed under the operation of taking tangential difference quotients δ_h^j for sufficiently small $|h|$, then we can repeat the argument above to again get estimates like those in inequality (3), and hence again the lemma of Lecture 8 can be applied to give estimates like those in Theorem 1 in this more general setting. In particular, all this is applicable to the case when H is one of the spaces H_k above. That is, we have:

Theorem 2. *Suppose that $k \in \{0, 1, \ldots, m\}$, the coercivity condition (C) and the boundedness condition* (B) *hold, and let Ω be a bounded Lipschitz domain. Then there is is a discrete set $\Lambda \subset \mathbb{R}$ (possibly empty), such that the conclusions of the* *theorem of Lecture 7 (including the additional smoothness conclusions in the corollary in case the coefficients and the domain are smooth) all hold with* (P^k_λ), H_k *in place of* $(\mathrm{P})_\lambda$, $H_0^m(\Omega)$ *respectively. The boundary regularity results of Theorem 1 above also hold if* H *is replaced by one of the spaces* H_k.

Notice that here we have adopted the reverse procedure to that of Lecture 7 in the sense that in Lecture 7 we started with a classical problem (namely, the classical Dirichlet problem) and then gave a weak formulation. Here on the other hand we start with the weak formulation (P^k_λ), and the natural question of course is what classical problem does the weak problem (P^k_λ) correspond to. To answer this question in full, it is first necessary to make a brief discussion of integration by parts over $\partial\Omega$. We here assume that $\partial\Omega$ is smooth enough, say C^ℓ for some $\ell \ge 2$. We let d be the distance function for $\partial\Omega$, defined by $d(x) = \operatorname{dist}(x, \partial\Omega)$ for $x \in \overline{\Omega}$. Recall that this function is C^ℓ in a boundary strip $S_\sigma = \{x \in \overline{\Omega} : d(x) \le \sigma\}$ for some $\sigma > 0$, and the unit normal of $\partial\Omega$ at points $y \in \partial\Omega$ is just $\eta(y) \equiv Dd(y)$. Let us assume that each component η^j of η is extended to be $C^1(\overline{\Omega})$ in such a way that $\eta \equiv Dd$ on the boundary strip S_σ. Then on $\overline{\Omega}$, we define the first order operators $\delta_1, \dots, \delta_n$ by

$$\delta_j = \textstyle\sum_{i=1}^n (\delta_{ij} - \eta^i\eta^j) D_i, \quad j = 1, \dots, n,$$

and for $\psi \in C^1(\partial\Omega)$, we define

$$\delta\psi(x) = (\delta_1\psi(x), \dots, \delta_n\psi(x)).$$

δ is tangential on $\partial\Omega$ in the sense that $\delta\psi(x)$ is tangent to $\partial\Omega$ (i.e. normal to η) at each boundary point $x \in \partial\Omega$. As a matter of fact, $\delta\psi(x)$ is evidently exactly the orthogonal projection of the ordinary gradient $D\psi(x)$ onto the tangent space of $\partial\Omega$ at x—i.e. the orthogonal projection of $D\psi(x)$ onto the subspace of $\mathbb{R}^n$ normal to $\eta(x)$.

We claim now that if $\psi, \varphi \in C^1(\overline{\Omega})$, then on $\partial\Omega$ we have the integration by parts formulae

$$\text{(I)} \qquad \int_{\partial\Omega} \psi\delta_j\varphi = -\int_{\partial\Omega}\varphi\delta_j\psi + \int_{\partial\Omega}\psi\varphi H\eta^j, \quad j = 1, \dots, n,$$

where $H = \Delta d|_{\partial\Omega}$. (Geometrically, H is just n times the mean curvature of $\partial\Omega$, but we won't make use of this here.)

These integration by parts formulae are easily proved. We just use the divergence theorem over Ω to evaluate each side of the identity $\int_\Omega (d_k\varphi)_{ki} = \int_\Omega (d_k\varphi)_{ik}$ (where subscripts denote partial derivatives in the indicated variable), and this gives the required identity after replacing φ by $\varphi\psi$. See Exercise 9.3 below. (Notice that it is necessary to use the identity $\sum_k d_k d_{ik} \equiv 0$, which follows from differentiation of the identity $|Dd|^2 \equiv 1$.) Once we have the integration by parts formulae (I) above, it is easy to see (by repeated use of the divergence theorem to integrate by parts in Ω and by repeated use of the formulae (I) on $\partial\Omega$), that there is a general identity of the form

$$\int_\Omega (Lu)\varphi = \int_\Omega \textstyle\sum_{|\alpha|,|\beta|\le m}(-1)^{m+|\beta|} a_{\alpha\beta} D^\alpha u D^\beta\varphi - \displaystyle\int_{\partial\Omega}\textstyle\sum_{j=0}^{m-1} B_{2m-j-1}(u) D^j_\eta\varphi,$$

for $u, \varphi \in C^\infty(\overline{\Omega})$, where L is the order $2m$ operator

$$Lu = (-1)^m \textstyle\sum_{|\alpha|,|\beta|\le m} D^\beta(a_{\alpha\beta}D^\alpha u), \quad u \in C^\infty(\overline{\Omega}),$$

where $D^j_\eta f$ means $\left.\dfrac{d^j}{ds^j}\right|_{s=0} f(x + s\eta) \equiv \sum_{|\alpha|=j}\eta^\alpha D^\alpha f(x)$ for $x \in \partial\Omega$, and where each B_{2m-j-1} is a linear partial differential operator of order $2m - j - 1$, $j = 0, \dots, m-1$. (See Exercise 9.2 below.) The operators B_j, $j = m, \dots, 2m-1$ are called natural boundary operators associated with L. (They are uniquely determined by the coefficients $a_{\alpha\beta}$, but one should keep in mind that there are many choices of $a_{\alpha\beta}$ which give the same operator L. (See Problem 6.4 above.)

Now if we take φ to lie in $C^{m-1}(\overline{\Omega}) \cap H_k$ (so that $D^\alpha\varphi \equiv 0$ on $\partial\Omega$, $|\alpha| \le k-1$ by Exercise 5.7), if the coefficients $a_{\alpha\beta} \in C^m(\overline{\Omega})$ if $f_\beta \in C^m$, and if $u \in C^{2m}(\overline{\Omega})$ is a classical solution of the equation $Lu = 0$, then (in view of the arbitrariness of the function φ), the identity $(**)$ evidently implies that

$$B_{2m-j-1}(u) = 0, \text{ on } \partial\Omega, \quad j = k, \dots, m-1.$$

Thus in particular if $a_{\alpha\beta}, f_\beta \in C^\infty(\overline{\Omega})$ and if $\partial\Omega$ is C^∞, then any weak solution of the problem (P^k_λ) is a classical solution of the problem

$$\begin{cases} Lu = \sum_{|\beta|\le m} D^\beta f_\beta \text{ on } \Omega \\ D^\alpha u = 0 \text{ on } \partial\Omega,\ |\alpha| \le k-1,\ B_j(u) = 0 \text{ on } \partial\Omega,\ j = m, \dots, 2m-k-1. \end{cases}$$

Notice that of course there is redundancy in the first group of boundary conditions here; the above problem can be equivalently written in terms of exactly m boundary conditions as

$$\begin{cases} Lu = \sum_{|\beta|\le m} D^\beta f_\beta \text{ on } \Omega \\ D^j_\eta u = 0 \text{ on } \partial\Omega,\ 0 \le j \le k-1,\ B_j(u) = 0 \text{ on } \partial\Omega,\ j = m, \dots, 2m-k-1. \end{cases}$$

We now want to discuss some examples which illustrate the kind of boundary-value problems which are included in the above discussion.

Example 1: The Neumann problem for second order elliptic opertors.

Here $m = 1$, so the operator Lu is given by

$$Lu = -\left(\sum_{i,j=1}^n D_j(a_{ij} D_i u) + \sum_{j=1}^n D_j(b_j u) + \sum_{j=1}^n c_j D_j u + du\right).$$

With the coefficients as indicated, we get natural boundary operator B_1 given by

$$B_1(u) = \sum_{i,j=1}^n \eta^j a_{ij} D_i u + \sum_{j=1}^n \eta^j b_j u,$$

and hence the natural boundary operator $B_1(u)$ is given by

$$B_1(u) = \sum_{i,j=1}^n \eta^j a_{ij} D_i u + \left(\sum_{j=1}^n \eta^j b_j\right) u,$$

so if $b_j = 0$ the above discussion enables us to treat the boundary-value problem

$$\begin{cases} Lu = \lambda u + f \text{ on } \Omega \\ \sum_{i,j=1}^n \eta^j a_{ij} D_i u = 0 \text{ on } \partial\Omega, \end{cases}$$

which is called the Neumann problem for the operator L.

Example 2: The second order oblique derivative problem.

Here again $m = 1$ and the operator is as in Example 1 above. Now there are many alternative ways of choosing coefficients to represent the operator L. For example, we can choose any γ, ε_{ij} smooth on $\overline{\Omega}$ with $\varepsilon_{ij} = -\varepsilon_{ji}$, and then represent the operator L in the form

$$Lu = -\left(\sum_{i,j=1}^n D_j(\widetilde{a}_{ij} D_i u) + \sum_{j=1}^n D_j(\widetilde{b}_j u) + \sum_{j=1}^n \widetilde{c}_j D_j u + \widetilde{d} u\right),$$

where we use the notation

$$\widetilde{a}_{ij} = a_{ij} + \varepsilon_{ij},\ \widetilde{b}_j = b_j + \eta^j \gamma,\ \widetilde{c}_j = c_j - \sum_j D_j \varepsilon_{ij} - \gamma \eta^j,\ \widetilde{d} = d - \sum_j D_j(\eta^j \gamma),$$

and note that in this case the natural boundary operator B_1 is given by

$$B_1 = \sum_{i,j=1}^n \eta_j (a_{ij} + \varepsilon_{ij}) D_i u + \sum_{j=1}^n \eta^j b_j u + \gamma u.$$

This illustrates clearly how the natural boundary operator can be changed while the operator L itself remains fixed; in this case we can choose ε_{ij}, γ to ensure that B_1 is actually given by

$$\widetilde{B}_1(u) = \alpha \cdot Du + \beta u,$$

where α, β are any preassigned smooth functions on $\partial\Omega$ such that $\alpha \cdot \eta \neq 0$. (See Exercise 9.5 below.) Thus we can include in the framework considered above the boundary-value problem

$$\begin{cases} Lu = \lambda u + f \text{ on } \Omega \\ \alpha \cdot Du + \beta u = 0 \text{ on } \partial\Omega, \end{cases}$$

where α, β are any smooth functions on $\partial\Omega$ such that $\alpha \cdot \eta \neq 0$ on $\partial\Omega$. This is the oblique derivative problem for the operator L. (Notice that it formally includes the Neumann problem discussed in Example 1 above.)

Example 3: Boundary-value problems for the Biharmonic Operator

We consider the operator $Lu \equiv \Delta^2 u$. In this case $m = 2$, so there are two natural boundary operators, of order 2 and 3. As in the previous example, we get different natural boundary operators according as to how we select the coefficients. For example if we write the operator L in the form

$$Lu = \sum_{i,j,k,\ell} D_k D_\ell (\delta_{k\ell} \delta_{ij} D_i D_j u),$$

then (with coefficients $\delta_{k\ell}\delta_{ij}$) after integration by parts twice and by using the integration by parts formula (I) as in the general procedure for computing the natural boundary operators described above, we obtain

$$\int_\Omega \varphi \Delta^2 u = \int_\Omega \sum_{i,j,k,\ell=1}^n \delta_{ij}\delta_{k\ell}(D_i D_j u)(D_k D_\ell \varphi) + \int_{\partial\Omega} (D_\eta(\Delta u)\varphi + \Delta u D_\eta \varphi),$$

so the natural boundary operators in this case are

$$B_3(u) = D_\eta \Delta u, \quad B_2(u) = \Delta u.$$

Therefore for this choice of coefficients the theory developed above makes it possible for us to treat the boundary-value problems

(1) $$\begin{cases} \Delta^2 u = \lambda u + f \text{ on } \Omega \\ u = 0, \quad D_\eta u = 0 \text{ on } \partial\Omega \end{cases}$$

(2) $$\begin{cases} \Delta^2 u = \lambda u + f \text{ on } \Omega \\ u = 0, \quad \Delta u = 0 \text{ on } \partial\Omega \end{cases}$$

(3) $$\begin{cases} \Delta^2 u = \lambda u + f \text{ on } \Omega \\ \Delta u = 0, \quad D_\eta \Delta u = 0 \text{ on } \partial\Omega \end{cases}$$

Notice that (1) is just the Dirichlet problem. On the other hand if we write the operator $\Delta^2 u$ in the form

$$\Delta^2 u = \sum_{i,j,k,\ell=1}^n \delta_{ik}\delta_{j\ell}(D_i D_j u)(D_k D_\ell \varphi),$$

then after two integration by parts we get the identity

$$\int_\Omega \varphi\Delta^2 u = \int_\Omega \textstyle\sum_{i,j,k,\ell=1}^n \delta_{ij}\delta_{k\ell}(D_iD_ju)(D_kD_\ell\varphi) +$$
$$\int_{\partial\Omega}((D_\eta(\Delta u) - H\eta^i\eta^jD_iD_ju - \delta_j(\eta^jD_iD_ju))\varphi - \eta^j\eta^iD_iD_juD_\eta\varphi).$$

Thus in this case, after some simplification (see Exercise 9.7), we get natural boundary operators

$$B_3(u) = D_\eta^3 u - \textstyle\sum_{i,j=1}^n \alpha_{ij}D_iD_ju,\quad B_2(u) = D_\eta^2 u,$$

where we continue to use the notation $D_\eta^2 = \dfrac{d^2}{ds^2}\Big|_{s=0} u(x+s\eta)$, and where $\alpha_{ij} = \delta_j\eta^i$ ($\equiv D_iD_jd$ in terms of the distance function). Then we can write down new boundary-value problems corresponding to (2), (3) above. There are of course a large class of other boundary conditions corresponding to other choices for the coefficients. (See Exercise 9.6.)

LECTURE 8 PROBLEMS

9.1 Discuss the solvability of each of the problems (2), (3) above in case $\lambda = 0$.

9.2 Prove using the divergence theorem over Ω and the integration by parts formulae (I) over $\partial\Omega$ that the boundary operators B_j, $j = m,\dots,2m-1$ introduced above always exist in case the domain Ω is at least C^2 and the coefficients $a_{\alpha\beta} \in C^m(\overline{\Omega})$.

9.3 (i) Check the proof of the integration by parts formulae (I), and that the function $H : \partial\Omega \to \mathbb{R}$ in that formula is unique.

(ii) The Laplacian $\Delta_{\partial\Omega}$ on $\partial\Omega$ is defined by $\Delta_{\partial\Omega}\psi = \sum_{j=1}^n \delta_j(\delta_j\psi)$ on $\partial\Omega$, assuming $\psi \in C^2(\overline{\Omega})$. Prove that $\Delta_{\partial\Omega}\psi$ is given by

$$\textstyle\sum_{i,j=1}^n(\delta_{ij} - \eta^i\eta^j)D_iD_j\psi - H\sum_{j=1}^n \eta^jD_j\psi,\quad \psi \in C^\infty(\overline{\Omega}).$$

9.4 If the coordinate transformation $x \mapsto \xi$ is as described in the proof of Theorem 1 above, prove that (as claimed in that proof) the coercivity condition (C) continues to hold with $\tilde{\mu} = C\mu$, where C depends only on M, n.

9.5 Check the claim made above that for any given a smooth domain Ω and functions $a_{ij}, b_j, \alpha_j, \beta \in C^\infty(\overline{\Omega})$ with (a_{ij}) positive definite, we can find $\varepsilon_{ij}, \gamma \in C^\infty(\overline{\Omega})$ such that

$$\textstyle\sum_{j=1}^n \eta^j(a_{ij}+\varepsilon_{ij}) = \alpha_i,\ i = 1,\dots,n,\quad \sum_{j=1}^n \eta^jb_j + \gamma = \beta \text{ on } \partial\Omega.$$

9.6 Show that for the biharmonic operator Δ^2 it is possible to select coefficients such that the boundary operator B_2 is given by

$$B_2(u) = \Delta u + a\cdot Du + bu$$

for any given $a = (a^1,\dots,a^n)$, $b \in C^\infty(\overline{\Omega})$. Show also that it can be arranged that

$$B_2(u) = D_\eta^2 u + a\cdot Du + bu.$$

9.7 (i) Check the claim above that the choice of coefficients $\delta_{ij}\delta_{k\ell}$ for the biharmonic operator Δ^2 leads to the natural boundary operators

$$B_2(u) = D_\eta^2 u,\quad B_3(u) = D_\eta^3 - \textstyle\sum_{i,j=1}^n \alpha_{ij}D_iD_ju.$$

(ii) If $a_{ij} = a_{ji} \in C^\infty(\overline{\Omega})$ and if $\sum_{j=1}^n \eta^ja_{ij} \equiv 0$ on $\partial\Omega$, prove that it is possible to choose coefficients for Δ^2 such that the natural boundary operators are

$$B_2(u) = D_\eta^2 u,\quad B_3(u) = D_\eta^3 u - \textstyle\sum_{i,j=1}^n a_{ij}D_iD_ju.$$

Lecture 10

Spectrum of Self-adjoint Operators

Here we want to consider the nature of the set Λ introduced in Lecture 7 in case the problem $\mathrm{P}_\lambda(0)$ is self-adjoint; that is in case $A(\varphi,\psi) = A(\psi,\varphi)\ \forall \varphi,\psi \in \mathcal{H} \equiv H_0^m(\Omega)$. Recall that the notation is as follows: Ω is a bounded domain in $\mathbb{R}^n$, Λ is the set of $\lambda \in \mathbb{R}$ such that the problem $\mathrm{P}_\lambda(0)$ has non-trivial (i.e. non-zero) solutions; that is, there exists $u \in \mathcal{H} \setminus \{0\}$ ($\mathcal{H} = H_0^m(\Omega)$) such that

$$\mathrm{P}_\lambda(0) \qquad A(u,\psi) = \lambda\langle u,\psi\rangle_{L^2(\Omega)}\ \forall \psi \in \mathcal{H},$$

where $A(\varphi,\psi)$ is the bilinear form given by

$$A(\varphi,\psi) = \int_\Omega \sum_{|\alpha|,|\beta|\le m} (-1)^{m+|\beta|} a_{\alpha\beta} D^\alpha \varphi D^\beta \psi\, dx.$$

As in Lecture 7 we assume $a_{\alpha\beta} \in L^\infty(\Omega)$—that is, there is constant $M > 0$ such that

$$\text{(B)} \qquad |a_{\alpha\beta}| \le M, \quad |\alpha|,|\beta| \le m,$$

and we also assume that either condition (E) of Lecture 7 holds or alternatively that the $a_{\alpha,\beta}$ with $|\alpha| = |\beta| = m$ are $C^0(\overline{\Omega})$ and the weaker ellipticity condition (E)$'$ of Lecture 7 holds. From the discussion of Lecture 7 recall that either alternative gives us the coercivity condition that there are constants $\mu > 0$, $\gamma \ge 0$ such that

$$\text{(C)} \quad \int_\Omega \sum_{|\alpha|=|\beta|=m} a_{\alpha\beta} D^\alpha \varphi D^\beta \varphi\, dx \ge \mu \int_\Omega \sum_{|\alpha|=m} (D^\alpha \varphi)^2\, dx - \gamma \int_\Omega \varphi^2\, dx,$$

for all $\varphi \in H_0^m(\Omega)$. Thus for the remainder of this lecture we assume (B), (C).

Recall also from Lecture 6 that any solution of $P_\lambda(0)$ is automatically $C^\infty(\Omega)$ provided that $a_{\alpha\beta} \in C^\infty(\Omega)$, and if in addition the domain is smooth then (by the results of Lectures 8,9) u is actually $C^\infty(\overline{\Omega})$ provided $a_{\alpha\beta} \in C^\infty(\overline{\Omega})$, and in this case u satisfies the classical Dirichlet problem

$$Lu = \lambda u \text{ in } \Omega$$
$$D^\alpha u = 0 \text{ on } \partial\Omega,\ |\alpha| \le m-1,$$

where $Lu = \sum_{|\alpha|,|\beta|\le m}(-1)^m D^\beta(a_{\alpha\beta}D^\alpha u) = 0$.

As indicated above, we are here interested in the self-adjoint case, when

$$(*) \qquad A(\varphi,\psi) = A(\psi,\varphi)\ \forall\, \varphi,\psi \in \mathcal{H}.$$

Notice that since we are dealing with Dirichlet boundary conditions here, in case $a_{\alpha\beta} \in C^m(\Omega)$ the self-adjointness condition $*$ is equivalent to the requirement that

$$\textstyle\sum_{|\alpha|,|\beta|\le m} D^\beta(a_{\alpha\beta}D^\alpha) = \sum_{|\alpha|,|\beta|\le m}(-1)^{|\alpha|+|\beta|}D^\alpha(a_{\alpha\beta}D^\beta),$$

which translates into various identities involving the different order coefficients and their derivatives. (See Exercise 7.4 in Lecture 7.) In particular it is sufficient that $a_{\alpha\beta} = a_{\beta\alpha}$ for $|\alpha| = |\beta|$ and $a_{\alpha\beta} = 0$ for $|\alpha| \ne |\beta|$.

The first main result here is summarized in the following theorem:

Theorem 1. *Suppose the assumptions are as above; that is, Ω is bounded, $a_{\alpha\beta} \in L^\infty(\Omega)$, the self-adjointness condition $*$ above holds and the coercivity and boundedness conditions (C), (B_0) above hold. Then the set Λ forms an increasing sequence*

$$\lambda_1 \le \lambda_2 \le \dots,$$

where $\lambda_k \to \infty$ as $k \to \infty$ and where q repetitions of an eigenvalue λ_j indicates that the space of solutions of the problem $P_{\lambda_j}(0)$ has dimension q (we say λ_j has "multiplicity q"). Furthermore, corresponding to the sequence $\lambda_1, \lambda_2, \dots$ there is a sequence $\varphi_1, \varphi_2, \dots \subset \mathcal{H}$ with $A(\varphi_k,\psi) = \lambda_k\langle\varphi_k,\psi\rangle_{L^2(\Omega)}$ and with $\varphi_1, \varphi_2, \dots$ forming a complete orthonormal set in $L^2(\Omega)$ in the sense that $\langle\varphi_i,\varphi_j\rangle_{L^2(\Omega)} = \delta_{ij}$ and

$$\psi = \textstyle\sum_{j=1}^\infty \langle\psi,\varphi_j\rangle_{L^2(\Omega)}\varphi_j,\ \ \psi \in L^2(\Omega).$$

Here the convergence of the series is with respect to the $L^2(\Omega)$ norm. If $\psi \in \mathcal{H}$, then the convergence is with respect to the H^m norm.

Proof: First note that by the coercivity condition (C) there is a real number λ_1 such that

$$\lambda_1 = \inf_{u\in\mathcal{H}\setminus\{0\}} \frac{A(u,u)}{\|u\|_0^2}.$$

We claim that there is a $\varphi_1 \in \mathcal{H}$ with $\|\varphi_1\|_0 = 1$ such that φ_1 causes this inf to be realized; that is,

$$(1) \qquad A(\varphi_1,\varphi_1) = \lambda_1.$$

To prove this claim, we let $\{u_k\}$ be a sequence in $\mathcal{H}$ with $\|u_k\|_0 = 1\ \forall k$ and with $A(u_k,u_k) \to \lambda_1$, so for given $\varepsilon > 0$ we have

$$(2) \qquad A(u_k,u_k) \le \lambda_1 + \varepsilon,$$

for k sufficiently large. In particular by (C) we have that $\|u_k\|_m$ is bounded independent of k, and hence by Rellich's theorem there is a subsequence, still denoted u_k, which converges in $L^2(\Omega)$ to some φ_1 with $\varphi_1 \in \mathcal{H}$:

$$(3) \qquad \|u_k - \varphi_1\|_0 \to 0,\ \ \varphi_1 \in \mathcal{H}.$$

Since A is bilinear, we have (using $*$) the parallelogram identity

(4)
$$A(u_k-u_\ell,u_k-u_\ell)+A(u_k+u_\ell,u_k+u_\ell) = 2(A(u_k,u_k)+A(u_\ell,u_\ell))\ \forall\, k,\ell \ge 1.$$

and, by definition of λ_1,

$$(5) \qquad A(\varphi,\varphi) \ge \lambda_1\|\varphi\|_0^2\ \forall\,\varphi \in \mathcal{H}.$$

Combining (2), (3),(4),(5) we have

$$(6) \qquad \lambda_1\|u_k+u_\ell\|_0^2 + A(u_k-u_\ell,u_k-u_\ell) \le 4\lambda_1 + 4\varepsilon$$

for all sufficiently large $k,\ \ell$. By (3) and the fact that $\|u_k\|_0 = 1\ \forall k$, we see $\|u_k+u_\ell\|_0^2 \to 4$ as $k,\ell \to \infty$, whereupon (6) gives

$$A(u_k-u_\ell,u_k-u_\ell) \le 5\varepsilon$$

for all sufficiently large $k,\ \ell$. Then using again the coercivity (C) and the convergence of u_k with respect to $\|\ \|_0$, we have

$$\|u_k-u_\ell\|_m^2 \le 6\varepsilon$$

for all sufficiently large $k,\ \ell$; that is, $\{u_k\}$ is a Cauchy sequence with respect to $\|\ \|_m$. Thus the convergence to φ_1 is in $\mathcal{H}$, and hence $A(u_k,u_k) \to A(\varphi_1,\varphi_1)$ as $k \to \infty$. This completes the proof of the fact that $A(\varphi_1,\varphi_1) = \lambda_1$.

Next we claim that φ_1 satisfies the problem $\mathrm{P}_{\lambda_1}(0)$. To see this we note that by the definition of λ_1 and the bilinearity of A, we have

$$\lambda_1 \le \frac{A(\varphi_1 + t\psi, \varphi_1 + t\psi)}{\|\varphi_1 + t\psi\|_0^2} = \frac{A(\varphi_1,\varphi_1) + 2tA(\varphi_1,\psi) + t^2 A(\psi,\psi)}{1 + 2t\langle \varphi_1,\psi\rangle_{L^2} + t^2\|\psi\|_0^2}$$

for any $\psi \in \mathcal{H}$ and any $t \in \mathbb{R}$ small enough to ensure that $\|\varphi_1 + t\psi\|_0 \neq 0$ (which is true if for example $|t|\,\|\psi\|_0 < 1/2$). But this guarantees that the expression on the right has a local minimum at $t = 0$, hence has zero derivative at $t = 0$. Checking, one finds that this gives the required equation $A(\varphi_1,\psi) = \lambda_1\langle\varphi_1,\psi\rangle$.

Now by an essentially identical argument, working in the Hilbert space $\mathcal{H}_1 = \{u \in \mathcal{H} : \langle u,\varphi_1\rangle_0 = 0\}$, we can prove that there is $\varphi_2 \perp \varphi_1$ (orthogonality with respect to the $L^2(\Omega)$ inner product) such that $\|\varphi_2\|_0 = 1$ and

$$A(\varphi_2,\varphi_2) = \lambda_2 = \inf_{u\in\{\varphi\in\mathcal{H}\setminus\{0\}:\varphi\perp\varphi_1\}} \frac{A(u,u)}{\|u\|_0^2},$$

and this φ_2 satisfies $A(\varphi_2,\psi) = \lambda_2\langle\varphi_2,\psi\rangle_0$ for every $\psi \in \mathcal{H}_1$, and hence for every $\psi \in \mathcal{H}$ because by construction $\langle\varphi_2,\varphi_1\rangle_0 = 0$.

Continuing inductively we get a sequence $\{\varphi_k\} \subset \mathcal{H}$ with $\varphi_k \perp \varphi_\ell,\ k \neq \ell$ $\|\varphi_k\|_0 = 1$ for $k = 1,2,\ldots$, and

$$A(\varphi_k,\varphi_k) = \lambda_k = \inf_{u\in\{\varphi\in\mathcal{H}\setminus\{0\}:\varphi\perp\varphi_\ell,\,\ell=1,\ldots,k-1\}} \frac{A(u,u)}{\|u\|_0^2},$$

and where φ_k satisfies

$$A(\varphi_k,\psi) = \lambda_k\langle\varphi_k,\psi\rangle_{L^2(\Omega)} \ \ \forall\,\psi \in \mathcal{H}.$$

(Notice that by definition λ_k is an increasing sequence.)

The next step is to show that the sequence λ_k so constructed has the property that $\lambda_k \uparrow \infty$ as $k \uparrow \infty$. This is true, because otherwise by the coercivity condition (C) we would have that $\sup_{k\ge1}\|\varphi_k\|_m < \infty$, which by Rellich's theorem would give us a subsequence converging strongly in $L^2(\Omega)$, thus implying in particular that a subsequence of $\{\varphi_k\}$ is Cauchy with respect to the norm $\|\ \|_0$, contradicting the fact that

$$\|\varphi_k - \varphi_\ell\|_0^2 = \|\varphi_k\|_0^2 + \|\varphi_\ell\|_0^2 - 2\langle\varphi_k,\varphi_\ell\rangle_{L^2(\Omega)} = 2\ \forall\,k \neq \ell.$$

To check the completeness of the orthonormal set $\{\varphi_k\}$ as claimed in the statement of the theorem, first assume that $\psi \in \mathcal{H}$ and note that for each $N \ge 1$

we have $\psi - \sum_{j=1}^N a_j\varphi_j \perp \varphi_k\ \forall\,k = 1,\ldots,N$, where $a_j = \langle\psi,\varphi_j\rangle_{L^2}$. Then by definition of λ_N, we have

$$\begin{aligned}(1)\quad \lambda_{N+1}\|\psi - \textstyle\sum_{j=1}^N a_j\varphi_j\|_0^2 &\le A(\psi - \textstyle\sum_{j=1}^N a_j\varphi_j,\ \psi - \sum_{j=1}^N a_j\varphi_j)\\ &= A(\psi,\psi) - 2\textstyle\sum_{j=1}^N a_jA(\varphi_j,\psi) + \sum_{j,\ell=1}^N a_ja_\ell A(\varphi_j,\varphi_\ell)\\ &= A(\psi,\psi) - \textstyle\sum_{j=1}^N \lambda_j a_j^2 \le A(\psi,\psi) + \sum_{j=1}^{N_0}|\lambda_j|a_j^2,\end{aligned}$$

where $N_0 \ge 1$ is any integer such that $\lambda_j \ge 0$ for $j \ge N_0$. Then since $\lambda_N \to \infty$ as $N \to \infty$, we have the required eigenfunction expansion for ψ. To handle the general case when $\psi \in L^2(\Omega)$, we note that $\mathcal{H}$ is dense in $L^2(\Omega)$ with respect to the L^2 norm (indeed it is a standard fact that $C_c^\infty(\Omega)$ is dense in L^2), and hence the above result implies that for each $\varepsilon > 0$ there exists $N \ge 1$ and constants $c_1,\ldots,c_N \in \mathbb{R}$ such that

$$\|\psi - \textstyle\sum_{j=1}^N c_j\varphi_j\|_0 \le \varepsilon.$$

On the other hand, we have that for any integer $Q \ge 1$

$$\min_{d_1,\ldots,d_Q}\|\psi - \textstyle\sum_{j=1}^Q d_j\varphi_j\|_0^2 = \|\psi - \sum_{j=1}^Q a_j\varphi_j\|_0^2,$$

where $a_j = \langle\varphi_j,\psi\rangle_{L^2}$. (See problem 10.3.) Hence for $Q \ge N$ we conclude that

$$\|\psi - \textstyle\sum_{j=1}^Q a_j\varphi_j\|_0^2 \le \varepsilon;$$

that is, the required eigenfunction expansion holds for ψ. The proof that the convergence is with respect to the H^m-norm when $\psi \in H_0^m(\Omega)$ is left as an exercise.

Finally we have to check that $\{\lambda_1,\lambda_2,\ldots\} = \Lambda$ and that the multiplicities are correct. This is accomplished by virtue of the following important fact; namely, if $\lambda,\mu \in \Lambda$ with corresponding eigenfunctions $\varphi,\psi \in \mathcal{H}$, then $\lambda \neq \mu$ implies that φ,ψ are orthogonal with respect to the L^2 inner product. This is clear because we have

$$\lambda\langle\varphi,\psi\rangle = A(\varphi,\psi) = A(\psi,\varphi) = \mu\langle\psi,\varphi\rangle.$$

Using this fact we obtain the following:

Lemma 1. *If the assumptions are as in the above theorem and if $\lambda_1,\lambda_2,\ldots \subset \Lambda$ is any sequence such that there exists a corresponding sequence $\varphi_1,\varphi_2\ldots$ of eigenfunctions $\in \mathcal{H}$ (i.e. $A(\varphi_j,\psi) = \lambda_j\langle\varphi_j,\psi\rangle_{L^2(\Omega)}\ \forall\,\psi \in \mathcal{H},\ j = 1,2,\ldots$) which is orthonormal and complete in $L^2(\Omega)$, then $\{\lambda_1,\lambda_2,\ldots\} = \Lambda$.*

Proof: Let $\lambda \in \Lambda$. We have to show that $\lambda = \lambda_j$ for some j. Otherwise, by the remark preceding the statement of the lemma we would have a non-zero $\varphi \in \mathcal{H}$ with $A(\varphi, \psi) = \lambda\langle\varphi, \psi\rangle_0 \;\forall\, \psi \in \mathcal{H}$ and $\langle\varphi, \varphi_j\rangle_0 = 0$ for each j. But the $\varphi_1, \varphi_2, \ldots$ are complete, so in particular $\varphi = \sum_{j=1}^{\infty} c_j\varphi_j$, with $c_j = \langle\varphi, \varphi_j\rangle_0 = 0$ for each j, contradicting the fact that $\varphi \neq 0$.

Remark: Notice that if Ω is at least Lipschitz, then all of the above generalizes without change in the proofs to the partially free boundary conditions treated in Lecture 9, simply by replacing the space $\mathcal{H} = H_0^m(\Omega)$ by any of the spaces $\mathcal{H}_k$ as in Lecture 9.

We now want to discuss the min-max principle for the eigenvalues λ_j, and give some applications. The min-max principle, due to Poincaré, is given in the following lemma:

Lemma. *If $\lambda_1, \lambda_2, \ldots$ are the eigenvalues introduced above, then we have*

$$\lambda_j = \min_{S \in \mathcal{S}_j} \Big\{ \max_{S\setminus\{0\}} \frac{A(u,u)}{\|u\|_0^2} \Big\},$$

where $\mathcal{S}_j$ denotes the set of all j-dimensional subspaces of $\mathcal{H}$.

Proof: Let

$$\mu_j = \inf_{S \in \mathcal{S}_j} \Big\{ \max_{u \in S\setminus\{0\}} \frac{A(u,u)}{\|u\|_0^2} \Big\}.$$

Then we want to show that $\mu_j = \lambda_j$ and that the inf is attained. We have trivially

$$\lambda_j \geq \mu_j,$$

because one possible $S \in \mathcal{S}_j$ is the subspace spanned by the eigenfunctions $\varphi_1, \ldots, \varphi_j$ introduced above.

To prove the reverse inequality, let $v_1, \ldots, v_j$ be any linearly independent elements of $\mathcal{H}$, and choose real $c_1, \ldots, c_j$ not all zero such that $u = \sum_{i=1}^{j} c_i v_i$ is orthogonal to $\varphi_1, \ldots, \varphi_{j-1}$ and $\|u\|_0 = 1$. (Notice that we can do this by solving the $j-1$ equations $\sum_{i=1}^{j} c_i \langle v_i, \varphi_k\rangle = 0$, $k = 1, \ldots, j-1$ in the j unknowns $c_1, \ldots, c_j$.) Hence

$$A(u,u) \geq \lambda_j$$

by the Rayleigh quotient definition of λ_j given above. Thus we have $\lambda_j \leq \mu_j$, and hence $\lambda_j = \mu_j$ as required. Also, we can attain the value λ_j by taking S to be the subspace spanned by $\varphi_1, \ldots, \varphi_j$ and $u = \varphi_j$.

This completes the proof.

There is also a max-min principle which is a kind of dual of the above min-max principle, as follows.

Lemma.

$$\lambda_j = \max_{\ell_1,\ldots,\ell_{j-1}\in\mathcal{L}} \Big\{ \min_{\ell_1(u)=0,\ldots,\ell_{j-1}(u)=0,\, u\neq 0} \frac{A(u,u)}{\|u\|_0^2} \Big\},$$

where $\mathcal{L}$ denotes the set of continuous linear functionals on $\mathcal{H}$; in case $j = 1$, we interpret the right side to be $\min_{u\in\mathcal{H}\setminus\{0\}} \frac{A(u,u)}{\|u\|_0^2}$.

Proof: Let

$$\gamma_j = \sup_{\ell_1,\ldots,\ell_{j-1}\in\mathcal{L}} \Big\{ \inf_{\ell_1(u)=0,\ldots,\ell_{j-1}(u)=0,\, u\neq 0} \frac{A(u,u)}{\|u\|_0^2} \Big\},$$

so we have to prove that $\gamma_j = \lambda_j$ and that the sup and inf are attained.

Evidently this is so with $j = 1$. Also for any $j \geq 2$, we have

$$\gamma_j \geq \lambda_j,$$

by taking $\ell_i = \langle\varphi_i, \cdot\rangle_{L^2}$ for $i = 1, \ldots, j-1$ and $u = \varphi_j$.

To prove the reverse inequality, we let $\ell_1, \ldots, \ell_{j-1}$ be any elements of $\mathcal{L}$, and choose $u \in \mathcal{H}$ with $u = \sum_{i=1}^{j} c_i\varphi_i$, and such that $\ell_i(u) = 0$, $i = 1, \ldots, j-1$, and $\|u\|_0 = 1$. (Evidently such constants $c_1, \ldots, c_{j-1}$ exist, because the c_i are required to satisfy a sytem of $j-1$ linear equations and there are j unknowns c_i.) Then the quotient $A(u,u)/\|u\|_0^2$ equals $\sum_{i=1}^{j} c_i^2\lambda_i \leq \lambda_j$, so we have shown $\gamma_j \leq \lambda_j$ as required.

Further the sup and inf are attained, by the choice of ℓ_i and u described in the proof of the inequality $\gamma_j \geq \lambda_j$.

As a first application of the min-max principle, we have the following monotonicity lemma:

Lemma. *If $\Omega_1 \subset \Omega_2$ (bounded), and if λ_j^1, λ_j^2 denote the j-th eigenvalues for the problem $P_\lambda(0)$ on $H_0^m(\Omega_1)$, $H_0^m(\Omega_2)$ respectively. Then*

$$\lambda_j^1 \geq \lambda_j^2, \;\; j \geq 1.$$

Proof: Notice first that

$$H_0^m(\Omega_1) \subset H_0^m(\Omega_2) \tag{1}$$

in the sense that if $u \in H_0^m(\Omega_1)$ and if we extend u to $\widetilde{u}$ on Ω_2 by defining $\widetilde{u} = 0$ on $\Omega_2 \setminus \Omega_1$, then $\widetilde{u} \in H_0^m(\Omega_2)$. (This follows trivially from the definition of H_0^m.)

Let $\mathcal{S}_j^i$ be the set of all j-dimensional subspaces of $H_0^m(\Omega_i)$, $i = 1, 2$. In view of (1) we have

$$\mathcal{S}_j^1 \subset \mathcal{S}_j^2$$

in the obvious sense (extending functions in $\mathcal{S}_j^1$ to be zero on $\Omega_2 \setminus \Omega_1$). Hence the required inequality follows from the min-max principle.

Another very important application of the max-min principle is the following estimate for λ_j, which gives upper and lower bounds for the asymptotic behaviour of λ_j as $j \to \infty$.

Theorem 2. *There is a fixed constant $C > 1$, $C = C(n, m, L, \Omega)$, such that*

$$C^{-1} j^{2m/n} \le \lambda_j \le C j^{2m/n}$$

for all j such that $\lambda_j \ge 1$.

Before we begin the proof, we note the following corollary, whose proof is left as an exercise based on Theorems 1, 2 (see problem 10.2 below):

Corollary 1. *Let $\lambda_1, \lambda_2, \ldots$ be the eigenvalues introduced above, and $\varphi_1, \varphi_2, \ldots$ the corresponding complete orthonormal set of eigenfunctions. Then*

$$H_0^m(\Omega) = \{\psi \in L^2(\Omega) : \textstyle\sum_{j=1}^\infty j^{2m/n} a_j^2 < \infty\}$$

where $a_j = \langle \psi, \varphi_j \rangle_{L^2(\Omega)}$. Also $\| \ \|_{m,\Omega}$ and $(\sum_{j=1}^\infty j^{2m/n} a_j^2)^{1/2}$ are equivalent norms.

Proof of Theorem 2: By the coercivity condition (C) and the interpolation inequality (Lemma 6 of Lecture 5) we have

$$\text{(1)} \qquad \mu A^0(u,u) - C\|u\|_0^2 \le A(u,u) \le C(A^0(u,u) + \|u\|_0^2) \ \forall\, u \in \mathcal{H}$$

for suitable constants μ, $C > 0$, where A^0 is the bilinear form associated with the operator $(-1)^m \Delta^m$. That is, A^0 is given by

$$A^0(u,v) = \int_\Omega \textstyle\sum_{j_1,\ldots,j_m=1}^n (D_{j_1}\cdots D_{j_m} u)(D_{j_1}\cdots D_{j_m} v), \ u, v \in \mathcal{H}.$$

For any domain U, let A_U^0 be as above with U in place of Ω, and let $\lambda_j(U)$ be the j^{th} eigenvalue of A^0 on $H_0^m(U)$. Then by the monotonicity result proved

above, we have

$$\lambda_j(C_2) \le \lambda_j(\Omega) \le \lambda_j(C_1)$$

for any cubes C_1, C_2 with $C_1 \subset \Omega \subset C_2$. Also, in view of (1) (with C_1, C_2 in place of Ω), we can apply the min/max principle to give

$$C^{-1}\lambda_j^0(C_i) \le \lambda_j(C_i) \le C\lambda_j^0(C_i)$$

for sufficiently large j, so to complete the proof it is evidently enough to show there is a constant $c = c(n,m) \ge 1$ such that

$$\text{(2)} \qquad c^{-1} j^{2m/n} \le \lambda_j(U) \le c j^{2m/n}, \ j \ge 1,$$

where $U = (0,\pi) \times \cdots \times (0,\pi)$.

We first check this in case $m = 1$. Notice that the classical version of the eigenvalue problem $\mathrm{P}_\lambda(0)$ in the case $m = 1$ on U is

$$\begin{cases} -\Delta u = \lambda u \text{ on } U \\ \quad u = 0 \text{ on } \partial U. \end{cases}$$

We get classical solutions of this by separating variables (i.e. looking for solutions of the special form $u(x^1, \ldots, x^n) = u_1(x^1) \cdot u(x^2) \cdots u_n(x^n)$). We in fact see that the products $\sin(k_1 x^1) \cdots \sin(k_n x^n)$ are such solutions for any positive integers $k_1, \ldots, k_n$—the corresponding value of λ is $\lambda = k_1^2 + \cdots + k_n^2$. As a matter of fact

$$\{\sin(k_1 x^1) \cdots \sin(k_n x^n) : (k_1, \ldots, k_n) \in \mathbb{Z}_+^n\}, \ \mathbb{Z}_+ = \{1, 2, \ldots\},$$

is a complete orthonormal set for $L^2(U)$ (see Exercise 10.6), hence by Lemma 1 the required sequence of eigenvalues $\lambda_1 \le \lambda_2 \le \ldots$ (with correct multiplicity) is given by ordering the set

$$\{k_1^2 + \cdots + k_n^2 : k_1, \ldots, k_n \in \mathbb{Z}_+\}$$

and taking the multiplicity of $k_1^2 + \cdots + k_n^2$ to be the number of elements in the set $\{(j_1, \ldots, j_n) \in Z_+ : j_1^2 + \cdots + j_n^2 = k_1^2 + \cdots + k_n^2\}$. Hence for any given j we get

$$\lambda_j = k_1^2 + \cdots + k_n^2 \text{ for some } (k_1, \ldots, k_n) \in \mathbb{Z}_+$$

and $N \le j \le \overline{N}$, where

$$N = \text{ number of elements in } \{(j_1, \ldots, j_n) \in Z_+ : j_1^2 + \cdots + j_n^2 < k_1^2 + \cdots + k_n^2\}$$

and $\overline{N}$ is the same quantity when we replace $<$ by $\leq$. Since, for any given real λ, the number N_λ of elements in the set $\{(j_1,\ldots,j_n) \in Z_+ : j_1^2+\cdots+j_n^2 \leq \lambda\}$ satisfies

$$c^{-1}\lambda^{n/2} \leq N_\lambda \leq c\lambda^{n/2}$$

for some fixed constant $c = c(n)$ (see Exercise 10.7), we have

$$c^{-1}\lambda_j^{n/2} \leq j \leq c\lambda_j^{n/2},$$

or in other words

$$c^{-1}j^{2/n} \leq \lambda_j \leq cj^{2/n}$$

as required.

Next to handle the case $m \geq 2$, we note first the inclusion

$$(*) \qquad H_0^m(U) \subset H^m(U) \cap H_0^1(U),$$

so that by the min-max principle we get

$$\lambda_j \geq \widetilde{\lambda}_j(U),$$

where $\widetilde{\lambda}_j$ is the j^{th} eigenvalue for A_0 with the space $\mathcal{H}_1$ in place of $H_0^m(\Omega)$ (see the remark after Lemma 1 above). Since the eigenfunctions from the case $m = 1$ are all smooth on $\overline{U}$, indeed they are of the form $\sin(k_1x^1)\cdots\sin(k_nx^n)$, hence in $\mathcal{H}_1 \equiv H^m(U) \cap H_0^1(U)$, it is easy to check (with the aid of Lemma 1) that

$$\widetilde{\lambda}_j(U) \geq \mu_j^m,$$

where μ_j is the j^{th} eigenvalue of Δ on U. See Exercise 10.8 below. That is, since we showed $\mu_j \geq c^{-1}j^{2/n}$ above, we conclude

$$\lambda_j \geq c^{-1}j^{2m/n}.$$

Finally we note the inclusion

$$\{\sin^m(k_1x^1)\cdots\sin^m(k_nx^n) : k_1,\ldots,k_n \in \mathbb{Z}_+\} \subset H_0^m(U),$$

hence, by a counting argument similar to that used in the case $m = 1$ (see Exercise 10.9), we conclude also that $\lambda_j \leq cj^{2m/n}$.

LECTURE 10 PROBLEMS

10.1 Check the claim made in the proof of Theorem 1 that the sequence $\{u_k\}$ is bounded independent of k.

10.2 Check that the claim made in Theorem 1 that if $\psi \in \mathcal{H}$ then the eigenfunction expansion of ψ converges in $\mathcal{H}$ with respect to the norm $\|\ \|_{m,\Omega}$, and hence (using also Theorem 2) give the proof of Corollary 1.

Hint: To prove the convergence, use the coercivity condition (C).

10.3 Prove that if H is any real Hilbert space, and if $\varphi_1,\ldots,\varphi_Q$ are any orthonormal set in H, then

$$\min_{\mu_1,\ldots,\mu_Q\in\mathbb{R}} \|\psi - \textstyle\sum_{j=1}^Q \mu_j\varphi_j\|_H^2 = \|\psi - \sum_{j=1}^Q a_j\varphi_j\|_H^2, \quad a_j = (\psi,\varphi_j)_H.$$

10.4 Compute the spectrum (with respect to Dirichlet boundary data) for the operator $-\Delta = -\frac{d^2}{dx^2}$ on $\Omega = (0,1) \subset \mathbb{R}$, and discuss the consequences of the main theorem above in this special case.

10.5 Find the spectrum for the problem

$$\Delta^2 u = \lambda u \text{ on } \Omega = (0,1)\times(0,1) \subset \mathbb{R}^2$$
$$u = 0 \text{ and } \Delta u = 0 \text{ on } \partial\Omega.$$

Of course you should use the weak formulation of this problem—possible because Δu is a natural boundary operator for

$$\Delta^2 = \textstyle\sum_{k,\ell,i,j} D_k D_\ell(\delta_{k\ell}\delta_{ij}D_iD_j).$$

10.6 Check as claimed (using the completeness of $\sin(kt)$, $k = 1,2,\ldots$ in $L^2(0,\pi)$) that $\{\sin(k_1x^1)\cdots\sin(k_nx^n)\}$ is a complete orthonormal set in $L^2(U)$.

10.7 Prove the inequality (claimed above)

$$c^{-1}\lambda^{n/2} \leq N_\lambda \leq c\lambda^{n/2}.$$

Note: Geometrically, N_λ is just the number of points of the integer lattice $\mathbb{Z}_+^n$ which lie in the closure of the ball $B_{\sqrt{\lambda}}(0)$ of radius $\sqrt{\lambda}$ and centre 0.

10.8 Check the claim made at the end of the above lecture that $\widetilde{\lambda}_j(U) \geq \mu_j$ by using the inclusion $*$ above and the min-max principle.

10.9 Check the claim that $\sin^m(k_1x^1)\cdots\sin^m(k_nx^n) \in H_0^m(U)$ and that hence $\lambda_j \leq cj^{2m/n}$ as claimed above.

Lecture 11

The Initial-boundary-value Problem for Parabolic Equations, the Heat Kernel & Weyl's Asymptotic Formula

As we mentioned earlier in Lecture 1, heat flow (or more specifically the temperature at any given time t and any given point x) in a 3-dimensional anisotropic inhomogeneous medium is modelled by an equation of the form

$$u_t - \gamma \sum_{i,j=1}^{3} D_i(a_{ij} D_j u) = 0 \text{ on } \Omega \times (0, \infty),$$

where $\gamma > 0$ and a_{ij} is a given positive definite matrix (γ, a_{ij} are determined by the material in which the temperature distribution exists).

In case of a homogeneous isotropic medium, we have $\gamma =$ const. and $a_{ij} = \delta_{ij}$, so that after a change of scale $t \mapsto \gamma t$ the equation becomes

$$u_t - \Delta u = 0.$$

This equation is called the heat equation. It is the prototypical example of a (linear) parabolic equation—by which we mean an equation of the form

$$(*) \qquad u_t + Lu = f \text{ on } \Omega \times (0, \infty)$$

where f is given on $\Omega \times (0, \infty)$ and where L is an elliptic operator on Ω; we here consider L of the form considered in Lecture 10, with the same self-adjointness property; i.e.

$$Lu = \sum_{|\alpha|,|\beta|\le m}(-1)^m D^\beta(a_{\alpha\beta}D^\alpha u) \equiv \sum_{|\alpha|,|\beta|\le m}(-1)^{m+|\alpha|+|\beta|}D^\beta(a_{\beta\alpha}D^\alpha u).$$

We assume that $a_{\alpha\beta}$ are bounded on Ω and that a coercivity condition (as in Lecture 10) holds as follows:

(C)
$$\int_\Omega \sum_{|\alpha|=|\beta|=m} a_{\alpha\beta}D^\alpha\varphi D^\beta\varphi \ge \mu\int_\Omega \sum_{|\alpha|=m}|D^\alpha\varphi|^2 - C\|\varphi\|^2_{0,\Omega}, \quad \varphi \in H_0^m(\Omega).$$

Recall (Exercise 6.5(ii)) that this implies the weaker ellipticity condition

$$\text{(E)}' \qquad \sum_{|\alpha|,|\beta|=m} a_{\alpha\beta}\xi^\alpha\xi^\beta \ge \mu|\xi|^{2m} \;\forall\, \xi \in \mathbb{R}^n,$$

where $\mu > 0$ is constant, and conversely (see Gårding's inequality in Lecture 7) (C) is implied by (E)′ if the coefficients $a_{\alpha\beta}$ with $|\alpha| = |\beta| = m$ are continuous on $\overline{\Omega}$. In any case it is implied by the stronger ellipticity condition (E) of Lecture 6. We also assume that L is self-adjoint in the sense of the previous lecture.

Notice that we allow a given function f on the right in $(*)$; this corresponds to allowing *heat sources* and *sinks* in the physical situation.

In addition to their applications in physics, parabolic equations are important in non-linear problems in Riemannian geometry.

We begin with some general results about the initial-boundary-value problem for the heat equation. For the present introductory discussion we take $f \equiv 0$ in $(*)$. (See Exercise 11.1 for some discussion relating to the case of non-zero f.)

Consider the initial-boundary-value problem:

$$(**) \qquad \begin{cases} u_t + Lu = 0 \text{ on } \Omega \times (0,\infty) \\ D^\alpha u = 0 \text{ on } \partial\Omega \times (0,\infty), \; |\alpha| \le m-1 \\ u = \varphi \text{ on } \Omega \times \{0\}. \end{cases}$$

Notice that the boundary conditions here are precisely Dirichlet boundary conditions; with no essential change in the discussion, we could equally well consider any of the partially free boundary conditions of Lecture 9.

Since we don't necessarily want to assume smoothness of the domain Ω and the functions $a_{\alpha\beta}$, φ, we first need to give a weak formulation of the problem $(*)$. To do this we take $\mathcal{H}$ to denote the set of functions w on $\Omega \times (0,\infty)$ such that $w|\Omega \times (0,T) \in L^2(\Omega \times (0,T))$, $\forall T > 0$, and such that $w|\Omega \times \{t\} \in H_0^m(\Omega)$ a.e. $t \in (0,\infty)$ and $D_x^\alpha w(x,t) \in L^2(\Omega \times (0,T))$ for each $T > 0$ and each $|\alpha| \le m$. Then a suitable weak formulation of $(**)$ is the following:

Find a $u \in \mathcal{H}$ such that

$(**)'$
$$-\langle u, \psi_t\rangle_{L^2(\Omega\times(0,\infty))} + \overline{A}(u,\psi) = 0, \;\forall\psi \in C_c^\infty(\Omega\times(0,\infty)) \text{ and } \lim_{t\downarrow 0} u(\cdot,t) = \varphi,$$

where

$$\overline{A}(\varphi,\psi) = \int_0^\infty \int_\Omega \sum_{|\alpha|,|\beta|\le m}(-1)^{m+|\beta|}a_{\alpha\beta}(x)D_x^\alpha\varphi D_x^\beta\psi\, dxdt,$$

and where the limit on the right is required to hold in the sense that

$$\lim_{t\downarrow 0}\langle u(\cdot,t),\psi\rangle_{L^2(\Omega)} = \langle\varphi,\psi\rangle_{L^2(\Omega)} \;\forall\,\psi \in H_0^m(\Omega).$$

We now state a general existence and uniqueness result for the problem $(**)'$.

Theorem 1. *Given $\varphi \in L^2(\Omega)$ and L as described above (with $a_{\alpha\beta} \in L^\infty(\Omega)$ and with the coercivity condition (C) and self-adjointness condition $*$ of Lecture 10 holding), there is a unique $u \in \mathcal{H}$ satisfying $(**)'$. Furthermore, if $a_{\alpha\beta} \in C^\infty(\Omega)$, then $u \in C^\infty(\Omega \times (0,\infty))$ and if Ω is a C^∞ domain and $a_{\alpha\beta} \in C^\infty(\overline{\Omega})$, then $u \in C^\infty(\overline{\Omega} \times (0,\infty))$. Also, if $\varphi \in C_c^\infty(\Omega)$, $\Omega \in C^\infty$, and $a_{\alpha\beta} \in C^\infty(\overline{\Omega})$, then $u \in C^\infty(\overline{\Omega} \times [0,\infty))$ (and it satisfies the classical problem $(**)$).*

Proof: We look for separated-variable solutions $a(t)b(x)$. By direct calculation we see that $e^{-\lambda_j t}\varphi_j(x)$ (for any fixed $j \ge 1$) is such a solution; here λ_j denote the eigenvalues of L as discussed in the previous lecture and φ_j are the corresponding complete orthonormal set of eigenfunctions.

We now look for a solution to our initial-boundary-value problem as a superposition of these—that is, we look for u in the form

$$u(x,t) = \sum_{j=1}^\infty c_j e^{-\lambda_j t}\varphi_j(x).$$

In order to ensure that the initial condition holds, we evidently want to select

$$c_j = \langle\varphi,\varphi_j\rangle_{L^2(\Omega)},$$

because we know that $\varphi = \sum_{j=1}^\infty\langle\varphi,\varphi_j\rangle_{L^2(\Omega)}\varphi_j$, by virtue of the fact that the φ_j are a complete orthonormal set.

In view of the facts that $\lambda_j \ge C^{-1}j^{2m/n}$ for sufficiently large j (by Lecture 10) and $\sum_{j=1}^\infty c_j^2 < \infty$, it is easy to check that the series for $u(x,t)$ converges absolutely at each point of $\Omega \times (0,\infty)$, and converges in $L^2(\Omega \times (0,T))$ for each

$T > 0$. Since $\|\varphi_j\|_{m,\Omega} \le C\lambda_j$ for j sufficiently large (by the Rayleigh quotient construction of the φ_j), it is also easy to check that the series for $u(x,t)$ converges with respect to the Sobolev norm $\| \ \|_{m,\Omega}$ for each fixed $t > 0$, and this convergence holds uniformly for $t \ge \varepsilon$ for any fixed $\varepsilon > 0$. Thus we have established that $u \in \mathcal{H}$. To check that the initial condition holds in the appropriate sense, we have only to use the fact that if $b_j = \langle \varphi_j, \psi\rangle_{L^2(\Omega)}$, then by the orthonormality of the φ_j, the series $\sum_{j=1}^\infty b_j c_j$ is absolutely convergent, and hence

$$\lim_{t\downarrow 0}\langle u(t),\psi\rangle_{L^2(\Omega)} \equiv \lim_{t\downarrow 0}\sum_{j=1}^\infty e^{-\lambda_j t} b_j c_j = \sum_{j=1}^\infty b_j c_j \equiv \langle \varphi,\psi\rangle_{L^2(\Omega)},$$

as required. Further, by the convergence in $L^2(\Omega\times(0,T))$ and in the norm $\| \ \|_{m,\Omega}$ for fixed t, together with the fact that each term of the series defining u satisfies the identity $\langle u,\psi_t\rangle_{L^2(\Omega\times(0,\infty))} = \overline{A}(u,\psi)\ \forall\,\psi\in C_c^\infty(\Omega\times(0,\infty))$, we easily check that u itself satisfies the same identity. Thus we have the required solution of $(**)'$.

We leave the uniqueness of u as an exercise. (See Exercise 11.3.)

Thus it remains to check the regularity statements. This will be done using the series representation for u directly, but first we need a lemma concerning L^2 estimates for the eigenfunctions φ_j.

Lemma 1. *If φ_j are a complete orthonormal set of eigenfunctions for the operator L, where we assume the coercivity condition (C) and that the coefficients $a_{\alpha\beta}$ are bounded and of class $C^\infty(\Omega)$, then for any ball $B_\rho(y)\subset\Omega$ and any $\theta\in(0,1)$*

$$(\ddagger)\qquad \|\varphi_j\|_{qm,B_{\theta\rho}(y)} \le C\,j^{2qm/n},\quad q\ge 1,\ j\ge 1,$$

where $C = C(\theta,\rho,q,a_{\alpha\beta})$; C does not depend on j. If $\Omega\in C^\infty$ and if $a_{\alpha\beta}\in C^\infty(\overline{\Omega})$ then the same holds provided the hypothesis $B_\rho(y)\subset\Omega$ is replaced by the hypothesis $y\in\partial\Omega$, and provided $\|\varphi_j\|_{qm,B_{\theta\rho}(y)}$ is replaced by $\|\varphi_j\|_{qm,B_{\theta\rho}(y)\cap\Omega}$ in the conclusion; in this case the constant C depends also on Ω.

Remark 1: Since for any integer $\ell\ge 0$ we can choose an integer $q\ge 1$ such that $n/2+\ell < qm \le n/2+\ell+m$, by virtue of the Sobolev Embedding Theorem (Lecture 5) we conclude from $(\ddagger)$ that

$$(\ddagger\ddagger)\qquad |\varphi_j|_{C^\ell(B_{\theta\rho}(y))} \le C\,j^{1+2(\ell+m)/n},\quad j\ge 1,\ell\ge 0,$$

where C depends only on $\ell,\rho,\theta,a_{\alpha\beta}$; again C does not depend on j. If $\Omega\in C^\infty$ and $a_{\alpha\beta}\in C^\infty(\overline{\Omega})$ then again this is valid provided the hypothesis $B_\rho(y)\subset\Omega$ is replaced by the hypothesis $y\in\partial\Omega$ and provided $|\varphi_j|_{C^\ell(B_{\theta\rho}(y))}$ is replaced by $|\varphi_j|_{C^\ell(\Omega\cap B_{\theta\rho}(y))}$ in the conclusion.

Proof of Lemma 1: First note that since $L\varphi_j = \lambda_j\varphi_j$, we have by the regularity theory of Lecture 6 (with $f\equiv\lambda_j\varphi_j$ in place of $\sum_{|\beta|\le m}D^\beta f_\beta$) that for any ball $B_\rho(y)\subset\Omega$, for any $k\ge 0$, and for any $\theta\in(0,1)$ we have $\|\varphi_j\|_{m+k,B_{\theta\rho}(y)} \le C(|\lambda_j|\cdot\|\varphi_j\|_{k,B_\rho(y)} + \|\varphi_j\|_{0,B_\rho(y)})$, and by choosing $k = (q-1)m$ we have $\|\varphi_j\|_{qm,B_{\theta\rho}(y)} \le C(|\lambda_j|\cdot\|\varphi_j\|_{(q-1)m,B_\rho(y)} + \|\varphi_j\|_{0,B_\rho(y)})$ for each $q\ge 1$, with $C = C(q,n,m,\rho,\theta,a_{\alpha\beta})$. By induction on q, this gives $\|\varphi_j\|_{qm,B_{\theta\rho}(y)} \le C(1+|\lambda_j|)^{q-1}\|\varphi_j\|_{m,B_\rho(y)}$. On the other hand by construction of the φ_j described in Lecture 10 we have $\|\varphi_j\|_{m,\Omega}\le C(1+|\lambda_j|)$, where C, $\widetilde{C}$ depend on $a_{\alpha\beta}$ but not on j, and hence $\|\varphi_j\|_{qm,B_{\theta\rho}(y)} \le C(1+|\lambda_j|)^q$ for some constant $C = C(\rho,\theta,q,a_{\alpha\beta})$. Since $\lambda_j\le C\,j^{2m/n}$ (see Lecture 10), we then have $(\ddagger)$ as required.

To obtain the stated variants of $(\ddagger)$ in case $\Omega\in C^\infty$ and $a_{\alpha\beta}\in C^\infty(\overline{\Omega})\cap L^\infty(\Omega)$, we simply use the boundary L^2 estimates of Lecture 6 in place of the above interior estimates.

This completes the proof of the lemma.

We can now check the regularity statements of Theorem 1. In case $a_{\alpha\beta}\in C^\infty(\Omega)$, the estimate $(\ddagger\ddagger)$ of Remark 1 above can be applied. Since $\lambda_j\ge C^{-1}j^{2m/n}$, the presence of the negative exponentials in $\sum_{j=1}^\infty c_j e^{-\lambda_j t}\varphi_j$ then enables us to justify termwise differentiation as many times as we please for $t>0$, thus showing that $u\in C^\infty(\Omega\times(0,\infty))$ as required. The remaining regularity statements are established similarly, using the variant of $(\ddagger\ddagger)$ with $\Omega\cap B_{\theta\rho}(y)$ in place of $B_{\theta\rho}(y)$ as discussed in Remark 1. The final regularity statement uses also the fact that $|c_j|\le C\,j^{-\ell}$ for any $\ell\ge 1$ in case $\varphi\in C_c^\infty(\Omega)$ (where C depends on ℓ but not on j). (See Exercise 11.2.)

Notice that we can modify the above existence program to work with a Borel measure μ of compact support in place of the initial value φ. Indeed we take the series $\sum_{j=1}^\infty c_j e^{-\lambda_j t}\varphi_j(x)$ as before, but now $c_j = \int_\Omega \varphi_j\,d\mu$. An important special case of this occurs when we choose a point $y\in\Omega$ (fixed) and $\mu=\delta_y$. (δ_y = the Dirac delta measure concentrated at y, given by $\delta_y(\varphi)=\varphi(y)$ for any $\varphi\in C^0(\overline{\Omega})$.) In this case the solution u is given by

$$p(x,y,t)\equiv\sum_{j=1}^\infty e^{-\lambda_j t}\varphi_j(y)\varphi_j(x),\ x \text{ in } \Omega,\ t>0.$$

(Notice the symmetry in the variables x,y.)

The function $p(x,y,t)$ so defined is called the heat kernel for the operator L on the domain Ω, for the following reason:

Let $\varphi \in C_c^\infty(\Omega)$ be arbitrary and define

$$u(x,t) = \int_\Omega p(x,y,t)\varphi(y)\,dy.$$

(Notice that p plays the role of a "kernel" here.) We claim that u solves problem $(**)'$ (and the classical problem $(**)$ in case $a_{\alpha\beta} \in C^\infty(\overline{\Omega})$ and $\Omega \in C^\infty$). In fact, by definition, for $t > 0$ we have

$$u(x,t) = \textstyle\sum_{j=1}^\infty b_j e^{-\lambda_j t}\varphi_j(x),$$

where $b_j = \langle \varphi_j, \varphi\rangle_{L^2(\Omega)}$. We showed above that this indeed is the solution of $(**)'$ (and a solution of the classical problem $(**)$ in case $a_{\alpha\beta} \in C^\infty(\overline{\Omega})$ and $\Omega \in C^\infty$).

Notice of course that the heat kernel is unique, because we showed above that the solution of $(**)'$ is unique, hence if $p_1(x,y,t)$ also had the property that $\int_\Omega p_1(x,y,t)\varphi(y)\,dy$ solves $(**)'$ with respect to any given $\varphi \in C_c^\infty(\Omega)$, then we would have to have the identity $\int_\Omega (p_1(x,y,t) - p(x,y,t))\varphi(y)\,dy = 0$ $\forall \varphi \in C_c^\infty(\Omega)$, thus implying that $p_1 = p$, as claimed.

In case $a_{\alpha\beta} \in C^\infty(\Omega)$, the above discussion gives us a nice way of of approximating the heat kernel $p(x,y,t)$ by a sequence $\{q_i\}$ of $C^\infty(\Omega \times [0,\infty))$ solutions of $(**)'$ (or $C^\infty(\overline{\Omega} \times [0,\infty))$ solutions of $(**)$ in case Ω is C^∞ and $a_{\alpha\beta} \in C^\infty(\overline{\Omega})$). Namely, we take any sequence $\psi_i \in C_c^\infty(\Omega)$ with $\psi_i \to \delta_y$ in the sense of measures ($y \in \Omega$ fixed); thus we have $(\psi_i, \psi)_{L^2(\Omega)} \to \psi(y)$ for each $\psi \in C_c^\infty(\Omega)$. Let us also require that support $\psi_i \subset \{x \in \Omega : |x-y| < 1/i\}$ and $\psi_i \geq 0$ everywhere. (See Exercise 11.4.) Then define

$$q_i(x,t) = \int_\Omega p(x,y,t)\psi_i(y)\,dy.$$

By the discussion above we have

$$q_i(x,t) = \textstyle\sum_{j=1}^\infty c_j^i e^{-\lambda_j t}\varphi_j(x),$$

and q_i is a $C^\infty(\Omega \times [0,\infty))$ solution of $(**)'$, where

$$c_j^i \equiv \langle \psi_i, \varphi_j\rangle_{L^2(\Omega)} \to \varphi_j(y).$$

Now let K be an arbitrary compact subset of Ω, and restrict y to be in K. By virtue of the bound on the L^1 norms of ψ_i together with the bound $\sup_K |\varphi_j| \leq C\, j^{2km/n}$ (by (‡‡) above), we have

$$|c_j^i| \leq C\, j^{2km/n}$$

for all sufficiently large j, with C depending only on K and not on i or j. We thus have

$$q_i(x,t) \to p(x,y,t)$$

smoothly on compact subsets of $\Omega \times (0,\infty)$.

Actually in case Ω is not smooth we could also approximate Ω by an increasing sequence Ω_i of C^∞ domains, with $\cup_{i=1}^\infty \Omega_i = \Omega$. Then we take ψ_i as above, and define

$$q_i(x,t) = \int_{\Omega_i} p_i(x,y,t)\psi_i(y)\,dy,$$

where p_i is the heat kernel for Ω_i. Again (by essentially the same argument) we get convergence of q_i to $p(x,y,t)$, but now we have the nice feature that q_i is the $C^\infty(\overline{\Omega}_i \times [0,\infty))$ solution of $(**)$ with Ω_i in place of Ω and with ψ_i in place of φ.

We now want to give Weyl's asymptotic formula for the eigenvalues λ_j of $-\Delta$; recall that in the previous lecture we already established that $C^{-1} j^{2/n} \leq \lambda_j \leq C\, j^{2/n}$ for some fixed constant $C > 0$. Here we want to prove a much more precise result about the asymptotics of λ_j as $j \uparrow \infty$. Specifically, we shall prove:

Theorem 2 (Weyl.) *For $\lambda \in \mathbb{R}$, let N_λ denote the number of eigenvalues λ_j of $-\Delta$ relative to Dirichlet boundary conditions which are $\leq \lambda$ (i.e. the number of eigenvalues $\lambda_j \leq \lambda$ of the problem $P_\lambda(0)$ on $H_0^1(\Omega)$ in case $L = -\Delta$), then*

$$N_\lambda \sim \frac{\lambda^{n/2}|\Omega|}{(4\pi)^{n/2}\Gamma(n/2+1)} \quad \text{as } \lambda \to \infty,$$

where we use the notation $a_\lambda \sim b_\lambda$ as $\lambda \to \infty$ to mean $a_\lambda/b_\lambda \to 1$ as $\lambda \to \infty$, and where $|\Omega|$ means the volume of Ω.

Remark: Note that this says exactly that

$$\lambda_j \sim \frac{4\pi\Gamma(n/2+1)^{2/n} j^{2/n}}{|\Omega|^{2/n}} \quad \text{as } j \to \infty.$$

(Recall that we are using the notational convention that if an eigenvalue λ has multiplicity ℓ, then there is a string of ℓ eigenvalues λ_j each equal to λ in the sequence $\lambda_1 \leq \lambda_2 \leq \cdots$)

The main idea in the proof is to get asymptotic estimates for the behaviour of the heat kernel $p(x,y,t)$ as $t \downarrow 0$. We are particularly interested in asymptotic behaviour of $p(y,y,t)(\equiv \sum_{j=1}^\infty e^{-\lambda_j t}\varphi_j^2(y))$, because we will then have asymptotic behaviour of $\int_\Omega p(y,y,t)\,dy \equiv \sum_{j=1}^\infty e^{-\lambda_j t}$ as $t \downarrow 0$, and this (via a Tauberian theorem) gives information about asymptotic behaviour of λ_j as $j \uparrow \infty$, as we shall see below.

We are going to get our asymptotic estimates by comparison with the heat kernel $K(x,y,t)$ for $\mathbb{R}^n$, which we now discuss. We begin by considering the initial-value problem:

$$\begin{cases} u_t - \Delta u = 0 \text{ in } \mathbb{R}^n \times (0,\infty) \\ u = \delta_y \text{ on } \mathbb{R}^n \times \{0\}, \end{cases}$$

where (as above) δ_y denotes the Dirac delta measure concentrated at y and where the initial value is to be taken in the measure theoretic sense that

$$\lim_{t\downarrow 0}(u(\cdot,t),\psi)_{L^2(\Omega)} = \psi(y)\ \forall \psi \in C_c^\infty(\mathbb{R}^n).$$

To solve this initial value problem, we first use Fourier transform to solve the problem

$$\begin{aligned} u_t - \Delta u &= 0 \text{ on } \mathbb{R}^n \times (0,\infty) \\ u &= \varphi \text{ on } \mathbb{R}^n \times \{0\}, \end{aligned}$$

where $\varphi \in C_c^\infty(\mathbb{R}^n)$ is given. Indeed, denoting Fourier transform of $u(x,t)$ with respect to x by $\hat{u}(\xi,t)$ (assuming for the moment that all quantities make sense and have the appropriate differentiability properties with respect to t), we get

$$\begin{aligned} \hat{u}_t + |\xi|^2\hat{u} &= 0 \text{ on } \mathbb{R}^n \times (0,\infty) \\ \hat{u}(\xi,0) &= \hat{\varphi}(\xi) \text{ on } \mathbb{R}^n, \end{aligned}$$

which integrates to give

$$\hat{u}(\xi,t) = \hat{\varphi}(\xi)e^{-|\xi|^2 t}.$$

Now the point is that the right side here is the Schwarz class $\mathcal{S}$ of smooth and rapidly decaying functions, so we can define $u(x,t)$ by taking inverse Fourier transform. (See the discussion of Lecture 5.) Thus we define

$$(1) \qquad u(x,t) = \frac{1}{(2\pi)^{n/2}}\int_{\mathbb{R}^n} e^{ix\cdot\xi - |\xi|^2 t}\hat{\varphi}(\xi)\,d\xi.$$

Using the definition of $\hat{\varphi}$ we thus have

$$u(x,t) = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{i(x-y)\cdot\xi - |\xi|^2 t}\,d\xi\,\varphi(y)\,dy.$$

Notice that by completing the square we can write

$$|\xi|^2 - \frac{i(x-y)\cdot\xi}{t} = \sum_{j=1}^n(\xi_j - \frac{i(x_j-y_j)}{2t})^2 + \frac{|x-y|^2}{4t^2},$$

so by change of variable $\zeta = \xi - i(x-y)/2$ (formally justified by using the Cauchy integral theorem over a suitable contour for each of the variables ξ_j), we get

$$u(x,t) = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{-t|\zeta|^2}e^{-|x-y|^2/4t}\varphi(y)\,dy\,d\zeta.$$

Now by change of variable $\eta = \zeta/t$ we get

$$\int_{\mathbb{R}^n} e^{-t|\zeta|^2}\,d\zeta = \frac{1}{t^{n/2}}\int_{\mathbb{R}^n} e^{-|\eta|^2}\,d\eta = \frac{\pi^{n/2}}{t^{n/2}},$$

and thus finally

$$u(x,t) = \frac{1}{(4\pi t)^{n/2}}\int_{\mathbb{R}^n} e^{-|x-y|^2/4t}\varphi(y)\,dy,$$

or

$$(2) \qquad u(x,t) = \int_{\mathbb{R}^n} K(x,y,t)\varphi(y)\,dy,$$

where

$$K(x,y,t) = \frac{e^{-\frac{|x-y|^2}{4t}}}{(4\pi t)^{n/2}}.$$

We should now check that u defined in (1) (or (2)) does furnish the solution of the initial value problem. Indeed it is evident by (1) that the initial condition holds. Also since $\varphi \in C_c^\infty(\mathbb{R}^n)$ we have $|\hat{\varphi}(\xi)| \le C_\ell(1+|\xi|)^{-\ell}$ for every $\ell \ge 1$, so we can differentiate under the integral in (1) as many times as we wish for $t \ge 0$, so (1) (or (2)) gives us a $C^\infty(\mathbb{R}^n \times [0,\infty))$ solution of the heat equation.

As a matter of fact the function $K(x,y,t)$ is the heat kernel we are seeking, because, firstly, direct computation shows that $K(x,y,t)$ satisfies the heat equation with respect to the variables x,t, and, secondly, by (1), (2) we see that

$$\lim_{t\downarrow 0} K(\cdot,y,t) = \delta_y$$

in the required measure-theoretic sense.

As in the case of the heat kernel on the bounded domain Ω, we can also approximate $K(x,y,t)$ by $C^\infty(\mathbb{R}^n \times [0,\infty))$ solutions of the heat equation with initial data ψ_i, where ψ_i is again a sequence in $C_c^\infty(\mathbb{R}^n)$ with support $\psi_i \subset$

$\{x : |x-y| < 1/i\}$, $\psi_i \ge 0$, and $(\psi_i, \psi)_{L^2} \to \psi(y)$ for any $\psi \in C_c^\infty(\mathbb{R}^n)$. By (1), (2) such a sequence k_i is given explicitly by

$$k_i(x,t) = \int_{\mathbb{R}^n} K(x,y,t)\psi_i(y)\,dy.$$

Our aim is to prove Weyl's theorem by showing that the heat kernel $p(x,y,t)$ of the previous lecture is well approximated by $K(x,y,t)$ as $t \downarrow 0$. The proof of this will be based on the maximum principle for the heat equation, which we now present.

Lemma (Parabolic Maximum Principle.) *Suppose Ω is a bounded domain, $u \in C^0(\overline{\Omega} \times [0,T]) \cap C^2(\Omega \times (0,T))$, and $u_t - \Delta u \le 0$ in $\Omega \times (0,T)$ (where $T > 0$). Then for each $t \le T$,*

$$\sup_{\Omega\times(0,t)} u = \sup_{(\partial\Omega\times(0,t))\cup(\Omega\times\{0\})} u.$$

Remark: There is also a parabolic maximum principle in case $-\Delta$ is replaced by a general second order elliptic operator, assuming the coefficients are appropriately bounded—see problem 11.5 below.

Proof of lemma: Let $\varepsilon > 0$ be arbitrary and note that if $v = u - \varepsilon t$ then

$$\text{(i)} \qquad v_t - \Delta v < 0 \text{ on } \Omega \times (0,T).$$

Now let $0 < t_1 < T$. We know that $v|\overline{\Omega}\times[0,t_1]$ attains a maximum somewhere in $\overline{\Omega} \times [0,t_1]$. If this maximum is attained at a point (x_0,t_0) with $x_0 \in \Omega$ and $0 < t_0 < t_1$, then the gradient $(v_t, D_x v)$ vanishes at (x_0,t_0), and the Hessian matrix $(D_{x^i x^j} v)$ is semi-negative definite at (x_0,t_0), hence in particular $\Delta v \le 0$ at (x_0,t_0). Thus we would have

$$\text{(ii)} \qquad v_t - \Delta v \ge 0 \text{ at } (x_0,t_0),$$

thus contradicting (i).

If on the other hand $x_0 \in \Omega$ and $t_0 = t_1$, then we get the same facts at (x_0,t_0), except that now $v_t \ge 0$ (rather than $v_t = 0$) at (x_0,t_0). Thus we still get (ii), and (i) is again contradicted. Thus v attains its maximum on $\overline{\Omega} \cup \partial\Omega \times [0,t_1]$. Letting $\varepsilon \downarrow 0$, we deduce the same for u, as required.

Proof of Weyl's Theorem: Let $q_i \in C^\infty(\overline{\Omega}_i \times [0,\infty))$ and $k_i \in C^\infty(\mathbb{R}^n \times [0,\infty))$ be the approximations of the heat kernels $p(x,y,t)$ of Ω and $K(x,y,t)$ of $\mathbb{R}^n$ as described above. Then

$$q_i(x,0) \equiv k_i(x,0) \equiv \psi_i(x) \text{ on } \Omega_i,$$

and

$$q_i|\partial\Omega_i \times [0,\infty) = 0 \le k_i|\partial\Omega_i \times [0,\infty).$$

Hence by the maximum principle (applied to the functions q_i, $k_i - q_i$ and $q_i - k_i$), we conclude that

$$0 \le q_i \le k_i \text{ and } \sup_{\Omega_i\times(0,T)} (k_i - q_i) \le \sup_{\partial\Omega_i\times(0,T)} k_i.$$

Passing to the limit as $i \to \infty$, we thus conclude that

$$0 \le p(x,y,t) \le K(x,y,t) \text{ and}$$
$$\sup_{(x,t)\in\Omega\times(0,T)} (K(x,y,t) - p(x,y,t)) \le \sup_{(x,t)\in\partial\Omega\times(0,t)} K(x,y,t)$$

for any fixed $y \in \Omega$. Taking $y \in \Omega_\sigma = \{x \in \Omega : \operatorname{dist}(x,\partial\Omega) < \sigma\}$, we thus have in particular that

$$0 \le p(y,y,t) \le K(y,y,t) \equiv \frac{1}{(4\pi t)^{n/2}}$$
$$\text{and} \sup_{y\in\Omega_\sigma, t\in(0,T)} \Big(\frac{1}{(4\pi t)^{n/2}} - p(y,y,t)\Big) \le Q(\sigma),$$

where

$$Q(\sigma) = \sup_{0<t<\infty} \frac{e^{-\sigma^2/(4t)}}{(4\pi t)^{n/2}} < C\sigma^{-n},$$

where $C = C(n)$ is a fixed constant. Notice that by integration over Ω_σ, the second inequality above gives

$$0 \le \int_{\Omega_\sigma} \textstyle\sum_{j=1}^\infty e^{-\lambda_j t}\varphi_j^2(y)\,dy - \dfrac{|\Omega_\sigma|}{(4\pi t)^{n/2}} \le Q(\sigma,T), \quad 0 < t < T,$$

and by the first inequality

$$\int_{\Omega\sim\Omega_\sigma} \textstyle\sum_{j=1}^\infty e^{-\lambda_j t}\varphi_j^2(y)\,dy \le \dfrac{|\Omega \setminus \Omega_\sigma|}{(4\pi t)^{n/2}}.$$

Combining these inequalities and using the fact that $\int_\Omega \varphi_j^2(y)\,dy = 1$, we get that then

$$\lim_{t\downarrow 0} (4\pi t)^{n/2} \textstyle\sum_{j=1}^\infty e^{-\lambda_j t} = |\Omega|.$$

We now use a Tauberian Theorem (from Feller Vol. 2, p.443, Th. 1):

Theorem. *For any non-negative Borel measure μ on $[0,\infty)$ and any constants $\beta, q > 0$, the following two statements are equivalent:*

(i) $$\int_0^\infty e^{-t\lambda}\, d\mu(\lambda) \sim \beta t^{-q} \text{ as } t \downarrow 0$$

(ii) $$\mu([0,R]) \sim \frac{\beta R^q}{\Gamma(q+1)} \text{ as } R \uparrow \infty.$$

In our case the hypothesis (i) holds with $\mu = \sum_{j=1}^\infty \delta_{\lambda_j}$, where δ_{λ_j} denotes the Dirac measure with unit mass concentrated at λ_j, with $q = n/2$, and with $\beta = (4\pi)^{-n/2}|\Omega|$. Taking $R = \lambda$ (so that $\mu([0,R]) = N_\lambda$ if λ differs from each of the eigenvalues λ_j), Weyl's theorem is then established.

LECTURE 11 PROBLEMS

11.1 Discuss the extension of Theorem 1 to the case when there is a given function f on the right side of $(**)$. Include in your discussion at least the points (i) existence of weak solutions if $f \in L^2(\Omega \times (0,T))$ for each $T > 0$, and (ii) regularity in case the coefficients are in $C^\infty(\overline{\Omega})$, $\Omega \in C^\infty$, $\varphi \in C_c^\infty(\Omega)$ and $f \in C^\infty(\overline{\Omega} \times [0,\infty))$ such that support of $f(\cdot,t)$ is a compact subset of Ω for each $t \in [0,\infty)$. Hint for (i): Let $f_j(t) = \langle f(t), \varphi_j\rangle_{L^2(\Omega)}$ and look for a weak solution in the form $\sum_{j=1}^\infty a_j(t)\varphi_j(x)$, where the $a_j(t)$ are to be determined in terms of f, φ.

11.2 If $a_{\alpha\beta} \in C^\infty(\Omega)$, if $\psi \in C_c^\infty(\Omega)$, and if φ_j is the j^{th} eigenfunction for L as in Lecture 10 (where as usual we assume the coercivity condition (C) and that the $a_{\alpha\beta}$ are bounded on Ω), prove that $|\langle \psi, \varphi_j\rangle_{L^2(\Omega)}| \le C_\ell\, j^{-\ell}$ for each $\ell \ge 1$, where C_ℓ depends on the coefficients $a_{\alpha\beta}$ and on ψ but does not depend on j. Hint: Consider $\langle \psi, L^q\varphi_j\rangle_{L^2(\Omega)}$ for any $q = 1, 2, \ldots$

11.3 Prove that the solution u of $(**)'$ obtained in Theorem 1 is unique as claimed in the statement of the theorem. Hint: If u_1, u_2 are two solutions, by considering $\psi(x,t) = \varphi_j^{(k)}(x)e^{\lambda_j t}\gamma(t)$ in the weak formulation of the problem for u_1, u_2, where $\gamma \in C_c^\infty(0,\infty)$ and $\{\varphi_j^{(k)}\}_{k=1,2,\ldots}$ is a sequence of $C_c^\infty(\Omega)$ functions converging in the Sobolev norm $\|\ \|_{m,\Omega}$ to φ_j (the j^{th} eigenfunction of L as in Lecture 10), prove that $\int_0^\infty \dot\gamma(t)e^{\lambda_j t}a_j(t)\,dt = 0$, where $a_j(t) = \langle (u_1(t) - u_2(t)), \varphi_j\rangle_{L^2(\Omega)}$.

11.4 Prove that if $y \in \Omega$, then there is a sequence of $C_c^\infty(\Omega)$ functions $\{\psi_j\}$ with $\psi_j \ge 0$, support $\psi_j \subset \{x : |x-y| < 1/j\}$, and with $\psi_j \to \delta_y$ (the Dirac delta measure) in the sense that $\lim \int_\Omega \psi_j(x) f(x)\,dx = f(y)$, $f \in C_c^0(\Omega)$. Hint: Consider $\varphi^{(\sigma_j)}$ for a suitable sequence $\sigma_j \downarrow 0$, where φ is the usual mollifier.

11.5 Generalize the parabolic maximum principle to the case when Δu is replace by $\sum_{i,j=1}^n a_{ij}(x,t)D_iD_ju$, where $\sum a_{ij}(x,t)\xi_i\xi_j \ge \mu|\xi|^2$ and $|a_{ij}| \le \mu^{-1}$ for some fixed constant $\mu > 0$.

Lecture 12

Schauder Theory

Here we want to develop the Schauder estimates for the class of equations considered in Lecture 6; that is we consider weak solutions of

$$(\ddagger) \qquad \sum_{|\alpha|,|\beta|\le m} D^\beta(a_{\alpha\beta} D^\alpha u) = \sum_{|\beta|\le m} D^\beta f_\beta,$$

where f_β are given functions. We also consider classical solutions of the non-divergence form equation

$$(\ddagger\ddagger) \qquad \sum_{|\alpha|,|\beta|\le m} a_{\alpha\beta} D^{\alpha+\beta} u = f.$$

We establish the interior Schauder estimates, and also the boundary estimates for the case of Dirichlet boundary conditions and also the "partially free" boundary conditions considered in Lecture 9.

We assume here, for the interior estimates and for the boundary estimates subject to Dirichlet data, only the weaker ellipticity condition

$$(\mathrm{E})' \qquad \sum_{|\alpha|=|\beta|=m} a_{\alpha\beta}(x)\xi^\alpha\xi^\beta \ge \lambda|\xi|^{2m}, \quad \xi \in \mathbb{R}^n,$$

where $\lambda > 0$. For the boundary regularity results for the partially free boundary data discussed in Lecture 9 it is necessary to assume the stronger coercivity condition (C) of Lecture 11 (which is of course guaranteed by the stronger ellipticity condition (E) of Lecture 6).

For the Schauder estimates it is necessary to impose appropriate continuity restrictions on all the coefficients $a_{\alpha\beta}$, $|\alpha|, |\beta| \le m$ and on the given functions f_β.

We first need to discuss some preliminary facts about Hölder continuity and Hölder norms and semi-norms. First, if $f : \Omega \to \mathbb{R}$, if $A \subset \Omega$, and if $\mu \in (0,1]$, we write

$$[f]_{\mu,A} = \sup_{x,y\in A,\, x\neq y} \frac{|f(x)-f(y)|}{|x-y|^{\mu}}.$$

The class of all functions $f : \Omega \to \mathbb{R}$ which satisfy the condition $[f]_{\mu,K} < \infty$ for each compact $K \subset \Omega$ will be denoted $C^{0,\mu}(\Omega)$. Of course this means in particular that the $f \in C^{0,\mu}(\Omega)$ are continuous in Ω: $C^{0,\mu}(\Omega) \subset C^0(\Omega)$. The class of $f \in C^0(\Omega)$ such that $[f]_{\mu,\Omega} < \infty$ will be denoted $C^{0,\mu}(\overline{\Omega})$. More generally, for any integer $k \geq 0$, we let $C^{k,\mu}(\Omega)$ denote the set of all functions in $u \in C^k(\Omega)$ with $D^\beta u \in C^{0,\mu}(\Omega)$ for each multi-index β with $|\beta| = k$ and $C^{k,\mu}(\overline{\Omega})$ will denote the set of $C^k(\Omega)$ functions u such that $D^\beta u \in C^{0,\mu}(\overline{\Omega})$ for each multi-index β with $|\beta| = k$. For $u \in C^{k,\mu}(\overline{\Omega})$ we let

$$|u|_{k,\mu;\Omega} = |u|_{k;\Omega} + [D^k u]_{\mu;\Omega},$$

where $|u|_{k,\Omega} = \sum_{j=0}^{k} |D_j u|_{0,\Omega}$ and where we use the notation $|f|_{0,A} = \sup_A |f|$, $|D^j u|_{0,A} = \sum_{|\beta|=j} |D^\beta u|_{0,A}$ and $[D^k u]_{\mu,A} = \sum_{|\beta|=k} [D^\beta u]_{\mu,A}$.

Functions in $C^{0,\mu}(\overline{\Omega})$ are said to be uniformly Hölder continuous with exponent μ; the functions in $C^{0,\mu}(\Omega)$ are said to be locally Hölder continuous with exponent μ on Ω. With this terminology, we can thus describe the functions in $C^{k,\mu}(\overline{\Omega})$ as the class of $C^k(\Omega)$ functions such that all partial derivatives of order k are uniformly Hölder continuous of exponent μ on Ω.

Notice that if Ω is a bounded Lipschitz domain then any $u \in C^{k,\mu}(\Omega)$ has the property that all its partial derivatives $D^\beta u$, $|\beta| \leq k-1$ are Lipschitz on Ω, so in particular $C^\ell(\overline{\Omega}) \subset C^{k,\mu}(\Omega)$ for any bounded Lipschitz domain and any $\mu \in (0,1]$ and $\ell \in \{0,\ldots,k\}$.

The interior Schauder estimates for equations in "divergence" and "non-divergence" form are then as follows; here and subsequently we require $\mu \in (0,1)$, so that $\mu = 1$ is not included in the discussion.

Theorem 1. *If $u \in C^{m,\mu}(\overline{B}_R(x_0))$ is a weak solution of the equation (‡), where the ellipticity condition (E)′ holds and where $|a_{\alpha\beta}|_{0,\mu,B_R(x_0)} \leq \Lambda$, then*

$$|u|_{m,\mu,B_{\theta R}(x_0)} \leq C\big(|u|_{0,B_R(x_0)} + \textstyle\sum_{|\beta|\leq m} |f_\beta|_{0,\mu,B_R(x_0)}\big)$$

for each $\theta \in (0,1)$, where C depends on $n, m, R, \theta, \gamma, \mu, \Lambda$.

Theorem 2. *If $u \in C^{2m,\mu}(\overline{B}_R(x_0))$ is a weak solution of the equation (‡‡), where*

the ellipticity condition (E)′ holds and where $|a_{\alpha\beta}|_{0,\mu,B_R(x_0)} \leq \Lambda$, then

$$|u|_{2m,\mu,B_{\theta R}(x_0)} \leq C\big(|u|_{0,B_R(x_0)} + |f|_{0,\mu,B_R(x_0)}\big)$$

for each $\theta \in (0,1)$, where C depends on $n, m, R, \theta, \gamma, \mu, \Lambda$.

Remark: One can also give a scale-invariant version of Theorems 1,2; see problem 12.8.

For the proofs of both Theorem 1 and Theorem 2 we need the following lemma concerning solutions of the equations of the form

$$(*) \qquad \textstyle\sum_{|\alpha|=|\beta|=m} a^0_{\alpha\beta} D^{\alpha+\beta} u = \sum_{|\beta|=2m-k} D^\beta f_\beta,$$

where $k \in \{0,\ldots,2m\}$ is an integer and $a^0_{\alpha\beta}$ are constants satisfying the ellipticity and boundedness condition

$$(**) \qquad \begin{cases} \sum_{|\alpha|=|\beta|=m} a^0_{\alpha\beta}\xi^\alpha\xi^\beta \geq \gamma|\xi|^{2m} & \forall\, \xi \in \mathbb{R}^n \\ |a^0_{\alpha\beta}| \leq \Lambda. \end{cases}$$

Then we have

Lemma 1 (Main Lemma of Interior Schauder Theory.) *Suppose $(**)$ holds and u is a weak solution of the equation $(*)$. Assume also that $[D^k u]_{\mu,\mathbb{R}^n} < \infty$. Then*

$$[D^k u]_{\mu,\mathbb{R}^n} \leq C \textstyle\sum_{|\beta|=2m-k} [f_\beta]_{\mu,\mathbb{R}^n},$$

where $C = C(\Lambda, \lambda, k, n, \mu)$.

Proof of Lemma 1: If the theorem is false, there are sequences $a^{(j)}_{\alpha\beta}, u_j, f^{(j)}_\beta$ such that u_j is a weak solution of $(*)$, $(**)$ with $a^{(j)}_{\alpha\beta}$, $f^{(j)}_\beta$ in place of $a^0_{\alpha\beta}$, f_β respectively, and with

$$[f^{(j)}_\beta]_{\mu,\mathbb{R}^n} < j^{-1}[D^k u_j]_{\mu,\mathbb{R}^n} < \infty.$$

Now by definition of $[D^k u_j]_{\mu,\mathbb{R}^n}$ there are points $x_j \neq y_j$ such that $|D^k u_j(y_j) - D^k u_j(x_j)| \geq (2N)^{-1}[D^k u_j]_{\mu,\mathbb{R}^n}$, where N is the number of terms in the set of multi-indices $\{\alpha\}_{|\alpha|=k}$. Then, taking $\lambda_j = [D^k u_j]_{\mu,\mathbb{R}^n}$ and $\sigma_j = |y_j - x_j|$, and defining $\widetilde{u}_j(x) = \sigma_j^{-\mu-k}\lambda_j^{-1} u_j(\sigma_j x + x_j)$ and $\widetilde{f}_\beta(x) = \sigma_j^{-\mu}\lambda_j^{-1} f^j_\beta(\sigma_j x + x_j)$ we see that the equation $(*)$, $(**)$ holds with $a^{(j)}_{\alpha\beta}, \widetilde{u}_j, \widetilde{f}^j_\beta$ in place of $a^0_{\alpha\beta}, u, f_\beta$ respectively, and

$$(1) \qquad [D^k \widetilde{u}_j]_{\mu,\mathbb{R}^n} = 1, \quad [\widetilde{f}^{(j)}_\beta]_{\mu,\mathbb{R}^n} \leq j^{-1},$$

and

$$(2) \qquad |D^k\widetilde{u}_j(\eta_j) - D^k\widetilde{u}_j(0)| \ge (2N)^{-1}$$

for some sequence η_j with $|\eta_j| = 1$ for each j. (In fact this holds with $\eta_j = \sigma_j^{-1}(y_j - x_j)$.) Now let $p_j(x)$ by the degree-k Taylor polynomial for $\widetilde{u}_j$ at 0, so that $p_j(x) = \sum_{|\alpha|\le k}(\alpha!)^{-1}D^\alpha\widetilde{u}_j(0)x^\alpha$. Since (by Taylor's theorem) $\widetilde{u}_j(x) = p_j(x) + R_j$, with $\sup_{B_R(0)}|R_j| \le R^k \sup_{B_R(0)}|D^k\widetilde{u}_j(x) - D^k\widetilde{u}_j(0)|$, we then have

$$(3) \qquad \sup_{B_R(0)}|v_j| \le R^k \sup_{x\in B_R(0)}|D^k\widetilde{u}_j(x) - D^k\widetilde{u}_j(0)| \le R^{k+\mu}[D^k\widetilde{u}_j]_{\mu,\mathbb{R}^m} = R^{k+\mu},$$

where

$$(4) \qquad v_j = \widetilde{u}_j - p_j.$$

Notice that v_j is a weak solution of the equation $(*)$ with $a^{(j)}_{\alpha\beta}$, $\widetilde{f}^{(j)}_\beta - \widetilde{f}^{(j)}_\beta(0)$ in place of $a^0_{\alpha\beta}$, f_β respectively, and (by (1), (2))

$$(5) \qquad [D^k v_j]_{\mu,\mathbb{R}^n} = 1, \quad |D^k v_j(x_j)| \ge (2N)^{-1}, \; D^k v_j(0) = 0.$$

Also by taking the $(k-\ell)$-order Taylor approximation to $D^\ell v_j$ (at 0) we deduce from (3) that

$$(6) \qquad |D^\ell v_j|_{0,B_R(0)} \le CR^{k-\ell+\mu}.$$

Thus by (5), (6) and the Arzela-Ascoli lemma there is a subsequence $v_{j'}$ converging uniformly, together with its derivatives up to and including order k, on compact subsets of $\mathbb{R}^n$ to a function $v \in C^{k,\mu}(\mathbb{R}^n)$ such that

$$(7) \qquad [D^k v]_{\mu,\mathbb{R}^n} \le 1, \quad D^k v(0) = 0, \quad |D^k v(\eta)| = 1$$

for some η with $|\eta| = 1$. (In fact $\eta = \lim \eta_{j'}$.) Also (using (1) and (3)) we have

$$(8) \qquad |v|_{0,B_R(0)} \le R^{k+\mu} \quad \forall R > 0$$

and v is a weak solution of the equation

$$\textstyle\sum_{|\alpha|=|\beta|=m} a^0_{\alpha\beta} D^{\alpha+\beta} v = 0$$

locally on $\mathbb{R}^n$, where $a^0_{\alpha\beta} = \lim_{j'\to\infty} a^{(j')}_{\alpha\beta}$. By virtue of the L^2 estimates of Lecture 8 we then have that $v \in C^\infty$ and that for each $R > 0$ and any $q = 1, 2, \ldots$,

$$|D^q v|_{0,B_{R/2}(0)} \le CR^{-q}|v|_{0,B_R(0)},$$

where C does not depend on R. But then by virtue of (8) we deduce

$$|D^q v|_{0,B_{R/2}(0)} \le CR^{k+\mu-q},$$

and taking $q = k+1$ gives a negative exponent on the right, so we can let $R \to \infty$ to conclude that $D^{k+1}v \equiv 0$ on $\mathbb{R}^n$, so that v is a polynomial of degree $\le k$. Thus $D^k v$ is constant, which evidently contradicts (7). This completes the proof.

We are now ready to prove Theorem 1 and Theorem 2; we give only the proof of Theorem 1. The proof of Theorem 2 is very similar (in fact a bit simpler), and is left as an exercise. (See problem 12.2.)

In the proof of Theorem 1 we shall need the following interpolation inequality:

$$(\dagger) \qquad R^\ell|D^\ell u|_0 \le \varepsilon R^{k+\mu}[D^k u]_\mu + C|u|_0$$

for each $\varepsilon > 0$ and $1 \le \ell \le k$, where $C = C(\varepsilon, \mu, k, n)$. This is valid for any $u \in C^{k,\mu}(\overline{B}_R(x_0))$ and all norms and semi-norms being over $B_R(x_0)$. The proof involves nothing more than one-dimensional calculus along line segments in $\mathbb{R}^n$. (See problem 12.5 below.)

Proof of Theorem 1: By rescaling and translation (and since we are presently allowing the constant C in the conclusion of Theorem 1 to depend on R) it suffices to prove Theorem 1 under the assumption that $B_R(x_0) = B_1(0)$ (i.e. $R = 1$ and $x_0 = 0$). Λ continues to denote a constant such that

$$(\ddagger) \qquad |a_{\alpha\beta}|_{0,B_1(0)} + [a_{\alpha\beta}]_{\mu,B_1(0)} \le \Lambda.$$

Let $B_\sigma(y)$ be an arbitrary ball with $B_\sigma(y) \subset B_1(0)$. We are going to get some estimates on the various Hölder norms over $B_\sigma(y)$ (or actually over $B_{\sigma/2}(y)$) and it is convenient to make another scaling, taking $B_\sigma(y)$ to the unit ball. Thus we introduce new variable $z = \sigma^{-1}(x-y)$ and let $\widetilde{u}(z) = u(x)$ (i.e. $\widetilde{u}(z) = u(y+\sigma z)$); then in terms of $\widetilde{u}$ the equation can be written

$$(1) \qquad \textstyle\sum_{|\alpha|,|\beta|\le m} D^\beta(\widetilde{a}_{\alpha\beta}D^\alpha\widetilde{u}) = \sum_{|\beta|\le m} D^\beta \widetilde{f}_\beta,$$

where $\widetilde{a}_{\alpha\beta}(z) = \sigma^{2m-|\alpha|-|\beta|}a_{\alpha\beta}(y+\sigma z)$ and $\widetilde{f}_\beta = \sigma^{2m-|\beta|}f_\beta(y+\sigma z)$ for $z \in B_1(0)$. Notice that then (by (‡) and the fact that $\sigma \le 1$) we have

$$(2) \qquad |\widetilde{a}_{\alpha\beta}|_{0,B_1(0)} \le \Lambda \text{ and } [\widetilde{a}_{\alpha\beta}]_{\mu,B_1(0)} \le \Lambda\sigma^\mu.$$

We have in mind to apply Lemma 1 to $\widetilde{u}$, but Lemma 1 only applies to constant coefficients, so we "freeze" the coefficients of the operator at the center point y of $B_\sigma(y)$, taking $a^0_{\alpha\beta} = \widetilde{a}_{\alpha\beta}(0)(= a_{\alpha\beta}(y))$, and put all the correction terms on the right. Thus we get

(3)
$$\textstyle\sum_{|\alpha|,|\beta|\le m} D^\beta(a^0_{\alpha\beta}D^\alpha\widetilde{u}) = \sum_{|\alpha|,|\beta|\le m} D^\beta((a^0_{\alpha\beta}-\widetilde{a}_{\alpha\beta})D^\alpha\widetilde{u}) + \sum_{|\beta|\le m} D^\beta \widetilde{f}_\beta.$$

Next we must address the fact that Lemma 1 applies to functions over all of $\mathbb{R}^n$, whereas $\widetilde{u}$ is defined only on the unit ball $B_1(0)$. To remedy this in an appropriate way we are going to instead look at the function $v = \varphi\widetilde{u}$, where φ is a cutoff function in $B_1(0)$. In fact let $\varphi \in C^\infty_c(\mathbb{R}^n)$ have the properties that $\varphi \equiv 1$ on $B_{1/2}(0)$ and support $\varphi \subset B_1(0)$, and the usual property that $|D^\alpha\varphi|_{0,B_1(0)} \le C$, $C = C(n,\alpha)$. Since, by the Leibniz formula, $v = \varphi\widetilde{u}$ has derivatives $D^\alpha(v\varphi) = \varphi D^\alpha\widetilde{u} + \sum_{\gamma+\delta=\alpha,|\delta|<|\alpha|}\frac{\alpha!}{\gamma!\delta!}D^\gamma\varphi D^\delta\widetilde{u}$, we see that (weakly) v satisfies

$$\textstyle\sum_{|\alpha|,|\beta|\le m} D^\beta(a^0_{\alpha\beta}D^\alpha v) = \sum_{|\alpha|,|\beta|\le m} D^\beta\bigl(\varphi a^0_{\alpha\beta}D^\alpha\widetilde{u} + \sum_{|\gamma|\le m,|\delta|<m} c_{\alpha\beta\gamma\delta}D^\gamma\varphi D^\delta\widetilde{u}\bigr)$$

for suitable constants $c_{\alpha\beta\gamma\delta}$ with $|c_{\alpha\beta\gamma\delta}| \le C\Lambda$. Since

$$\textstyle D^\beta(\varphi a^0_{\alpha\beta}D^\alpha\widetilde{u}) = \varphi D^\beta(a^0_{\alpha\beta}D^\alpha\widetilde{u}) + \sum_{|\gamma|\le m,|\delta|<m} D^\delta(\widetilde{c}_{\alpha\beta\gamma\delta}D^\gamma\varphi D^\alpha u)$$

for suitable constants $\widetilde{c}_{\alpha\beta\gamma\delta}$ with $|\widetilde{c}_{\alpha\beta\gamma\delta}| \le C\Lambda$ (see problem 12.7 below), this can be written

(4)
$$\textstyle\sum_{|\alpha|,|\beta|\le m} D^\beta(a^0_{\alpha\beta}D^\alpha v) = \sum_{|\alpha|,|\beta|\le m}\bigl(\varphi D^\beta(a^0_{\alpha\beta}D^\alpha\widetilde{u}) + \sum_{|\gamma|\le m,|\delta|<m} D^\delta(\widetilde{c}_{\alpha\beta\gamma\delta}D^\gamma\varphi D^\alpha\widetilde{u}) + D^\beta(\sum_{|\gamma|\le m,|\delta|<m} c_{\alpha\beta\gamma\delta}D^\gamma\varphi D^\delta\widetilde{u})\bigr).$$

Since $\widetilde{c}_{\alpha\beta\gamma\delta}$ are constants and $|\delta| < m$, the term $D^\delta(\widetilde{c}_{\alpha\beta\gamma\delta}D^\gamma\varphi D^\alpha\widetilde{u})$ can be (weakly) rewritten $D^{\widetilde{\delta}}(\widetilde{c}_{\alpha\beta\gamma\delta}D^\gamma\varphi D^\alpha\widetilde{u}) - D^\delta(\widetilde{c}_{\alpha\beta\gamma\delta}D^{\widetilde{\gamma}}\varphi D^{\widetilde{\alpha}}\widetilde{u})$ with $\widetilde{\delta} = \delta + e_j$ for some j (so $|\widetilde{\delta}| \le m$), $\widetilde{\gamma} = \gamma + e_j$, and $\widetilde{\alpha} = \alpha - e_j$ (so $|\widetilde{\alpha}| = m-1$). Similarly any term $D^\beta(a^0_{\alpha\beta}D^\alpha v)$ on the left side with $|\beta| < m$ and $|\alpha| = m$ can be rewritten $D^{\widetilde{\beta}}(a^0_{\alpha\beta}D^{\widetilde{\alpha}}v)$ with $|\widetilde{\beta}| \le m$ and $|\widetilde{\alpha}| = m-1$. Thus (4) can be rewritten

$$\textstyle\sum_{|\alpha|=|\beta|=m} D^\beta(a^0_{\alpha\beta}D^\alpha v) = \sum_{|\alpha|,|\beta|\le m}\varphi D^\beta(a^0_{\alpha\beta}D^\alpha\widetilde{u}) + \sum_{|\alpha|<m,|\beta|\le m,|\gamma|\le m} D^\beta(b_{\alpha\beta\gamma}D^\gamma\varphi D^\alpha\widetilde{u})$$

for suitable constants $b_{\alpha\beta\gamma}$ with $|b_{\alpha\beta\gamma}| \le C\Lambda$, with $C = C(n,m)$. Then using (3) we get

(5)
$$\textstyle\sum_{|\alpha|=|\beta|=m} D^\beta(a^0_{\alpha\beta}D^\alpha v) = \sum_{|\beta|\le m}\varphi D^\beta\widetilde{f}_\beta + \sum_{|\alpha|<m,|\beta|,|\gamma|\le m} D^\beta(b_{\alpha\beta\gamma}D^\gamma\varphi D^\alpha\widetilde{u}) + \sum_{|\alpha|,|\beta|\le m}\varphi D^\beta((a^0_{\alpha\beta}-\widetilde{a}_{\alpha\beta})D^\alpha\widetilde{u}).$$

Now (see problem 12.7 below) we can write

$$\textstyle\varphi D^\beta\widetilde{f}_\beta = \sum_{|\gamma|,|\delta|\le m} D^\gamma(c_{\beta\gamma\delta}D^\delta\varphi\widetilde{f}_\beta)$$

for suitable constants $c_{\beta\gamma\delta}$ with $|c_{\beta\gamma\delta}| \le C$, $C = C(n,m)$. Similarly

$$\textstyle\varphi D^\beta((a^0_{\alpha\beta}-\widetilde{a}_{\alpha\beta})D^\alpha\widetilde{u}) = \sum_{|\gamma|,|\delta|\le m} D^\gamma(d_{\alpha\beta\gamma\delta}D^\delta\varphi(a^0_{\alpha\beta}-\widetilde{a}_{\alpha\beta})D^\alpha\widetilde{u}))$$

for suitable constants $d_{\alpha\beta\gamma\delta}$ with $|d_{\alpha\beta\gamma\delta}| \le C$, $C = C(n,m)$. Thus (5) can be written

$$\begin{aligned}(6)\quad \textstyle\sum_{|\alpha|=|\beta|=m} D^\beta(a^0_{\alpha\beta}D^\alpha v) &= \textstyle\sum_{|\beta|,|\gamma|,|\delta|\le m} D^\gamma(c_{\beta\gamma\delta}D^\delta\varphi\widetilde{f}_\beta)\\ &\quad \textstyle+ \sum_{|\alpha|,|\beta|,|\gamma|,|\delta|\le m} D^\gamma(d_{\alpha\beta\gamma\delta}D^\delta\varphi(a^0_{\alpha\beta}-\widetilde{a}_{\alpha\beta})D^\alpha\widetilde{u})\\ &\quad \textstyle+ \sum_{|\alpha|<m,|\beta|,|\gamma|\le m} D^\beta(b_{\alpha\beta\gamma}D^\gamma\varphi D^\alpha\widetilde{u}),\end{aligned}$$

where $b_{\alpha\beta\gamma}, c_{\beta\gamma\delta}, d_{\alpha\beta\gamma\delta}$ are constants with $|c_{\beta\gamma\delta}|, |d_{\alpha\beta\gamma\delta}| \le C$ and $|b_{\alpha\beta\gamma}| \le C\Lambda$, $C = C(n,m)$.

Note that v has compact support in $B_1(0)$ so can be extended to all of $\mathbb{R}^n$ by taking $v = 0$ on $\mathbb{R}^n \setminus \{0\}$. Thus, keeping in mind that if $|\beta| < m$ then we can always write $D^\beta g = D^{\widetilde{\beta}}\widetilde{g}$ for suitable $\widetilde{\beta}$ with $|\widetilde{\beta}| = m$ and for suitable $\widetilde{g}$ such that $[\widetilde{g}]_\mu \le C(|g|_0 + [g]_\mu)$ (see problem 12.6), we see that we can now apply Lemma 1 (with $k = m$) to the equation (6), giving

(7)
$$\textstyle[D^m v]_\mu \le C\bigl(\sum_{|\alpha|,|\beta|\le m}|(a^0_{\alpha\beta}-a_{\alpha\beta}(x))D^\alpha\widetilde{u}|_0 + [(a^0_{\alpha\beta}-a_{\alpha\beta}(x))D^\alpha\widetilde{u}]_\mu + \sum_{|\alpha|<m}(|D^\alpha\widetilde{u}|_0 + [D^\alpha\widetilde{u}]_\mu) + \sum_{|\beta|\le m}(|\widetilde{f}_\beta|_0 + [\widetilde{f}_\beta]_\mu)\bigr),$$

where all norms and seminorms are over $B_1(0)$, and where we used the rules $[f+g]_\mu \le [f]_\mu + [g]_\mu$, $[fg]_\mu \le |f|_0[g]_\mu + [f]_\mu|g|_0$ to estimate the Hölder coefficients of the terms $b_{\beta\gamma\delta}D^\gamma\varphi D^\delta u$. Using the same rules on the first group of terms on the right, we then have

(8)
$$\textstyle[D^m\widetilde{u}]_\mu \le C\bigl(\sum_{|\alpha|,|\beta|\le m}|a_{\alpha\beta}-a_{\alpha\beta}(y)|_0[D^\alpha\widetilde{u}]_\mu + \sum_{|\alpha|,|\beta|\le m}[a_{\alpha\beta}]_\mu|D^\alpha\widetilde{u}|_0 + \sum_{|\alpha|<m}(|D^\alpha\widetilde{u}|_0 + [D^\alpha\widetilde{u}]_\mu) + \sum_{|\beta|\le m}(|\widetilde{f}_\beta|_0 + [\widetilde{f}_\beta]_\mu)\bigr),$$

where all norms and semi-norms are over $B_1(0)$ and C depends only on $\theta, \Lambda, \gamma, \mu, m$ and n.

Now, by (2), $\max_{|\alpha|=|\beta|=m} |\widetilde{a}_{\alpha\beta} - \widetilde{a}^0_{\alpha\beta}|_{0,B_1(0)} \le \Lambda\sigma^\mu$ and by 1-dimensional calculus (working along the line segment $x + t(y-x)$ joining points $x, y \in B_1(0)$) we have, $\forall\, x, y \in B_1(0)$,

$$|D^k\widetilde{u}(x) - D^k\widetilde{u}(y)| \le |D^{k+1}\widetilde{u}|_0 |x-y|, \quad k \ge 0,$$

and hence

$$(9) \qquad [D^k\widetilde{u}]_\mu \le 2|D^{k+1}\widetilde{u}|_0, \quad 0 \le k \le m-1$$

Then, using (9) and the interpolation inequalities (†) in (8)

$$(10) \qquad [D^m\widetilde{u}]_{\mu,B_{1/2}(0)} \le C\big((\varepsilon + \sigma^\mu)[D^m\widetilde{u}]_\mu + |\widetilde{u}|_0 + \textstyle\sum_{|\beta|\le m} |\widetilde{f}_\beta|_{0,\mu}\big),$$

where C depends only on $\mu, \theta, n, m, \Lambda, \gamma$, and where all norms and seminorms on the right are still over $B_1(0)$.

Now by scaling and translation to get back to the original function u on the original ball $B_\sigma(y) \subset B_1(0)$, we see that we have proved

(11)
$$\sigma^{m+\mu}[D^m u]_{\mu,B_{\sigma/2}(y)} \le C\big((\varepsilon + \sigma^\mu)\sigma^{m+\mu}[D^m u]_{\mu,B_\sigma(y)} + |\widetilde{u}|_{0,B_\sigma(y)} + \textstyle\sum_{|\beta|\le m} |f_\beta|_{0,B_\sigma(y)} + \sigma^\mu [f_\beta]_{\mu,B_\sigma(y)}\big),$$

which implies

$$(12) \qquad \sigma^{m+\mu}[D^m u]_{\mu,B_{\sigma/2}(y)} \le C(\varepsilon + \sigma^\mu)\sigma^{m+\mu}[D^m u]_{\mu,B_\sigma(y)} + \gamma$$

for every ball $B_\sigma(y) \subset B_1(0)$, where $\gamma = C(|u|_{0,B_1(0)} + \sum_{|\beta|\le m} |f_\beta|_{0,\mu,B_1(0)})$ and $C = C(m, n, \Lambda, \lambda, \mu)$.

The proof is now completed by taking $\sigma = \varepsilon$, with ε small enough, and applying the Absorption Lemma (Lemma 2 of Lecture 6) with $\ell = m + \mu$ and with $S(A) = [D^m u]_{\mu,A}$, in order to conclude from (12) that

$$[D^m u]_{\mu,B_\theta(0)} \le C\gamma \quad (\text{with } \gamma \text{ as in (12)}),$$

and hence, by the interpolation inequalities (†), $|u|_{m,\mu,B_\theta(0)} \le C\gamma$ as required. This completes the proof of Theorem 1.

Now we describe the extension of the above theory to give boundary estimates for solutions of the "partially free" boundary-value problems considered in Lectures 8, 9. We begin by looking at weak solutions of the appropriate

constant coefficient problem in the half-space $\mathbb{R}^n_+ \equiv \{x \in \mathbb{R}^n : x^n > 0\}$. So assume $\ell \in \{0, \dots, m\}$ and $u \in \mathcal{H}^\ell$ is a solution of

$$(*)' \qquad \int_{\mathbb{R}^n_+} \textstyle\sum_{|\alpha|=|\beta|=m} a^0_{\alpha\beta} D^\alpha u D^\beta \zeta = \int_{\mathbb{R}^n_+} \sum_{|\beta|=2m-k} f_\beta D^\beta \zeta, \quad \zeta \in \mathcal{H}^\ell_c,$$

where $\mathcal{H}^\ell = H^\ell_0(\mathbb{R}^n_+)$ and $\mathcal{H}^\ell_c$ denotes the set of $\mathcal{H}^\ell$ functions which vanish identically outside a bounded subset of $\mathbb{R}^n_+$. Then we have the following analogue of Lemma 1:

Lemma 2. *If $k \in \{\ell, \dots, 2m\}$, $\ell \in \{0, \dots, m\}$, if $u \in \mathcal{H}^\ell \cap C^{k,\mu}$ with $\mu \in (0,1)$, and if $(*)'$, $(**)$ hold, then*

$$[D^k u]_{\mu,\mathbb{R}^n_+} \le C \textstyle\sum_{|\beta|=2m-k} [f_\beta]_{\mu,\mathbb{R}^n_+},$$

where $C = C(k, \mu, n, \lambda, \Lambda)$.

The proof is left as an exercise—it is almost identical to the proof of Lemma 1, except that the L^2-estimates for the half-balls B^+_R are used in place of the L^2-estimates for the balls $B_R(0)$. See Lectures 8, 9 for these L^2 estimates; we actually give a discussion, independent of Lectures 8,9, of the necessary estimates in the supplement to this lecture. Thus the theory developed here is independent of the boundary regularity discussion of Lectures 8,9.

We now want to state the boundary Schauder estimates and for this we first need to give the definition of $C^{k,\mu}$ domain (analogous to the definition of C^k domain in Lecture 9) as follows:

Definition: Let $x_0 \in \partial\Omega$ and $R > 0$. We say that $\partial\Omega \cap B_R(x_0)$ is $C^{k,\mu}$ if there is a $C^{k,\mu}$ function $\psi : \mathbb{R}^{n-1} \to \mathbb{R}$ and coordinates $y = x_0 + Q(x - x_0)$, with Q an orthogonal transformation of $\mathbb{R}^n$, such that

$$\Omega \cap B_R(x_0) = \{x_0 + Q^{-1}(y - x_0) : y^n > \psi(y^1, \dots, y^{n-1})\} \cap B_R(x_0).$$

Notice that then in particular we have

$$\begin{aligned}\partial\Omega \cap B_R(x_0) &= \{x_0 + Q^{-1}(y - x_0) : y^n = \psi(y^1, \dots, y^{n-1})\} \cap B_R(x_0) \\ &\equiv \{x_0 + Q^{-1}(y - x_0) : y \in \operatorname{graph}\psi \cap B_R(0)\},\end{aligned}$$

We can of course assume for convenience of notation $x_0 = 0$ and the inward pointing unit normal of $\partial\Omega$ at $0 (= x_0)$ is e_n, so we can apply the above with Q equal to the identity. So the coordinate transformation $\Psi : x = (x^1, \dots, x^n) \mapsto \xi = (y^1, \dots, y^n)$, where the $y^i = x^i$ for $i = 1, \dots, n-1$

and $y^n = x^n - \psi(x^1, \ldots, x^{n-1})$, so that $\Psi(B_R(0))$ is an open set U containing 0, $\Psi(\Omega \cap B_R(0)) = U^+ = \{y \in U : y^n > 0\}$, and $\Psi(\partial\Omega \cap B_R(0)) = U \cap (R^{n-1} \times \{0\})$. (We say the transformation Ψ "flattens the boundary.")

Then we have the following analogues of Theorem 1 for the solutions $u \in \mathcal{H}^\ell(\Omega \cap B_R(x_0))$ of

$$\int_{\Omega\cap B_R(x_0)} \sum_{|\alpha|,|\beta|\le m} a_{\alpha\beta} D^\alpha u D^\beta \zeta = \int_{\Omega\cap B_R(x_0)} \sum_{|\beta|\le m} f_\beta D^\beta \zeta$$

for all $\zeta \in \mathcal{H}^\ell_c(\Omega \cap B_R(x_0))$, where $\mathcal{H}^\ell(\Omega \cap B_R(x_0))$ is the set of $u \in H^m(\Omega \cap B_R(x_0))$ such that $\varphi u \in H^\ell_0(\Omega \cap B_R(x_0))$ for every $\varphi \in C^\infty_c(B_R(x_0))$ and $\mathcal{H}^\ell_c(\Omega \cap B_R(x_0))$ the set of $\mathcal{H}^\ell(\Omega \cap B_R(x_0))$ functions which vanish identically in a neigbourhood of $\overline{\Omega} \cap \partial B_R(x_0)$.

Theorem 1′. *If $\partial\Omega \cap B_R(x_0)$ is $C^{m,\mu}$ in the sense that the above definition holds with $k = m$, and if $u \in \mathcal{H}^\ell(\Omega \cap B_R(x_0)) \cap C^{m,\mu}(\overline{\Omega} \cap B_R(x_0))$ is a weak solution of the equation (‡), where the ellipticity condition (E)′ holds and where $|a_{\alpha\beta}|_{0,\mu,\Omega\cap B_R(x_0)} \le \Lambda$, then*

$$|u|_{m,\mu,\Omega\cap B_{\theta R}(x_0)} \le C\left(|u|_{0,B_R(x_0)} + \textstyle\sum_{|\beta|\le m} |f_\beta|_{0,\mu,\Omega\cap B_R(x_0)}\right)$$

for each $\theta \in (0,1)$, where C depends on $n, m, R, \theta, \gamma, \mu, \Lambda$.

For the analogue of Theorem 2 (i.e., the boundary estimates in the non-divergence case) we need to discuss boundary conditions; we assume that either the Dirichlet boundary conditions $D^\gamma u|_{\partial\Omega} = 0$, $|\gamma| \le m-1$, or, more generally that

$$(**) \qquad \mathcal{B}u = 0 \text{ on } \partial\Omega,$$

where $\mathcal{B}u = 0$ is an abbeviation for $D^\gamma u|\partial\Omega = 0$, $|\gamma| \le \ell - 1$, and $B_j u|\partial\Omega = 0$, $j = 1, \ldots, 2m - \ell - 1$, where at the point $y \in \partial\Omega$, B_j are obtained as the natural boundary operators associated (as in Lecture 9) with the constant coefficients $a_{\alpha\beta}(y)$. Thus for example, in the case $m = 1$, the discussion here includes the general "oblique derivative" boundary condition discussed in Lecture 9.

Theorem 2′. *If $\partial\Omega \cap B_R(x_0)$ is $C^{2m,\mu}$ in the sense that the above definition holds with $k = 2m$, and $u \in C^{2m,\mu}(\overline{\Omega} \cap B_R(x_0))$ is a solution of theequation (‡‡), where the ellipticity condition (E)′ holds, where the boundary condition is either Dirichlet or one of the set of boundary conditions as in (∗∗) above, and where $|a_{\alpha\beta}|_{0,\mu,\Omega\cap B_R(x_0)} \le \Lambda$, then*

$$|u|_{2m,\mu,\Omega\cap B_{\theta R}(x_0)} \le C\left(|u|_{0,\Omega\cap B_R(x_0)} + |f|_{0,\mu,\Omega\cap B_R(x_0)}\right)$$

for each $\theta \in (0,1)$, where C depends on $n, m, R, \theta, \gamma, \mu, \Lambda$.

The proofs are left as an exercise, being almost identical to the proofs of Theorems 1, 2 after making the coordinate transformation which flattens the boundary mentioned above—of course at the key step in the proof we use Lemma 2 in place of Lemma 1.

Notice that by combining the above local and global estimates we obtain global schauder estimates for the norms $|u|_{m,\mu,\Omega}$ and $|u|_{2m,\mu,\Omega}$. For example for the "non-divergence form" equation

$$(*) \qquad \textstyle\sum_{|\alpha|,|\beta|\le m} a_{\alpha\beta} D^{\alpha+\beta} u = f$$

we have the following theorem:

Theorem 3 (Schauder's Global Estimate.) *Suppose $u \in C^{2m,\mu}(\overline{\Omega})$ and Ω is a bounded domain with $C^{2m,\mu}$ boundary, and suppose that (∗) holds toether with the ellipticity condition (E)′, and suppose that the boundary conditions are either Dirichlet $D^\gamma u|\partial\Omega = 0$, $|\gamma| \le m-1$ or (more generally) the "partially free" boundary conditions as in Theorem 2′. Then*

$$|u|_{2m,\mu,\Omega} \le C\left(|u|_{0,\Omega} + |f|_{0,\mu,\Omega}\right),$$

where $C = C(\mu, \lambda, \Lambda, \Omega)$.

Proof: Since $\partial\Omega$ is compact, there are fixed constants $0 < R < d$ and $\Gamma > 0$ such that the above definition of $C^{k,\mu}$ domain is satisfied with $k = 2m$ for each $x_0 \in \partial\Omega$, with functions $\psi : \mathbb{R}^{n-1} \to \mathbb{R}$ such that $|\psi|_{2m,\mu} \le \Gamma$. Then we can cover

$$S_{R/4} \equiv \{x \in \overline{\Omega} : \operatorname{dist}(x, \partial\Omega) \le R/4\}$$

by balls $\{B_{R/2}(y_j)\}_{j=1,\ldots,N}$ $(N \le c(n)(d/R)^n)$, where $y_j \in \partial\Omega$ for each $j = 1, \ldots, N$, and on each ball $B_R(y_j)$ we can apply Theorem 2′ above with $\theta = 1/2$.

The remaining volume

$$\Omega_{R/4} \equiv \{x \in \Omega : \operatorname{dist}(x, \partial\Omega) > R/4\}$$

can be covered by balls $\{B_{R/8}(z_k)\}_{k=1,\ldots,M}$ with $z_k \in \Omega_{R/4}$ for each k and with $M \le c(n)(d/R)^n$, and on each of the balls $B_{R/4}(z_k)$ we can apply Theorem 2 with $\theta = 1/2$.

Then summing $j = 1, \ldots, N$, $k = 1, \ldots, M$, we conclude the required estimate for $|u|_{2m,\mu,\Omega}$; notice that we here have to use the fact that the Hölder coefficient $[D^{2m}u]_{\mu,\Omega}$ is bounded above by the sum of the Hölder coefficients of $[D^{2m}u]_{\mu,\Omega}$ over the regions $\Omega \cap B_{R/2}(y_j)$ and $B_{R/2}(z_k)$.

Remark: Notice that the constant C on the right of the above theorem actually depends only on $n, m, \Lambda, \gamma, R, \Gamma$, and $R^{-n}\mathrm{vol}\,(\Omega)$ (where Γ, R are as in the above proof), because the number $N + M$ of balls needed in the above argument can evidently be bounded above by $CR^{-n}\mathrm{vol}\,(\Omega)$, with C depending only on n.

Using the global estimate of the above corollary together with the general existence results for boundary-value problems developed in Lectures 8, 9, we are going to develop some basic existence and regularity results in the Schauder setting. From now on we use the notation

$$Lu \equiv \textstyle\sum_{|\alpha|,|\beta|\le m} a_{\alpha\beta} D^{\alpha+\beta} u,$$

where the ellipticity (E)$'$ is assumed on $\overline{\Omega}$, and $\mathcal{B}u = 0$ will abbreviate the boundary conditions; thus $\mathcal{B}u = 0$ should be taken to mean either the Dirichlet boundary condition $D^\gamma u | \partial\Omega = 0,\ |\gamma| \le m-1$, or any of the "partially free" boundary conditions of Theorem 2$'$. (Of course there are analogous existence results for the divergence form operator $\sum D^\beta (a_{\alpha\beta} D^\alpha u)$.)

We here continue to assume that the coefficients $a_{\alpha\beta}$ satisfy $|a_{\alpha\beta}|_{0,\mu,\Omega} \le \Lambda$.

Theorem 4. *If the problem*

$$\begin{cases} Lu = 0 \text{ on } \Omega \\ \mathcal{B}u = 0 \text{ on } \partial\Omega \end{cases}$$

has no non-trivial $C^{2m,\mu}(\overline{\Omega})$ solutions, then for each $f \in C^{0,\mu}(\overline{\Omega})$ the problem

$$\begin{cases} Lu = f \text{ on } \Omega \\ \mathcal{B}u = 0 \text{ on } \partial\Omega \end{cases}$$

is solvable with $u \in C^{2m,\mu}(\overline{\Omega})$, and

$$|u|_{2m,\mu,\Omega} \le C\,|f|_{0,\mu,\Omega},$$

with $C = C(n, \mu, \Omega)$.

Before we begin the proof of Theorem 4, we need a definition which relates to approximation of $C^{2m,\mu}$ domains:

Definition: We say a sequence of $C^{2m,\mu}$ domains $\Omega_k \to \Omega$ in the $C^{2m,\mu}$-sense if

(i) there exists a sequence $\varepsilon_k \downarrow 0$ with $\{x \in \Omega : \mathrm{dist}\,(x, \partial\Omega) < \varepsilon_k\} \subset \Omega_k \subset \{x \in \mathbb{R}^n : \mathrm{dist}\,(x, \overline{\Omega}) < \varepsilon_k\}$, and

(ii) there exists a $\rho > 0$ such that $\forall x_0 \in \partial\Omega$ we can choose a coordinate axes with origin at x_0 and with $e_n =$ the inward-pointing unit normal for $\partial\Omega$ at x_0, such that

$$(*) \qquad \begin{cases} \Omega \cap B_\rho(0) = \{(y,t) : t > \psi(y)\} \cap B_\rho(0) \\ \Omega_k \cap B_\rho(0) = \{(y,t) : t > \psi_k(y)\}, \end{cases}$$

where $\psi, \psi_k \in C^{2m,\mu}(\mathbb{R}^{n-1})$ with $|\psi - \psi_k|_{2m,\mu,\mathbb{R}^{n-1}} \to 0$.

Preparatory to the proof of Theorem 4, we note that it is possible to find a sequence Ω_k of C^∞ domains with $\overline{\Omega}_k \subset \Omega$ for each k, and with $\Omega_k \to \Omega$ in the $C^{2m,\mu}$ sense of the above definition. Also, by mollification, we can find sequences $a^k_{\alpha\beta}$ in $C^{0,\mu}(\overline{\Omega}_k)$ with $C^{0,\mu}(\overline{\Omega}_k)$ norms bounded independent of k, converging to $a_{\alpha\beta}$ uniformly on compact subsets ofΩ.

Now, with $L_k u \equiv \sum_{|\alpha|,|\beta|\le m} a^k_{\alpha\beta} D^{\alpha+\beta} u$, we consider the corresponding boundary-value problem

$$(**) \qquad \begin{cases} L_k u = g \text{ on } \Omega_k \\ \mathcal{B}_k u = 0 \text{ on } \partial\Omega_k, \end{cases}$$

where g is any given $C^{0,\mu}(\overline{\Omega}_k)$ function and where $\mathcal{B}_k$ are the boundary operators corresponding to $\mathcal{B}$ in case $a_{\alpha\beta}$ is replaced by $a^k_{\alpha\beta}$.

We also need a lemma, as follows:

Lemma. *If the hypotheses are as in Theorem 4 then there is a constant $C > 0$, such that*

$$|u|_{0,\Omega} \le C|Lu|_{0,\mu,\Omega} \quad \forall u \in C^{2m,\mu}(\overline{\Omega})$$

with $\mathcal{B}u = 0$ on $\partial\Omega$.

If the sequences Ω_k, L_k are as described above, then C can be chosen so that for each k sufficiently large

$$|u|_{0,\Omega_k} \le C|L_k u|_{0,\mu,\Omega} \quad \forall u \in C^{2m,\mu}(\overline{\Omega}_k)$$

with $\mathcal{B}_k u = 0$ on $\partial\Omega_k$.

Proof: Notice that the second part of the lemma formally includes the first, so we need only prove the second part.

Let $u_k \in C^{2m,\mu}(\overline{\Omega}_k)$ with $\mathcal{B}_k u_k = 0$ on $\partial\Omega_k$ be arbitrary. By Theorem 3 (and the remark following its proof) we have C independent of k such that

$$(1) \qquad |u_k|_{2m,\mu,\Omega_k} \le C\left(|u_k|_{0,\Omega_k} + |L_k u_k|_{0,\mu,\Omega_k}\right)$$

for each k. Let $\lambda_k = |u_k|_{0,\Omega_k}$, and suppose, contrary to the lemma, that

$$\lambda_k^{-1}|L_k u_k|_{0,\alpha,\Omega_k} \to 0.$$

This gives

$$|v_k|_{2m,\mu,\Omega_k} \le C$$

where C is independent of k and where $v_k = \lambda_k^{-1}u_k$. By virtue of the Arzela-Ascoli lemma we then have a subsequence $v_{k'}$ converging uniformly on compact subsets of Ω to $v \in C^{2m,\mu}(\overline{\Omega})$ with $\mathcal{B}v = 0$ on $\partial\Omega$, $|v|_{0,\Omega} = 1$, and $Lv = 0$ on Ω, contrary to the hypotheses.

Thus it is impossible that $\lambda_k^{-1}|Lu_k|_{0,\alpha,\Omega_k} \to 0$, and hence there exists C independent of u and k such that $|u|_{0,\Omega_k} \le C|L_k u|_{0,\alpha,\Omega_k}$, and this, in view of (1), gives the required inequality.

Proof of Theorem 4: We let Ω_k, $a^k_{\alpha\beta}$, be as above. Since $a^k_{\alpha\beta}$ are in $C^\infty(\overline{\Omega}_k)$ (and since the operator L_k can be written in the divergence-form) we can use the general existence and regularity discussion of Lectures 8, 9 to prove that <u>either</u>

Case (a): $(**)$ has a non-trivial solution $u_k \in C^\infty(\overline{\Omega}_k)$, <u>or</u> **Case (b):** $(**)$ has a unique $C^{2m,\mu}(\overline{\Omega}_k)$ solution for every $g \in C^{0,\mu}(\overline{\Omega}_k)$.

In view of the above lemma we know that Case (a) can occur at most for finitely many values of k, and that if u_k is the unique solution of $(**)$ with $g = f_k$, then

$$|u_k|_{0,\Omega_k} \le C|f_k|_{0,\mu,\Omega_k} \le C',$$

where C, C' are independent of k. Then by Theorem 3 (and the remark following it), we have

$$|u_k|_{2m,\mu,\Omega_k} \le C,$$

where C is independent of k.

Then by the Arzela-Ascoli lemma, there is there is a subsequence $u_{k'}$ and $u \in C^{2m,\mu}(\overline{\Omega})$ with $\mathcal{B}u = 0$ on $\partial\Omega$ and $u_k \to u$ locally with respect to the C^{2m}-norm on Ω; it of course follows that $Lu = f$ on Ω. u is also unique by Lemma 1, so this completes the proof of Theorem 4.

LECTURE 12 PROBLEMS

12.1 Show that Lemma 1 is valid with the hypotheses that $u \in L^1(\mathbb{R}^n)$ in place of the hypothesis $[D^k u]_{\mu,\mathbb{R}^n} < \infty$, provided u has compact support. Hint: Use mollification.

12.2 By using appropriate modifications of the proof of Theorem 1 given above, give the proof of Theorem 2.

12.3 (a) If $K(x,y) = \Gamma(|x-y|)$ is the kernel of Lecture 4, if $f \in C^{0,\mu}_c(\mathbb{R}^n)$, and if $w(x) = \int_{\mathbb{R}^n} K(x-y)f(y)\,dy$, prove that $w \in C^{2,\mu}(\mathbb{R}^n)$, $\Delta w = f$, and

$$[D^2 w]_{\mu,\mathbb{R}^n} \le c[f]_{\mu,\mathbb{R}^n}.$$

Hint: Recall that we already know from Lecture 4 that $\Delta w = f$ in case f is smooth, so we can apply Lemma 1 in case f is smooth. (So apply Lemma 1 with f_σ in place of f.)

(b) If $f_\ell \in C^{0,\mu}_c(\mathbb{R}^n)$ for $\ell = 1,\ldots,n$, and if

$$w(x) = \int_{\mathbb{R}^n} \textstyle\sum_{\ell=1}^n (D_{x^\ell}K)(x,y) f_\ell(y)\,dy,$$

prove that $w \in C^{1,\mu}(\mathbb{R}^n)$, $\Delta w = \sum_{\ell=1}^n D_\ell f_\ell$ (weakly), and

$$[Dw]_{\mu,\mathbb{R}^n} \le C\textstyle\sum_{\ell=1}^n [f_\ell]_{\mu,\mathbb{R}^n}.$$

12.4 Let $u \in C^{1,\beta}(\overline{B}_R(x_0))$, where $\beta \in (0,1]$. Using the inequality

$$(*) \qquad |D_j u(y) - D_j u(x)| \le |y-x|^\beta [D_j u]_{\beta,B_R(x_0)}$$

whenever $x, y \in B_R(x_0)$, prove that for each $\sigma \in (0,R]$ there exists a ball $B_\sigma(y) \subset B_R(x_0)$ such that, for each $j = 1,\ldots,n$,

$$|D_j u|_{0,B_R(x_0)} \le \inf_{B_\sigma(y)} |D_j u| + (2\sigma)^\beta [D_j u]_{\beta,B_R(x_0)},$$

and hence prove $|Du|_{0,B_R(x_0)} \le n\sigma^{-1}|u|_{0,B_R(x_0)} + (2\sigma)^\beta [Du]_{\beta,B_R(x_0)}$ for each $\sigma \in (0,R]$. Here $|Du|_0 = \sum_j |D_j u|_0$ and $[Du]_\mu = \sum_j [D_j u]_\mu$.

12.5 Using the result of 12.4, prove the interpolation inequalities

$$(\dagger) \qquad R^\ell |D^\ell u|_{0,B_R(x_0)} \le \varepsilon R^{k+\mu}[D^k u]_{\mu,B_R(x_0)} + C|u|_{0,B_R(x_0)}, \quad C = C(\varepsilon,\mu,k,n),$$

for each $\varepsilon \in (0,1]$, $\mu \in (0,1]$, and $1 \le \ell \le k$.

Hint: First use the above inequality with $\beta = 1$ in order to establish that $|D^k u|_0 \le C(\sigma^{-k}|u|_0 + \sigma^{m-k}|D^m u|_0)$ (where $C = C(n,m)$), and then use the inequality with $D^m u$ in place of u and with $\beta = \mu$. Note: $[f]_\beta = |Df|_0$ in case $\beta = 1$ and $f \in C^1$.

12.6 If $[f]_{\mu,\mathbb{R}^n} < \infty$, and if f is supported in the ball $B_R(0)$, prove that for each $j \in \{1, 2, \ldots\}$ we have $f = D^j_{x^1} g_j$ for some g_j satisfying $[g_j]_\mu \le CR^j(R^{-\mu}|f|_0 + [f]_\mu)$.
Hint: If $j = 1$ we can define $g(x) = \int_0^{x^1} f(s, x^2, \ldots, x^n)\, ds$.

12.7 (a) For each multi-index β prove there are constants $c_{\beta\gamma\delta}$, $\gamma + \delta = \beta$, $|\delta| < |\beta|$, such that $\varphi D^\beta f = D^\beta(\varphi f) + \sum_{\gamma+\delta=\beta, |\delta|<|\beta|} D^\delta(c_{\beta\gamma\delta} D^\gamma \varphi\, f)$ for each pair φ, f of C^∞ functions on an open ball B

Hint: Induction on $|\beta|$.

(b) Show the identity of (a) continues to hold in the appropriate weak sense on B if $\varphi \in C^\infty(B)$ but f is merely in $L^1(B)$.

12.8 (Scaling.) If the hypotheses of Theorem 1 are stated in scale invariant form
$$R^{2m-|\alpha|-|\beta|}(|a_{\alpha\beta}|_{0,B_R} + R^\mu [a_{\alpha\beta}]_{\mu,B_R}) \le \Lambda$$
for $|\alpha|, |\beta| \le m$, prove that the conclusion can be written in the form
$$\sum_{j=1}^m R^j |D^j u|_{0,B_{\theta R}} + R^{m+\mu}[D^m u]_{\mu,B_{\theta R}} \le C\left(|u|_{0,B_R} + \sum_{|\beta|\le m} R^{2m-|\beta|}(|f_\beta|_{0,B_R} + R^\mu [f_\beta]_{\mu,B_R})\right)$$
with C independent of R.

Supplement to Lecture 12
Smoothness up to the boundary, and the Fredholm Alternative in $C^{2m,\mu}(\overline{\Omega})$

Our first aim here is to prove a result concerning smoothness of solutions up to the boundary; the basic result in this direction is stated in Lemma 1 below.

To begin with we assume we have an operator

(1) $$Lu = \sum_{|\alpha|,|\beta|\le m} D^\beta(a_{\alpha\beta} D^\alpha u) = f \text{ on } C_R^+,$$

where as usual
$$C_R = B_R^{n-1} \times (-R, R) \text{ and } C_R^+ = B_R^{n-1} \times (0, R).$$

Points in C_R^+ will be denoted (y, t) with $y = (y_1, \ldots, y_{n-1})$ and $t \in (0, R)$, and, for $j = 1, \ldots, n-1$ we let $\delta_{h,j}$ be the difference quotient operators with respect to the y variables:

(2) $$\delta_{h,j} u(y,t) = h^{-1} u(y + he_j, t) - u(y,t)), \quad h \ne 0.$$

$\mathcal{H}$ is a subspace of $H^m(C_R^+)$ with the property that, for any ζ which is the restriction to C_R^+ of a $C_c^\infty(C_R)$ function,

(3) $$\varphi \in \mathcal{H} \Rightarrow \zeta\varphi \text{ and } \zeta\delta_{h,j}\varphi \in \mathcal{H} \text{ for } j = 1, \ldots, n-1 \text{ and } 0 < |h| < \delta_0,$$

for suitable $\delta_0 = \delta_0(\zeta, \varphi) > 0$, and we assume the coercivity condition

(C) $$\int_{C_R^+} \sum_{|\alpha|=|\beta|=m} a_{\alpha\beta} D^\alpha \varphi D^\beta \varphi \ge \mu \int_{C_R^+} \sum_{|\alpha|=m} |D^\alpha \varphi|^2 - C\|\varphi\|^2_{0,C_R^+}, \quad \varphi \in \mathcal{H}.$$

Recall that this implies the ellipticity condition

(E)′ $$\sum_{|\alpha|=|\beta|=m} a_{\alpha\beta} \xi^\alpha \xi^\beta \ge C\mu|\xi|^{2m}, \quad \xi \in \mathbb{R}^n,$$

where $C = C(n, m)$, and, since we assume continuity (in fact smoothness) of $a_{\alpha\beta}$, it is implied by (E)′ if e.g. $\mathcal{H} = H_0^m(C_R^+)$.

Our aim here is to prove the following boundary smoothness result for weak solutions. Notice that more precise results were proved in Lecture 9, but the discussion here is independent of those results.

Lemma 1. *Assume that $a_{\alpha\beta}$, f are $C^\infty(\overline{C}_R^+)$ and that $u \in \mathcal{H}$ is a weak solution of $Lu = f$ in C_R^+. Then $u \in C^\infty(\overline{C}_{\theta R}^+)$ for each $\theta \in (0, 1)$.*

Proof: First observe that by the theory developed in Lecture 6 we have $u \in C^\infty(C_{\theta R}^+) \cap \mathcal{H}$ for each $\theta \in (0, 1)$. Furthermore, in view of coercivity (5) above, we can repeat exactly the difference quotient arguments of Lecture 6, with respect to the tangential difference quotient operators $\delta_{h,j}$, $j = 1, \ldots, n-1$, as in (2), to establish the bounds
$$\sum_{|\gamma|=N} \|D_y^\gamma u\|_{m,C_{\theta R}^+} \le C(\|u\|_{0,C_R^+} + \|\sum_{|\beta|\le N} D_y^\beta f\|_{0,C_R^+}),$$
where $C = C(n, N, m, \mu, \theta)$, so (since $f \in C^\infty(\overline{C}_R^+)$ and $\|u\|_{0,C_R^+} < \infty$) we have
$$\sum_{|\gamma|=N} \|D_y^\gamma u\|_{m,C_{\theta R}^+} < \infty \quad N = 0, 1, 2, \ldots,$$
and so

(5) $$\sum_{|\gamma|\le N} \|D_y^\gamma D_t^j u\|_{0,C_{\theta R}^+} < \infty, \quad j \in \{0, \ldots, m\},\ N = 0, 1, 2, \ldots.$$

We use the notation $u(t) : B_R^{n-1} \to \mathbb{R}$ according to $u(t)(y) = u(y,t)$ and observe that since $u \in C^\infty(C_{\theta R}^+)$ we can differentiate under the integral to compute, for $t \in (0, R/2)$,

$$D_t \int_{B_{\theta R}^{n-1}} (D_y^\gamma D_t^j u(y,t))^2\, dy = 2\int_{B_{\theta R}^{n-1}} (D_y^\gamma D_t^{j+1} u(y,t))(D_y^\gamma D_t^j u(y,t))\, dy,$$

and hence by using Cauchy-Schwarz on the right we have by (5) that

$$\int_0^{R/2} \Big| D_t \int_{B_{\theta R}^{n-1}} (D_y^\gamma D_t^j u(y,t))^2\, dy \Big| dt < \infty, \quad j \in \{0,\ldots,m-1\},\ \forall \gamma.$$

Thus by the general inequality $\sup |f(t)| \le |f(R/2)| + \int_0^{R/2} |f'(t)|\, dt$ we in fact have

$$\sup_{t\in(0,R/2)} \int_{B_{\theta R}^{n-1}} (D_y^\gamma D_t^j u(y,t))^2\, dy < \infty, \quad j \in \{0,\ldots,m-1\},\ \forall \gamma.$$

Thus in particular we can use the Sobolev embedding theorem with respect to the y variables (for any fixed $t \in (0, R/2)$) to deduce that

(6)
$$\sup_{t\in(0,R/2),|y|\le\theta R} \textstyle\sum_{|\gamma|\le N} |D_y^\gamma D_t^j u| < \infty, \quad j \in \{0,\ldots,m-1\},\ N = 0,1,2,\ldots.$$

Next we observe, since the $a_{\alpha\beta} \in C^\infty(\overline{C}_{\theta R}^+)$, and since the ellipticity

$$\textstyle\sum_{|\alpha|=|\beta|=m} a_{\alpha\beta}\xi^\alpha\xi^\beta \ge \mu|\xi|^{2m}$$

implies that the coefficient a_{me_n,me_n} (i.e. $a_{\alpha\beta}$ when $\alpha = \beta = (0,\ldots,0,m)$) is bounded from below by μ, we can write the equation (1) in the form

$$D_t^m\big(D_t^m + \textstyle\sum_{j=0}^{m-1} D_t^j (\sum_{|\gamma|\le 2m-j} b_{j\gamma} D_y^\gamma u)\big) = \\ \textstyle\sum_{j=0}^{m-1} D_t^j (\sum_{|\gamma|\le 2m-j} c_{j,\gamma} D_y^\gamma u) + \widetilde{f},$$

where $\widetilde{f}, b_{j,\gamma}, c_{j,\gamma}$ are all in $C^\infty(\overline{C}_{\theta R}^+)$. By applying the operator $\mathcal{I}_t$ defined by

$$(\mathcal{I}_t(v))(y) = -\int_t^{R/2} v(y,s)\, ds$$

for any continuous $v = v(y,t)$ on C_R^+, we see that then

$$\big(D_t^m + \textstyle\sum_{j=0}^{m-1} D_t^j (\sum_{|\gamma|\le 2m-j} b_{j\gamma} D_y^\gamma u)\big) = \\ \mathcal{I}_t^m \textstyle\sum_{j=0}^{m-1} D_t^j (\sum_{|\gamma|\le 2m-j} c_{j,\gamma} D_y^\gamma u) + \mathcal{I}_t^m \widetilde{f} + \sum_{j=0}^m d_j(y) t^j,$$

where each $d_j(y)$ is a $C^\infty(B_R^{n-1})$ function. Then by applying $D_y^\delta D_t^\ell$ to each side of the identity and rearranging terms we have

(7)
$$D_t^{m+\ell} D_y^\delta u = \textstyle\sum_{j=0}^{m-1} D_t^{j+\ell} (\sum_{|\gamma|\le 2m-j+|\delta|} \widetilde{b}_{j\gamma} D_y^\gamma u) \\ + \mathcal{D}_t^{(m,\ell)} \textstyle\sum_{j=0}^{m-1} D_t^j (\sum_{|\gamma|\le 2m-j+|\delta|} \widetilde{c}_{j,\gamma} D_y^\gamma u) + \mathcal{I}_t^m \widetilde{f} + D_y^\delta D_t^\ell \sum_{j=0}^m d_j(y) t^j,$$

where $\widetilde{b}_{j\gamma}, \widetilde{c}_{j\gamma} \in C^\infty(\overline{C}_R^+)$, $\mathcal{D}_t^{(m,\ell)} = \mathcal{I}_t^{m-\ell}$ if $\ell < m$ and $= D_t^{\ell-m}$ if $\ell \ge m$. In view of (6), (7) we can now use induction on ℓ, starting at $\ell = 0$, to conclude that

$$\sup_{t\in(0,R/2),|y|\le\theta R} |D_t^j D_y^\delta u| < \infty, \quad \forall\, j \ge 0,\ \delta = (\delta_1,\ldots,\delta_{n-1}),$$

so $u \in C^\infty(\overline{C}_{\theta R}^+)$ for each $\theta \in (0,1)$ as claimed.

From now on Ω will be a bounded $C^{2m,\mu}$ domain and for $u \in C^{2m,\mu}(\overline{\Omega})$ we write

$$Lu = \textstyle\sum_{|\alpha|,|\beta|\le m} a_{\alpha\beta} D^{\alpha+\beta} u.$$

We note that it is possible to find a sequence Ω_k of C^∞ domains with $\overline{\Omega}_k \subset \Omega$ for each k, and with $\Omega_k \to \Omega$ in the $C^{2m,\mu}$ sense of Lecture 12. Also, by mollification, we can find sequences $a_{\alpha\beta}^k$ in $C^{0,\mu}(\overline{\Omega}_k)$ with $C^{0,\mu}(\overline{\Omega}_k)$ norms bounded independent of k, and converging to $a_{\alpha\beta}$ uniformly on compact subsets of Ω.

We now wish to establish an important Fredholm Alternative result for $C^{2m,\mu}$ solutions of $Lu = f$ as follows:

(FA)
$$\underline{\text{either}}\begin{cases} \exists\, u \in C^{2m,\mu}(\overline{\Omega}) \setminus \{0\} \text{ with} \\ Lu = 0 \text{ in } \Omega \\ D^\gamma u = 0 \text{ on } \partial\Omega \text{ for } |\gamma| \le m-1 \\ |u|_{2m,\mu,\Omega} \le C\, |u|_{0,\Omega} \end{cases} \underline{\text{or}}\begin{cases} \forall f \in C^{0,\mu}(\overline{\Omega}),\ \exists\, u \in C^{2m,\mu}(\overline{\Omega}): \\ Lu = f \text{ in } \Omega \\ D^\gamma u = 0 \text{ on } \partial\Omega \text{ for } |\gamma| \le m-1 \\ |u|_{2m,\mu,\Omega} \le C\, |f|_{0,\mu,\Omega}. \end{cases}$$

Here $C = C(n, m, \mu, \{a_{\alpha\beta}\}, \Omega)$.

To prove this, let $\Omega_k, a_{\alpha\beta}^k$ be as above, so that Ω_k are C^∞ domains with $\Omega_k \to \Omega$ in the $C^{2m,\mu}$ sense, and $a_{\alpha\beta}^k$ are $C^\infty(\overline{\Omega}_k)$ functions with $a_{\alpha\beta}^k \to a_{\alpha\beta}$ in the $C^{0,\mu}$ sense. Likewise for given $f \in C^{0,\mu}(\overline{\Omega})$, let $f^k \in C^{0,\mu}(\overline{\Omega}_k)$ with $f^k \to f$ in the $C^{0,\mu}$ sense.

For $u \in C^{2m,\mu}(\overline{\Omega}_k)$, let $L_k u = \sum_{|\alpha|,|\beta|\le m} a_{\alpha\beta}^k D^{\alpha+\beta} u$. Since we can write L_k in the divergence form $L_k u = \sum_{|\alpha|,|\beta|\le m} D^\beta(\widetilde{a}_{\alpha\beta}^k D^\alpha u)$, with $\widetilde{a}_{\alpha\beta}^k \in C^\infty(\overline{\Omega}_k)$

and $\tilde{a}^k_{\alpha\beta} = a^k_{\alpha\beta}$ for $|\alpha| = |\beta| = m$, we can apply the general spectral theory established in Lecture 7: so for each k there is a discrete spectrum Λ_k such that either (a) $0 \in \Lambda_k$ and there is a weak $H^m_0(\Omega_k)$ solution $u_k \neq 0$ of $L_k u_k = 0$, or else (b) there is a weak $H^m_0(\Omega_k)$ solution u_k of $L_k u_k = f^k$. In either case (a) or case (b) we can apply Lemma 1 above to establish that $u_k \in C^\infty(\overline{\Omega}_k)$ and (since $u_k \in H^m_0(\Omega_k)$) we also have $D^\gamma u_k|_{\partial\Omega_k} = 0$ for $|\gamma| \le m-1$. By the Schauder estimates of Lecture 12 we also have

$$(*) \qquad |u_k|_{2m,\mu,\Omega_k} \le C|u_k|_{0,\Omega_k} \text{ in case (a)}$$

$$(**) \qquad |u_k|_{2m,\mu,\Omega_k} \le C(|u_k|_{0,\Omega_k} + |f_k|_{0,\mu,\Omega_k}) \text{ in case (b),}$$

where C is a constant which does not depend on k. If case (a) holds for infinitely many k we take $\lambda_k = |u_k|_{0,\Omega_k}$ and $\tilde{u}_k = \lambda_k^{-1} u_k$, so that by $(*)$ $|u_k|_{2m,\mu,\Omega_k} \le C$ for some subsequence of u_k and the Arzela-Ascoli lemma implies there is a subsequence u_{k_j} and a $C^{2m,\mu}(\overline{\Omega})$ function u with $u_{k_j} \to u$ in the C^{2m} sense, with $Lu = 0$, $|u|_{0,\Omega} = 1$, $D^\gamma u|_{\partial\Omega} = 0$ for $|\gamma| \le m-1$, and $|u|_{2m,\mu,\Omega} \le C$, which is the first alternative of (FA). Next we consider the possibility that (b) occurs with $|u_k|_{0,\Omega_k} \ge C_k |f^k|_{0,\mu,\Omega_k}$, where $C_k \to \infty$. Thus we can normalize as in case (a) to give $\tilde{u}_k = \lambda_k^{-1} u_k$ and $L_k \tilde{u}_k = \tilde{f}_k$ with $|\tilde{f}_k|_{0,\mu} \to 0$, and so the argument of case (a) can again be applied to give $Lu = 0$, $|u|_{0,\Omega} = 1$, $D^\gamma u|_{\partial\Omega} = 0$. (i.e. again the first alternative of (FA) is established.) The final possibility is that (b) occurs with $|u_k|_{0,\Omega_k} \le C|f^k|_{0,\mu,\Omega_k}$ with fixed C independent of k, and in this case we simply use the Schauder estimate $(**)$ to conclude $|u_k|_{2m,\mu,\Omega_k} \le C|f^k|_{0,\mu,\Omega_k} \to C|f|_{0,\mu,\Omega}$ and so Arzela-Ascoli can again be applied to give $u_{k_j} \to u$ in the C^{2m} sense, where $u \in C^{2m,\mu}(\overline{\Omega})$ with $Lu = f$ and $D^\gamma u|_{\partial\Omega} = 0$, $|\gamma| \le m-1$.

This completes the proof of (FA) except for the final inequality in the second alternative; we have not yet established that, since the final step of the above argument only shows that the u satisfies $|u|_{2m,\mu,\Omega} \le C|f|_{0,\mu,\Omega}$ with a constant that possibly depends on f. To show that we can establish (FA) with a fixed constant C in the second alternative, we assume that the first alternative of (FA) fails. Then the argument above shows that for each $f \in C^{0,\mu}(\overline{\Omega})$ we must have $u \in C^{2m,\mu}(\overline{\Omega})$ with $Lu = f$, $D^\gamma u|_{\partial\Omega} = 0$, $|\gamma| \le m-1$ and by the Schauder estimates

$$(8) \qquad |u|_{2m,\mu,\Omega} \le C(|u|_{0,\Omega} + |f|_{0,\mu,\Omega}), \quad C = C(n, m, \Omega, \mu, \Lambda).$$

Now we simply consider the possibility that there is a sequence $f^k \in C^{0,\mu}(\overline{\Omega})$ and $u_k \in C^{2m,\mu}(\overline{\Omega})$ with $Lu_k = f^k$ and with $|u_k|_{0,\Omega} \ge C_k |f^k|_{0,\mu,\Omega}$, where

$C_k \to \infty$. Then with $\tilde{u}_k = \lambda_k^{-1} u_k$ ($\lambda_k = |u_k|_{0,\Omega}$) we get $L\tilde{u}_k = \tilde{f}_k$ with $|\tilde{f}_k|_{0,\mu,\Omega} \to 0$ and $|\tilde{u}_k|_{2m,\mu,\Omega} \le C$ by (8), so $\tilde{u}_k \to u$ in $C^{2m}(\overline{\Omega})$ with $u \in C^{2m,\mu}(\overline{\Omega})$ and with the first alternative in (FA) holding, a contradiction. Thus $|u|_{0,\Omega} \le C|f|_{0,\mu,\Omega}$ for every $u \in C^{2m,\mu}(\overline{\Omega})$ with $Lu = f$ and $D^\gamma u|_{\partial\Omega} = 0$, $|\gamma| \le m-1$, and using this in (8) we see the inequality in the second alternative of (FA) does indeed hold.

Lecture 13

Maximum Principles for Second Order Equations

The simplest version of the maximum principle is the following weak maximum principle, which applies to operators of the form

$$Lu = \sum_{i,j=1}^{n} a_{ij} D_i D_j u + \sum_{j=1}^{n} b_j D_j u + cu,$$

where

(‡) $\qquad a_{ij},\ b_j,\ c$ are bounded functions on Ω, $\quad a_{ij} = a_{ji}$,

and where we assume that there is a constant $\mu > 0$ such that

(‡‡) $$\sum_{i,j=1}^{n} a_{ij}(x)\xi_i \xi_j \geq \mu|\xi|^2$$

for all $x \in \Omega$, $\xi \in \mathbb{R}^n$.

Theorem 1 (Weak maximum principle.) *Suppose Ω is bounded and let $u \in C^2(\Omega) \cap C^0(\overline{\Omega})$ satisfy $Lu \geq 0$ in Ω, where (‡), (‡‡) above hold. Suppose also that*

* $$c \leq 0 \text{ in } \Omega.$$

Then

(**) $$\max_{\overline{\Omega}} u \leq \max_{\partial\Omega} u_+,$$

where u_+ is the positive part of u, defined by $u_+(x) = \max\{u(x), 0\}$.

In case $c \equiv 0$, the inequality $(**)$ holds with u in place of u_+.

Remarks: (1) Notice that this is specifically a second order result; for example in case $n = 1$, one sees that the operator d^4/dx^4 satisfies no such maximum principle. (For example, the function $u(x) \equiv 3x^2 - 4x^3$ satisfies $u^{(iv)}(x) \equiv 0$, but has a positive maximum over the interval [0,1] at the point $x = 1/2$.)

(2) If $Lu \le 0$, then we can apply the above theorem to $-u$, thus giving

$$\min_{\overline{\Omega}} u \ge -\max_{\partial\Omega} u_-,$$

where $u_-(x) \equiv \max\{-u(x), 0\}$ denotes the negative part of u.

(3) In case $Lu = 0$ in Ω we can combine the two inequalities (for u and $-u$) to get

$$\max_{\overline{\Omega}} |u| = \max_{\partial\Omega} |u|.$$

Proof of weak maximum principle: We see that $L(u + \varepsilon e^{\gamma x^1}) = Lu + \varepsilon(\gamma^2 a_{11} + \gamma b_1 + c)e^{\gamma x^1} > 0$ for sufficiently large γ (independent of ε) because b_j, c are bounded by (‡) and $a_{11} \ge \mu$ by the ellipticity condition (‡‡). On the other hand if $v \in C^2(\Omega)$ and $Lv > 0$ in Ω, then v cannot attain a positive maximum in Ω. Indeed if $x_0 \in \Omega$ and v attains a positive maximum at x_0, then $Dv(x_0) = 0$ and Hessian $v(x_0)$ is negative semi-definite, so since $c \le 0$ we would have $Lv(x_0) \le \sum_{i,j=1}^n a_{ij}(x_0) D_i D_j v(x_0) \le 0$, contradicting the fact that $Lv > 0$. Thus in particular $u + \varepsilon e^{\gamma x_1}$ cannot attain a positive maximum in Ω, so

$$\sup_{\Omega}(u + \varepsilon e^{\gamma x^1}) \le \max_{\partial\Omega}(u + \varepsilon e^{\gamma x^1})_+.$$

Letting $\varepsilon \downarrow 0$, we then have the required result.

If $c \equiv 0$, then $L(u + k) \equiv Lu$ for any constant k; in particular taking k so that $u + k > 0$ everywhere in $\overline{\Omega}$ and applying the result just proved to $u + k$, we conclude that $\max_{\overline{\Omega}} u = \max_{\partial\Omega} u$ as required.

Before going further we note an important consequence of the above weak maximum principle in case u satisfies $Lu = f$ on Ω, where f is a given function on Ω.

Theorem 2. *Suppose $Lu = f$, where f is a given bounded function on Ω, and suppose that $c \le 0$ in Ω and that (‡), (‡‡) hold. Then*

$$\max_{\overline{\Omega}} |u| \le \max_{\partial\Omega} |u| + \mu^{-1} e^{2(1+\beta)} d^2 \sup_{\Omega} |f|,$$

with β any upper bound for $\mu^{-1}(d|b_1| + d^2|c|)$, and with d any constant such that $\Omega \subset \{x = (x^1, \ldots, x^n) : |x^1| < d\}$. (In particular we may take $d = \operatorname{diam}\Omega$ if we wish, because we can always translate so that $0 \in \Omega$.)

Corollary 1. *If $\varepsilon \in (0,1)$ and the hypotheses are as in the above theorem, except that the hypothesis $c \le 0$ is replaced by the condition $\mu^{-1} e^{2(1+\beta)} d^2 \sup c_+ \le \varepsilon$, then*

$$\max_{\overline{\Omega}} |u| \le (1-\varepsilon)^{-1} \max_{\partial\Omega} |u| + \mu^{-1}(1-\varepsilon)^{-1} e^{2(1+\beta)} d^2 \sup |f|.$$

Proof of Corollary 1: First note that the equation $Lu = f$ can be written $L_1 u = \tilde{f}$, where $\tilde{f} = f - c_+ u$ and where L_1 is the same as L with $-c_-$ (≤ 0) in place of c. Hence by the theorem

$$\max_{\overline{\Omega}} |u| \le \max_{\partial\Omega} |u| + \mu^{-1} e^{2(1+\beta)} d^2 \big(\sup_{\Omega} |f| + \sup_{\Omega} c_+ \max_{\overline{\Omega}} |u|\big),$$

so that

$$(1-\varepsilon) \max_{\overline{\Omega}} |u| \le \max_{\partial\Omega} |u| + \mu^{-1} e^{2(1+\beta)} d^2 \sup_{\Omega} |f|,$$

which evidently gives the corollary as claimed.

Proof of Theorem 2: Define the function v by

$$v = u + \mu^{-1} d^2 \sup_\Omega |f|\, e^{\gamma(1+x^1/d)}$$

(where $\gamma \ge 1$ is a constant to be chosen). By direct computation we have

$$\begin{aligned} Lv &= Lu + \sup_\Omega |f|\, \mu^{-1}(a_{11}\gamma^2 + b_1 d\gamma + cd^2) e^{\gamma(1+x^1/d)} \\ &\ge f + \sup_\Omega |f| \mu^{-1}(\gamma^2\mu - \gamma d|b_1| - |c|d^2) e^{\gamma(1+x^1/d)} \\ &\ge f + \sup_\Omega |f| \gamma(\gamma - \beta) e^{\gamma(1+x^1/d)}, \end{aligned}$$

so that, if we choose for example $\gamma = 1+\beta$, then we have $Lv \ge 0$ and hence the weak maximum principle applies to v. Thus $\max_{\overline{\Omega}} u \le \max_{\overline{\Omega}} v \le \max_{\partial\Omega} v_+ \le \max_{\partial\Omega} u_+ + \mu^{-1} e^{2(1+\beta)} \sup |f|\, d^2$.

Our principal aim now is to prove the strong (or Hopf) maximum principle. For this we first need to establish the following Hopf boundary point lemma, which is of considerable importance in itself.

Lemma 1. *Suppose $B = B_\rho(y)$ is an open ball in $\mathbb{R}^n$, $u \in C^2(B)$, $Lu \ge 0$ in B, $x_0 \in \partial B$, u is continuous at x_0, $u(x_0) \ge 0$, and $u(x) < u(x_0)$ for each $x \in B$.*

Then, if D_η denotes directional derivative in the direction of the inward pointing unit normal η of ∂B, we have

$$D_\eta u(x_0) < 0,$$

if this derivative exists, and in any case

$$\limsup_{h\downarrow 0} \frac{u(x_0+h\eta)-u(x_0)}{h} < 0.$$

Furthermore, if $c \equiv 0$ the condition $u(x_0) \ge 0$ can be dropped, while the hypothesis $c \le 0$ can be dropped if $u(x_0) = 0$.

Proof: First note the last pair of results follow from the first part of the lemma by virtue of the facts that $L(u+k) \equiv L(u)$ (k any constant) if $c \equiv 0$, and

$$\begin{aligned} Lu &= \textstyle\sum_{i,j=1}^n a_{ij} D_i D_j u + \sum_{j=1}^n b_j D_j u + cu \\ &\le \textstyle\sum_{ij=1}^n a_{ij} D_i D_j u + \sum_{j=1}^n b_j D_j u - c_- u \end{aligned}$$

in case $u(x_0) = 0$ (because in this case $u < 0$ in B by virtue of the hypothesis $u < u(x_0)$).

Thus it remains to prove only the first part of the lemma. For this, notice that we may without loss of generality assume that u is continuous on $\overline{B}$, otherwise replace B by a ball $\widetilde{B}$ with closure contained in $B \cup \{x_0\}$ and with $x_0 \in \partial\widetilde{B}$. Now let $r = |x-y|$, and note that if $w = e^{-\alpha r^2} - e^{-\alpha\rho^2}$ with $\alpha > 0$ a constant to be chosen, then by direct computation

$$\begin{aligned} Lw = e^{-\alpha r^2}\big(&4\alpha^2 \textstyle\sum_{i,j=1}^n a_{ij}(x^i-y^i)(x^j-y^j) \\ &- 2\alpha(\textstyle\sum_{i=1}^n a_{ii} + \sum_{i=1}^n b_i(x^i-y^i))\big) + cw, \end{aligned}$$

and so by (‡) and the fact that $c \le 0$ we have

$$Lw \ge e^{-\alpha r^2}\big(4\alpha^2 r^2 + 2\alpha(\textstyle\sum_{i=1}^n a_{ii} - \sum_{i=1}^n |b_i| r) + c\big).$$

Thus for α large enough we have

$$Lw > 0 \text{ in } A, \;\; A = B_\rho(y) \setminus B_{\rho/2}(y).$$

Then $Lv = Lu - cu(x_0) + \varepsilon Lw > 0$ because $c \le 0$ and $u(x_0) \ge 0$. Also, since $w = 0$ on $\partial B_\rho(y)$ and since $u < u(x_0)$ on $\partial B_{\rho/2}(y)$, we can choose $\varepsilon > 0$ such that $v \equiv u - u(x_0) + \varepsilon w \le 0$ on ∂A, and so by the weak maximum principle we then have $v \le 0$ in A; i.e.

$$(*) \qquad u(x) - u(x_0) \le -\varepsilon w(x), \quad x \in A.$$

Thus finally

$$\limsup_{h\downarrow 0} \frac{u(x_0+h\eta)-u(x_0)}{h} \le -\varepsilon D_\eta w(x_0) < 0.$$

This completes the proof of the Hopf boundary point lemma.

Remark: Notice that the above proof is actually a special case of the "barrier" principle, as follows:

Lemma 2. *Suppose Ω is bounded, $u \in C^2(\Omega) \cap C^0(\overline{\Omega})$ satisfies $Lu \ge 0$, and $g \in C^2(\Omega) \cap C^0(\overline{\Omega})$ satisfies $Lg \le 0$, and suppose $c \le 0$. Suppose further that $g \ge u$ on $\partial\Omega$. Then $g \ge u$ everywhere on $\overline{\Omega}$.*

Remarks: (1) In case $x_0 \in \partial\Omega$ and $g(x_0) = u(x_0)$, this implies $\liminf_{s\downarrow 0} u(x_0 + s\eta) \le \liminf_{s\downarrow 0} g(x_0+s\eta)$, for any η such that the segment $x_0 + s\eta \in \Omega$ for $s \in (0,\varepsilon)$ for some $\varepsilon > 0$; the above proof of the Hopf boundary point lemma involves explicit construction of such a barrier function g on the domain $A = B_\rho(y) \setminus B_{\rho/2}(y)$.
(2) If the strict inequality $g > u$ holds everwhere on $\partial\Omega$, then the condition $c \le 0$ can be dropped. (See problem 13.5 below.)

We now state and prove the strong (Hopf) maximum principle:

Theorem 3 (Hopf Maximum Principle.) *Suppose $u \in C^2(\Omega)$, $Lu \ge 0$ in Ω, $c \le 0$, (‡), (‡‡) above hold, and Ω is connected. Then u attains a non-negative maximum in $\Omega \Rightarrow u$ is constant in Ω.*

If $c \equiv 0$, the phrase "non-negative" can be dropped. Also, if the hypothesis that $c \le 0$ is dropped, it is still true that u cannot have a zero maximum in Ω unless it is identically zero.

Remark: Actually the above theorem, being local in character, remains true without change in the proof if we assume that (‡), (‡‡) merely hold locally (on compact subsets of Ω).

Proof: Let M be the maximum value of u in Ω, and let $S = \{x \in \Omega : u(x) = M\}$. Since u is continuous, S is relatively closed in Ω. Since Ω is connected, the proof of the first part will be complete if we can show that S is also open. Take any point $x_0 \in S$. If x_0 is not an interior point of S, we can find points $y \in \Omega \setminus S$ with $\operatorname{dist}(y,S) < \operatorname{dist}(y,\partial\Omega)$. Let $x_0 \in S$ be a closest point of S to y (such x_0 exists because $\operatorname{dist}(y,S) < \operatorname{dist}(y,\partial\Omega)$), and let $\rho = |x_0 - y|$. By the Hopf boundary point lemma, we conclude that $Du(x_0) \neq 0$, contradicting the fact that u has a maximum at x_0.

The remaining parts of the lemma follow in much the same way as the auxiliary statements in the boundary point lemma.

For the remainder of this lecture we want to discuss maximum principles for weak solutions of second order elliptic equations of the form

$$\sum_{|\alpha|,|\beta|\le 1} D^\beta(a_{\alpha\beta} D^\alpha u) = 0.$$

For this, we first need to establish a little more Sobolev space theory. In particular, we want to be able to make sense out of statements like $u \ge v$, $u \le v$, and $u = v$ on $\partial\Omega$, in case u, v are merely in some Sobolev space $W^{1,p}(\Omega)$ with Ω an arbitrary bounded open subset of $\mathbb{R}^n$. Actually the statement $u = v$ on $\partial\Omega$ is easy to interpret: we simply take it to mean $u - v \in W_0^{1,p}(\Omega)$. To get a good interpretation for the inequalities, we have to work a little harder. First we need the following consequence of the chain rule of Lecture 5 for $W^{1,p}$ functions.

Lemma 3. *If $u \in W^{1,p}(\Omega)$, then $u_+ \in W^{1,p}(\Omega)$, and the weak derivatives are given by*

$$D_j u_+ = \begin{cases} D_j u \text{ a.e. where } u(x) > 0 \\ 0 \text{ a.e. where } u(x) \le 0. \end{cases}$$

Also,

$$Du = 0 \text{ a.e. on } \{x : u(x) = 0\}.$$

Proof: For each $\varepsilon > 0$, let $\varphi_\varepsilon(t) = \sqrt{\varepsilon^2 + (t_+)^2}$. Then $\varphi'_\varepsilon(t) = \frac{t_+}{\sqrt{\varepsilon^2+(t_+)^2}}$. Thus the required result follows by applying the chain rule of Lecture 5 with φ_ε in place of φ, and then letting $\varepsilon \downarrow 0$.

Now we can make sense out of statements like $u \ge v$ on $\partial\Omega$ in case $u, v \in W^{1,p}(\Omega)$:

Definition: If $u \in W^{1,p}(\Omega)$, we adopt the following conventions:

(i) we say $u \le 0$ on $\partial\Omega$ if $u_+ \in W_0^{1,p}(\Omega)$.

(ii) we say $u \ge 0$ on $\partial\Omega$ if $-u \le 0$ on $\partial\Omega$; i.e. if $u_- \in W_0^{1,p}(\Omega)$.

(iii) if $v \in W^{1,p}(\Omega)$, then we say $u \le v$ on $\partial\Omega$ if $u - v \le 0$ on $\partial\Omega$ (i.e. if $(u-v)_+ \in W_0^{1,p}(\Omega)$).

(iv) $\sup_{\partial\Omega} u$ is defined to be $\inf\{k : u \le k \text{ on } \partial\Omega\}$, and taken to be $+\infty$ if this set is empty.

Notice that according to the above definitions we have $u \in W_0^{1,p}(\Omega)$ is equivalent to $u = 0$ on $\partial\Omega$ in the sense that both $u \ge 0$ on $\partial\Omega$ and $u \le 0$ on $\partial\Omega$, and if $u, v \in W^{1,p}(\Omega)$ with $u \le v$ on $\partial\Omega$, then $\sup_{\partial\Omega} u \le \sup_{\partial\Omega} v$. Also if $u \in W^{1,p}(\Omega)$, then $\sup_\Omega u$ (which is defined to be $\inf\{k : u \le k \text{ a.e. on } \Omega\}$) is $\ge \sup_{\partial\Omega} u$. (See problem 13.5 below.)

We shall also need the following inequality for $W_0^{1,1}(\Omega)$ functions.

Lemma 4. *If $u \in W_0^{1,1}(\Omega)$ then for any measurable set $E \subset \Omega$,*

$$\int_E |u|\, dx \le C|E|^{1/n} \int_\Omega |Du|\, dx,$$

where $|E|$ denotes the Lebesgue measure of E and C depends only on n.

Note: Since $C_c^\infty(\Omega)$ is dense in $W_0^{1,1}(\Omega)$ (by definition of $W_0^{1,1}(\Omega)$) relative to the norm $\|\ \|_{W^{1,1}}$, it suffices to prove the above inequality for $C_c^\infty(\Omega)$ functions u; also, we can of course assume that the functions in question are all non-negative. Accordingly we assume $u \ge 0$ and $u \in C_c^\infty(\Omega)$ in the proof which follows.

Proof of the Lemma: From the Green's identity (see Lecture 4), we have

$$\begin{aligned} u(y) &= \int_\Omega K(x,y)\Delta u(x)\, dx \\ &= \int_\Omega D_x K(x,y) \cdot Du(x)\, dx \\ &= c_n \int_\Omega \frac{x-y}{|x-y|^n} \cdot Du(x)\, dx, \end{aligned}$$

so

$$|u(y)| \le c_n \int_\Omega \frac{|Du(x)|}{|x-y|^{n-1}}\, dx.$$

Now we integrate this inequality over E with respect to y. This gives

$$(*) \qquad \int_E |u(y)|\, dy \le c_n \int_\Omega |Du(x)| \Big(\int_E \frac{1}{|x-y|^{n-1}}\, dy \Big)\, dx.$$

Now for any $\rho > 0$ and $x \in \Omega$ fixed we have

$$\begin{aligned} \int_E \frac{1}{|x-y|^{n-1}}\, dy &= \int_{E\cap B_\rho(x)} \frac{1}{|x-y|^{n-1}}\, dy + \int_{E\setminus B_\rho(x)} \frac{1}{|x-y|^{n-1}}\, dy \\ &\le \int_{B_\rho(x)} \frac{1}{|x-y|^{n-1}}\, dy + \frac{1}{\rho^{n-1}} \int_E 1\, dy \\ &= n\omega_n \rho + \frac{|E|}{\rho^{n-1}}. \end{aligned}$$

Now select $\rho = |E|^{1/n}$; then the right side here is $\le c|E|^{1/n}$, hence $(*)$ implies

$$\int_E |u(y)|\,dy \le c|E|^{1/n}\int_\Omega |Du(x)|\,dx,$$

as required.

We are now ready to prove the weak maximum principle for weak solutions. We let Lu be an operator of the form $\sum_{|\alpha|,|\beta|\le 1} D^\beta(a_{\alpha\beta}D^\alpha u)$, or, in other notation, an operator of the form

$$Lu = \sum_{i,j=1}^n D_j\big(a_{ij}D_iu\big) + \sum_{j=1}^n\big(a_jD_ju + D_j(b_ju)\big) + cu,$$

and, by analogy with the condition $c \le 0$ in the previous lecture, we require that the coefficient of the zero-order term be non-negative in the weak sense; that is

$$\text{(1)} \qquad \int_\Omega\big(c\psi - \textstyle\sum_{j=1}^n b_jD_j\psi\big) \le 0 \ \forall\, \psi \in C_c^\infty(\Omega) \text{ with } \psi \ge 0.$$

We also assume

$$\text{(2)} \qquad a_{ij},\, a_j,\, b_j,\, c \in L^\infty(\Omega),$$

and the ellipticity condition

$$\text{(3)} \qquad \textstyle\sum_{i,j=1}^n a_{ij}\xi^i\xi^j \ge \mu|\xi|^2 \ \forall\, \xi \in \mathbb{R}^n,$$

where $\mu > 0$ is a constant.

We also need to of course interpret inequalities like $Lu \ge 0$ in the weak sense; for $u \in H^1(\Omega)$, we say $Lu \ge 0 (\le 0)$ on Ω if

$$\int_\Omega\big(\textstyle\sum_{i,j=1}^n a_{ij}D_iuD_j\psi + \sum_{j=1}^n(-a_jD_ju\psi + ub_jD_j\psi) - cu\psi\big) \le 0 (\ge 0)$$

$\forall\, \psi \in C_c^\infty(\Omega)$ with $\psi \ge 0$ on $\partial\Omega$.

Theorem 4. *Suppose $u \in H^1(\Omega)$, $Lu \ge 0$ in the weak sense described above, and suppose (1), (2), (3) above all hold. Then*

$$\sup_\Omega u \le \sup_{\partial\Omega} u_+,$$

and if $b_j = c = 0$ then

$$\sup_\Omega u = \sup_{\partial\Omega} u.$$

Proof: If $\sup_{\partial\Omega} u_+ = \infty$, there is nothing to prove, hence assume $\sup_{\partial\Omega} u_+ \in [0,\infty)$. Assume contrary to the statement that $\sup_\Omega u > \sup_{\partial\Omega} u_+$ and select $k > 0$, with $\sup_{\partial\Omega} u_+ < k < \sup_\Omega u$. Then $u_k \equiv (u-k)_+ \in H_0^1(\Omega)$ (by definition of $\sup_{\partial\Omega} u_+$) and we can plug into the weak form of the equation, and the inequality (1) to get

$$\begin{aligned}\int_\Omega \textstyle\sum_{i,j=1}^n D_iu_kD_ju_k &\le \int_\Omega \textstyle\sum_{j=1}^n\big(a_jD_ju_ku_k + ub_jD_ju_k\big) + ucu_k\\ &\le C\int_\Omega |Du_k||u_k| \le C\|Du_k\|_0\Big(\int_{|Du_k|>0} u_k^2\Big)^{1/2},\end{aligned}$$

by the Schwarz inequality. Using the ellipticity condition (3) we thus have

$$\|Du_k\|_0^2 \le C\int_{\{|Du_k|>0\}} u_k^2.$$

Using this and the lemma, with $E = \{|Du_k| > 0\}$ and $v = u_k^2$, we have

$$\begin{aligned}\int_{\{|Du_k|>0\}} u_k^2 &\le C|\{|Du_k|>0\}|^{1/n}\int_\Omega |u_k||Du_k|\\ &\le C|\{|Du_k|>0\}|^{1/n}\Big(\int_{|Du_k|>0} u_k^2\Big)^{1/2}\|Du_k\|_0\\ &\le C|\{|Du_k|>0\}|^{1/n}\int_{|Du_k|>0} u_k^2,\end{aligned}$$

where C is independent of k. Since $|Du_k| = 0$ a.e. implies $u_k = 0$ a.e. by the Poincaré inequality of Lecture 5, and since $k < \sup_\Omega u$, we see that $|Du_k| \ne 0$ on a set of positive measure, and hence the above inequality gives

$$|\{|Du_k| > 0\}| \ge c > 0,$$

with c independent of k, which by the above corollary to the chain rule implies

$$|\{u > k,\ |Du| > 0\}| \ge c > 0,$$

with c independent of k. Now letting $k \uparrow \sup_\Omega u$, we see that $|Du| > 0$ on a set of positive measure in $\{x \in \Omega : u(x) = \sup_\Omega u\}$, which is evidently impossible by Lemma 3 above.

LECTURE 13 PROBLEMS

13.1 (i) Suppose $u, v \in C^2(\Omega) \cap C^0(\overline{\Omega})$ with $Lu \geq 0$, $Lv \leq 0$ in Ω, where L is a second order operator satisfying the conditions (‡), (‡‡) of the above lecture. If $v > 0$ in $\overline{\Omega}$ and $u \leq v$ on $\partial\Omega$, prove that $u \leq v$ everywhere in $\overline{\Omega}$.

Hint: Consider $v - \lambda u$, $\lambda \geq 0$.

Note: The point here is that we make no assumption on the sign of the coefficient c in the operator L.

(ii) Suppose Ω is a bounded connected C^2 domain, L is as in (i), $v \in C^2(\Omega) \cap C^1(\overline{\Omega})$ with $v > 0$ in Ω (but we allow the possibility that $v = 0$ at some points of $\partial\Omega$), $Lv \leq 0$ in Ω, and $u \in C^2(\Omega) \cap C^1(\overline{\Omega})$ with $Lu \geq 0$ in Ω, $u \leq 0$ on $\partial\Omega$. Prove that either u is a positive constant multiple of v or $u \leq 0$ everywhere in $\overline{\Omega}$. (Notice that the former alternative implies $Lu = Lv = 0$ in Ω and $u = v = 0$ on $\partial\Omega$.)

13.2 (i) Suppose $u \in C^3(\Omega)$ satisfies an equation of the form

$$\textstyle\sum_{i=1}^n D_i[A_i(Du)] = 0,$$

where $A_i = A_i(p)$, $p \in \mathbb{R}^n$, is a C^2 function on $\mathbb{R}^n$ such that the matrix $(\partial A_i/\partial p_j)$ is positive definite at each point of $\mathbb{R}^n$. Prove that $|Du|$ satisfies a strict maximum principle in Ω. Prove also the stronger result that if $|Du|$ attains a maximum at some point $y \in \Omega$, then u is an affine function (i.e. linear + const.) in Ω.

(ii) Show that the result of (i) continues to hold in case the equation is modified to be $\sum_{i=1}^n D_i[A_i(Du)] = f(u, Du)$, where $f = f(z, p)$ is a C^2 function of $(z, p) \in \mathbb{R} \times \mathbb{R}^n$ with $f_z \geq 0$ everywhere.

13.3 (i) Suppose u_1, u_2 are $C^2(\Omega)$ solutions of the equation

$$\textstyle\sum_{i=1}^n D_i[A_i(Du)] = 0,$$

with A_i as in 13.3(i). Prove that the difference $u_1 - u_2$ satisfies a strict maximum principle in Ω, and hence in particular that the problem

$$\begin{cases} \sum_{i=1}^n D_i[A_i(Du)] = 0 \text{ in } \Omega \\ u = 0 \text{ on } \partial\Omega \end{cases}$$

has at most one solution.

(ii) Show that in fact this continues to hold if we merely have

$$\textstyle\sum_{i=1}^n D_i[A_i(Du_1)] \geq 0 \text{ and } \sum_{i=1}^n D_i[A_i(Du_2)] \leq 0 \text{ in } \Omega.$$

(iii) Extend (i) and (ii) to the case when the equation $\sum_{i=1}^n D_i[A_i(Du)] = 0$ is replaced by the equation $\sum_{i=1}^n D_i A_i(Du) = f(u, Du)$, with f as in 13.2(ii).

13.4 Suppose Q is the quadrant $|y| < x$ in $\mathbb{R}^2$, and suppose $u \in C^2(Q) \cap C^0(\overline{Q})$ is harmonic in Q, $\sup_{\partial Q} |u| < \infty$, and

$$\liminf_{R\to\infty} R^{-2} \sup_{\sqrt{x^2+y^2}=R} |u(x, y)| = 0.$$

Prove that $\sup_{\overline{Q}} |u| = \sup_{\partial Q} |u|$.

Hint: For any fixed $R > 0$, $\arctan\left(\frac{2R^2(x^2-y^2)}{R^4-(x^2+y^2)^2}\right)$ is harmonic for $\sqrt{x^2+y^2} < R$. (In fact it's the imaginary part of $\log\left(\frac{R^2+iz^2}{R^2-iz^2}\right)$.)

Lecture 14

Some Initial Applications to Non-linear Problems: Small Data Problems & Method of Sub/super-solutions

First we consider the application to non-linear problems with small data. We consider the $(2m)^{\text{th}}$-order problem

$$(*)\qquad \begin{cases} F(x, \{D^{\alpha+\beta}u\}_{|\alpha|\le m, |\beta|\le m}) = 0 \text{ on } \Omega \\ D^\gamma u = 0 \text{ on } \partial\Omega, \quad 0 \le |\gamma| \le m-1 \end{cases}$$

where Ω is a bounded $C^{2m,\mu}$ domain, and F are given. In place of the Dirichlet boundary condition $D^\gamma u = 0$, $|\gamma| \le m-1$ one could, without significantly changing the present discussion, consider various other boundary conditions, including the "partially free" boundary conditions as in Lectures 8 and 9.

We assume $F = F(x,q)$ is a C^4 function of the variables $q = \{q_{\alpha\beta}\}_{|\alpha|,|\beta|\le m}$, for each given $x \in \overline{\Omega}$, and

(i) $\quad |D_q^j F(x,q)| \le \Lambda, \quad |D_q^j F(x,q) - D_q^j F(\overline{x},q)| \le \Lambda |x-\overline{x}|^\mu, \;\; j \le 4,$

for all x, $\overline{x} \in \overline{\Omega}$ and all $q = \{q_{\alpha\beta}\}_{|\alpha|,|\beta|\le m}$ with $|q| \le 1$. Notice that we here think of $q = \{q_{\alpha\beta}\}_{|\alpha|,|\beta|\le m}$ as a point in $\mathbb{R}^N$ (i.e. as an ordered N-tuple), where N is the number of distinct ordered multi-index pairs (α,β) with $|\alpha|, |\beta| \le m$.

We also assume the ellipticity condition

(ii) $$\sum_{|\alpha|=|\beta|=m} a_{\alpha\beta}(x)\xi^\alpha\xi^\beta \ge \gamma|\xi|^{2m}, \quad x\in\overline{\Omega},\ \xi\in\mathbb{R}^n,$$

where $\gamma>0$ and $a_{\alpha\beta}(x) = (\partial F(x,q)/\partial q_{\alpha\beta})|_{q=0}$,

(In accordance with our previous discussion in Lectures 8,9, for the other boundary conditions $\mathcal{B}u=0$ we need to assume the stronger ellipticity condition

(ii)′ $$\sum_{|\alpha|=|\beta|=m} a_{\alpha\beta}\lambda_\alpha\lambda_\beta \ge \gamma\sum_{|\alpha|=m}(\lambda_\alpha)^2, \quad \forall\,\{\lambda_\alpha\}_{|\alpha|=m},$$

or at least the coercivity condition

(C) $$\int_\Omega \sum_{|\alpha|=|\beta|=m} a_{\alpha\beta} D^\alpha v D^\beta v \ge \gamma\int_\Omega\sum_{|\alpha|=m}|D^\alpha v|^2 - C\int_\Omega |v|^2,$$

for all $v\in C^{2m}(\overline{\Omega})$ with $\mathcal{B}v=0$ on $\partial\Omega$; of course the stronger ellipticity condition (ii)′ implies this with $C=0$.)

The linearized equation (near $u=0$) corresponding to $(*)$ is

$$Lu=f,$$

where
$$Lu \equiv \frac{d}{ds}F(x,s\{D^{\alpha+\beta}u\}_{|\alpha|,|\beta|\le m})\,|_{s=0},\ \ u\in C^{2m}(\Omega),$$
and where $f(x)=F(x,0)$, $x\in\overline{\Omega}$. Thus (using the chain rule) we can directly compute
$$Lu\equiv\sum_{|\alpha|,|\beta|\le m} a_{\alpha\beta}D^{\alpha+\beta}u,$$
where $a_{\alpha\beta}(x) = D_{q_{\alpha\beta}}F(x,q)|_{q=0}$. Thus Lu is a $(2m)^{\text{th}}$-order elliptic operator with $C^{0,\mu}(\overline{\Omega})$ coefficients, and the theory of Theorems 3, 4 of Lecture 12 can be applied.

Our main theorem concerning the non-linear problem $(*)$ is as follows:

Theorem 1. *Suppose F satisfies* (i), (ii), *and suppose also that there are no non-zero $C^{2m,\mu}(\overline{\Omega})$ solutions of the linear problem*

$(**)$ $$\begin{cases} Lu=0 \text{ in } \Omega\\ D^\gamma u = 0 \text{ on } \partial\Omega, \quad |\gamma|\le m-1,\end{cases}$$

where L is the linearized operator described above. Then there is a positive $\varepsilon_0=\varepsilon_0(n,\gamma,\lambda,\Omega)$ such that the problem $(*)$ has a $C^{2m,\mu}(\overline{\Omega})$ solution u whenever $|f|_{0,\mu,\Omega}\le\varepsilon_0$, where $f(x)\equiv F(x,0)$ for $x\in\overline{\Omega}$.

Proof: We use the contraction mapping principle. First note that the equation can be written in the form

(1) $$Lu=\mathcal{N}(u)+f,$$

where
$$\mathcal{N}(u)=N(x,\{D^{\alpha+\beta}u\}_{|\alpha|,|\beta|\le m}),$$
with $N=N(x,q)$, $q=\{q_{\alpha\beta}\}_{|\alpha|,|\beta|\le m}$, given by
$$N(x,q)=F(x,q)-\sum_{|\alpha|,|\beta|\le m} q_{\alpha\beta}(D_{q_{\alpha\beta}}F)(x,0).$$

Notice that then, using the extended mean-value theorem $f(1)-f(0) = f'(0)+\int_0^1(1-s)f''(st)\,dt$ (applied to $f(t)=F(x,tq)$), we can write
$$N(x,q)=\int_0^1(1-s)(D_q^2F)|_{(x,sq)}(q,q)\,ds,$$
where $D_q^2F|_{(x,sq)}$ denotes the Hessian of F with respect to the q-variables, evaluated at (x,sq). (Thus $(D_q^2F)|_{(x,sq)}(p,p)$is a bilinear function of p.) Using this identity together with (i) and (ii), it is then straightforward to check that

(2) $$|\mathcal{N}(u)|_{0,\mu,\Omega}\le C|u|^2_{2m,\mu,\Omega}$$

and

(3) $$|\mathcal{N}(u_1)-\mathcal{N}(u_2)|_{0,\mu,\Omega}\le C\left(|u_1|_{2m,\mu,\Omega}+|u_2|_{2m,\mu,\Omega}\right)|u_1-u_2|_{2m,\mu,\Omega}$$

for all $u,u_1,u_2\in C^{2m,\mu}(\overline{\Omega})$ with $C^{2m,\mu}$-norm ≤ 1; here C depends on F only.

We define a map $T: C^{2m,\mu}(\overline{\Omega})\to C^{2m,\mu}(\overline{\Omega})$ as follows: For $v\in C^{2m,\mu}(\overline{\Omega})$, define $T(v)=w$, where
$$\begin{cases} Lw=\mathcal{N}(v)+f \text{ in } \Omega\\ D^\gamma w=0 \text{ on } \partial\Omega, \quad |\gamma|\le m-1.\end{cases}$$

Note that by the Fredholm Alternative (FA) of the supplement to Lecture 12, we know that such w exists, and satisfies

(4) $$|w|_{2m,\mu,\Omega}\le C|\mathcal{N}(v)+f|_{0,\mu,\Omega},$$

with C depending only on Ω and F. Furthermore if v_1,v_2 are given $C^{2m,\mu}(\overline{\Omega})$ functions, and if we write $w_j=T(v_j)$, $j=1,2$, then, since $L(w_1-w_2)=$

$\mathcal{N}(v_1)-\mathcal{N}(v_2)$, we use the same result (FA from the supplement to Lecture 12) again to give

$$(5)\qquad |w_1 - w_2|_{2m,\mu,\Omega} \le C|\mathcal{N}(v_1) - \mathcal{N}(v_2)|_{0,\mu,\Omega}.$$

In view of (2), (3), the inequalities (5), (6) imply

$$|w|_{2m,\mu,\Omega} \le C\left(|v|^2_{2m,\mu,\Omega} + |f|_{0,\mu,\Omega}\right)$$
$$|w_1 - w_2|_{2m,\mu,\Omega} \le C\left(|v_1|_{2m,\mu,\Omega} + |v_2|_{2m,\mu,\Omega}\right)|v_1 - v_2|_{2m,\mu,\Omega},$$

where C is fixed, so long as v, v_1, v_2 have $C^{2m,\mu}(\overline{\Omega})$-norm ≤ 1. It is therefore straightforward to check that, for small enough $\varepsilon_0 > 0$, T is a contraction map of $S_{\varepsilon_0} \equiv \{v \in C^{2m,\mu}(\overline{\Omega}) : \mathcal{B}v = 0 \text{ on } \partial\Omega,\ |v|_{2m,\mu,\Omega} \le \sqrt{\varepsilon_0}\}$ into itself, provided (i), (ii) hold with $\lambda \le \varepsilon_0$. Indeed by virtue of the above inequalities

$$|T(v)|_{2m,\mu,\Omega} \le C\varepsilon_0 < \sqrt{\varepsilon_0}$$
$$|T(v_1) - T(v_2)|_{2m,\mu,\Omega} \le C\sqrt{\varepsilon_0}|v_1 - v_2|_{2m,\mu,\Omega} \le \frac{1}{2}|v_1 - v_2|_{2m,\mu,\Omega},$$

for each $v,\ v_1,\ v_2 \in S_{\varepsilon_0}$.

By virtue of the contraction mapping principle, we then deduce that for ε_0 small enough there is a $v \in S_{\varepsilon_0}$ such that $T(v) = v$; that is, (1) is satisfied with $w = v$. This completes the proof of Theorem 1.

Remark: Note that the contraction mapping proof given above is constructive, in the sense that if we define v_k iteratively by $v_0 \equiv 0$, $v_{k+1} = T(v_k)$, $k = 0, 1, \ldots$, then the sequence $\{v_k\}$ converges in the $C^{2m,\mu}(\overline{\Omega})$-norm to the required solution u of (1), and $|u - v_k|_{2m,\mu,\Omega} \le 2^{-k+1}\sqrt{\varepsilon_0}$; this follows from the usual proof of the contraction mapping theorem.

Next we want to describe the <u>method</u> <u>of</u> <u>sub</u>/super-solutions for solving semi-linear problems; that is, problems of the form

$$(*)\qquad \begin{cases} Lu = G(x,u) \text{ in } \Omega \\ u = \psi \text{ on } \partial\Omega \end{cases}$$

where ψ is a given $C^{2,\mu}(\overline{\Omega})$ function, and

$$Lu = \textstyle\sum_{i,j=1}^n a_{ij} D_i D_j u + \sum_{j=1}^n b_j D_j u + cu$$

is a linear elliptic operator with Hölder continuous coefficients in $C^{0,\mu}(\overline{\Omega})$ (as above), but <u>without</u> the additional restriction that the problem $(*)(*)$ of Theorem 1 above has no non-trivial solutions. Concerning the function G we assume that there exist constants K, λ such that

$$\text{(i)}\qquad |G(x,z)| \le \lambda,\ \ |G(x_1,z_1) - G(x_2,z_2)| \le \lambda(|x_1 - x_2|^\mu + |z_1 - z_2|)$$

for all $x_1, x_2 \in \overline{\Omega}$ and $z_1, z_2 \in \mathbb{R}$ with $|z_1|$, $|z_2| \le K$.

What we want to prove is that $(*)$ has a $C^{2,\mu}(\overline{\Omega})$ solution whenever there exist $C^{2,\mu}(\overline{\Omega})$ super- and sub-solutions φ^+, φ^- with $-K \le \varphi^- \le \varphi^+ \le K$ everywhere on Ω. That is, we assume that there exist $C^{2,\mu}(\overline{\Omega})$ functions $\varphi^\pm$ with

$$\text{(ii)}\qquad -K \le \varphi^- \le \varphi^+ \le K,$$

$$\text{(iii)}\qquad \begin{cases} L\varphi^+ \le G(x,\varphi^+) \text{ in } \Omega \\ \varphi^+ \ge \psi \text{ on } \partial\Omega, \end{cases}$$

and

$$\text{(iv)}\qquad \begin{cases} L\varphi^- \ge G(x,\varphi^-) \text{ in } \Omega \\ \varphi^- \le \psi \text{ on } \partial\Omega, \end{cases}$$

Then we have the following theorem:

Theorem 2. *If* (i), (ii), (iii), (iv) *all hold, if Ω is a bounded $C^{2,\mu}$ domain, if $\psi \in C^{2,\mu}(\overline{\Omega})$, and if the operator $Lu \equiv \sum_{i,j=1}^n a_{ij} D_i D_j u + \sum_{j=1}^n b_j D_j u + cu$ is elliptic (i.e. (a_{ij}) is positive definite at each point of $\overline{\Omega}$), and if the coefficients $a_{ij}, b_j, c \in C^{0,\mu}(\overline{\Omega})$, then the problem $(*)$ has a $C^{2,\mu}(\overline{\Omega})$ solution.*

Remark: Actually we prove theorems of a more general character using Leray-Schauder existence program (see problem 14.1 of the next lecture); however the method here is especially elegant and is also somewhat more constructive than the Leray-Schauder approach—we solve the equation by iteratively constructing an increasing sequence of solutions of certain linear problems.

Proof of Theorem 2: We first note that we may assume without loss of generality that the coefficient c in the operator L is ≤ 0, and that $G(x,z)$ is decreasing in z for $z \in [-K, K]$; this is arranged simply by subtracting λu from each side of the equation. We therefore assume this from now on; also let us replace G by the function G_K defined by $G_K(x,z) = G(x,z)$ for $|z| \le K$, $G_K(x,z) = G(x,K)$ for $z > K$, and $G_K(x,z) = G(x,-K)$ for $z < -K$.

Now we inductively define a sequence $\{u_k\}_{k=0,1,\ldots}$ of $C^{2,\mu}(\overline{\Omega})$ functions as follows: Let $u_0 \equiv \varphi^-$, and assume $k \ge 0$ and that $u_0, \ldots, u_k$ are already defined as $C^{2,\mu}(\overline{\Omega})$ functions which satisfy the inequalities

$$u_0 = \varphi^- \le u_1 \le \cdots \le u_k \le \varphi^+. \tag{1}$$

Assuming we have already construction $u_0, \ldots, u_k \in C^{2,\mu}(\overline{\Omega})$, we in fact let u_{k+1} be defined to be the solution of the linear problem

$$\begin{cases} Lu_{k+1} = G_K(x, u_k) \text{ in } \Omega \\ u_{k+1} = \psi \text{ on } \partial\Omega. \end{cases}$$

(Note this can be written $L(u_{k+1} - \psi) = f$, $u_{k+1} - \psi|\partial\Omega = 0$, where $f = G_K(x, u_k) - L\psi$, so the Fredholm Alternative FA of the supplement Lecture 12 can be applied; since there are no nontrivial solutions of $Lu = 0$ with $u = 0$ on $\partial\Omega$ this does give us the existence of a solution $u_{k+1} \in C^{2,\mu}(\overline{\Omega})$.)

To complete the inductive definition we now have to check that the inequalities (1) hold with $k+1$ in place of k. To see this first note that if $k = 0$ then the above says that u_1 satisfies the problem

$$\begin{cases} Lu_1 = G_K(x, \varphi^-) \text{ in } \Omega \\ u_1 = \psi \text{ on } \partial\Omega, \end{cases}$$

and hence (in view of (iv) above) $v = u_1 - \varphi^-$ satisfies

$$\begin{cases} Lv \le 0 \text{ in } \Omega \\ v \ge 0 \text{ on } \partial\Omega, \end{cases}$$

and so by the maximum principle $v \ge 0$ everywhere in Ω. That is $u_1 \ge \varphi^-$ everywhere in Ω.

Next we note that if $k \ge 1$ and if (1) holds, then by the above definitions we have

$$\begin{cases} L(u_{k+1} - u_k) = G_K(x, u_k) - G_K(x, u_{k-1}) \text{ in } \Omega \\ u_{k+1} = u_k \text{ on } \partial\Omega \end{cases}$$

and hence from the fact that $G_K(x,z)$ is decreasing in z we deduce that $L(u_{k+1} - u_k) \le 0$ if $u_k \ge u_{k-1}$ in Ω, and hence $u_{k+1} \ge u_k$ in Ω whenever $u_k \ge u_{k-1}$ in Ω. Thus by induction $\{u_k\}$ is an increasing sequence. Finally, if $k \ge 0$, by (iii) and the definition of u_{k+1} we have that

$$\begin{cases} L(u_{k+1} - \varphi^+) \ge G_K(x, u_k) - G_K(x, \varphi^+) \text{ in } \Omega \\ u_{k+1} - \varphi^+ \le 0 \text{ on } \partial\Omega. \end{cases}$$

Thus $L(u_{k+1} - \varphi^+) \ge 0$ in Ω (and hence $u_{k+1} \le \varphi^+$ in Ω by the maximum principle) provided $u_k \le \varphi^+$ in Ω, so (since $u_0 = \varphi^- \le \varphi^+$) we have by induction that $u_k \le \varphi^+$ for all k. This completes the proof of (1).

Now that (1) is proved we see (since $-K \le \varphi^- \le \varphi^+ \le K$) that all the above actually holds with G in place of G_K everywhere. In view of the bounds (1), we now have that the sequence $\{u_k\}$ converges pointwise in $\overline{\Omega}$ to a function u. We want to show that the convergence is in the $C^2(\overline{\Omega})$-norm; this will complete the proof.

Using the definition of u_{k+1} together with Schauder estimates and the fact that (by (i)) $|G(x,u_k)|_{0,\mu,\Omega} \le 2\lambda + \lambda |u_k|_{0,\mu,\Omega}$, we see

$$|u_{k+1}|_{2,\mu,\Omega} \le C\left(\lambda + (1+\lambda)|u_k|_{0,\mu} + |\psi|_{2,\mu,\Omega}\right),$$

where C is independent of k. Since $|u_k|_{0,\mu} \le \varepsilon |u_k|_{2,\mu} + C(\varepsilon)|u_k|_0$ by interpolation (where $\varepsilon > 0$ is arbitrary), we then see that (since $|u_k|_{0,\Omega} \le |\varphi^+|_{0,\Omega} + |\varphi^-|_{0,\Omega}$ by (1))

$$|u_{k+1}|_{2,\mu,\Omega} \le \frac{1}{2}|u_k|_{2,\mu,\Omega} + C,$$

where C is independent of k. Iterating this, we obtain

$$|u_{k+1}|_{2,\mu,\Omega} \le 2C + \frac{1}{2^{k-1}}|u_1|_{2,\mu,\Omega}.$$

Thus, by the Arzela-Ascoli lemma, every subsequence of $\{u_k\}$ itself has a subsequence which converges in the C^2-norm on Ω, and since the pointwise limit is unique (namely u), the whole sequence must converge to u with respect to the $C^2(\overline{\Omega})$-norm. u then evidently satisfies the required equation and the given boundary condition.

LECTURE 14 PROBLEMS

14.1 If $\Omega \subset \mathbb{R}^2$ and if $u \in C^2(\Omega)$, then the Gauss curvature $K(u)$ of graph u is given by

$$K(u) = (1 + |Du|^2)^{-2}(u_{11}u_{22} - u_{12}^2),$$

where $u_{ij} = D_i D_j u$.

(i) Check (by direct computation) that the Gauss curvature of the hemisphere $z = \sqrt{1 - |x|^2}$ is equal to 1.

(ii) Using Theorem 1 of the above lecture, establish that for any $\rho \in (0, 1)$ there is an $\varepsilon = \varepsilon(\rho) > 0$ such that if $|f|_{0,\mu,B_\rho(0)} \le \varepsilon$, then there is a $u \in C^{2,\mu}(\overline{B}_\rho(0))$ satisfying

$$\begin{cases} K(u) = 1 + f \text{ on } B_\rho(0) \\ \quad u = 0 \text{ on } \partial B_\rho(0). \end{cases}$$

Hint: Write $w = u + \sqrt{1 - \rho^2} - \sqrt{1 - |x|^2}$, and work with the unknown w instead of u. Notice that the equation $K(u) = f$ is the equation of prescribed Gaussian curvature; it is of some geometric interest.

14.2 Use the method of sub/super-solutions to prove that there exists a $C^{2,\mu}(\overline{\Omega})$ solution u of the problem

$$\begin{cases} \Delta u + u^3 = 0 \text{ in } \Omega \\ \qquad u = \varphi \text{ on } \partial\Omega, \end{cases}$$

where Ω is a $C^{2,\mu}$ domain contained in the unit ball of $\mathbb{R}^n$ and $\varphi \in C^{2,\mu}(\overline{\Omega})$, in each of the cases: (i) $n = 3$ and $|\varphi| \le 5/6$, (ii) $n = 4$ and $|\varphi| \le 1$.

14.3 The minimal surface equation is the equation $\mathcal{M}(u) = 0$, where the operator $\mathcal{M}$ is defined by $\mathcal{M}(u) = \sum_{j=1}^n D_j(D_j u/\sqrt{1 + |Du|^2})$.

(i) Establish the alternative identity

$$\mathcal{M}(u) = (1 + |Du|^2)^{-1/2}\big(\Delta u - \textstyle\sum_{i,j=1}^n (1 + |Du|^2)^{-1} D_i u D_j u D_i D_j u\big).$$

(ii) If Ω is a bounded $C^{2,\mu}$ domain in $\mathbb{R}^n$, prove that there exists $\varepsilon_0 = \varepsilon_0(\Omega, \mu) > 0$ such that the problem

$$\begin{cases} \mathcal{M}(u) = 0 \text{ in } \Omega \\ u = \varphi \text{ on } \partial\Omega, \end{cases}$$

provided $\varphi \in C^{2,\mu}(\overline{\Omega})$ with $|\varphi|_{2,\mu,0} \le \varepsilon_0$.

Hint: In terms of the new unknown $\tilde{u} = u - \varphi$, show that the hypotheses of Theorem 1 hold with $m = 1$.

14.4 State and prove a version of Theorem 1 for equations

$$\textstyle\sum_{|\alpha| \le m} D^\alpha(F_{q^\alpha}(x, \{D^\gamma u\}_{|\gamma| \le m}))$$

on bounded $C^{m,\mu}$ domains.

14.5 Using the general theorem established in problem 14.4 above, generalize the result of 14.3(ii) to the case when the domain and φ are both of class $C^{1,\mu}$.

Lecture 15

The Leray-Schauder Approach to Non-linear Existence Theory

Here we want to describe the approach of Leray and Schauder to existence theory for non-linear problems—in particular we want to be able to discuss existence theory for the kind of equations which arise as Euler-Lagrange equations for functionals of the form $\mathcal{F}(u) = \int_\Omega F(x, u, Du)\, dx$.

The general method is based on the following Leray-Schauder fixed point theorem:

Theorem 1. *Suppose $\mathcal{B}$ is a Banach space with norm $\| \|$, $\mathcal{C}$ is a closed convex subset of $\mathcal{B}$, x_0 is a point of $\mathcal{C}$, $T : \mathcal{C} \times [0,1] \to \mathcal{C}$ is continuous and compact with $T(x, 0) \equiv x_0$, $x \in \mathcal{C}$, and suppose there is a fixed constant $M > 0$ such that*

$$(*) \qquad \{x \in \mathcal{C} : \exists \sigma \in [0,1] \text{ with } T_\sigma(x) = x\} \text{ is a bounded subset of } \mathcal{B}.$$

(i.e. the set of all points which are fixed points of at least one of the operators T_σ, $0 \le \sigma \le 1$, form a bounded subset of $\mathcal{B}$.) Then $T(\cdot, 1)$ has a fixed point; that is $T(x, 1) = x$ for some $x \in \mathcal{C}$.

Remarks: (1) The case $\mathcal{C} = \mathcal{B}$ is important.

(2) Notice that either the hypothesis $(*)$ holds (and consequently T_1 has a fixed point), or there are sequences $\{x_j\} \subset \mathcal{C}$ and $\sigma_j \in [0,1]$ with $T_{\sigma_j}(x_j) = x_j$ for each j and $\|x_j\| \to \infty$ as $j \to \infty$.

Abstractly the Leray-Schauder approach to non-linear problems can be described as follows:

(1) Given a non-linear problem (P), we set up a 1-parameter family $(\mathrm{P}_\sigma)_{0\le\sigma\le 1}$ of problems with $(\mathrm{P}_1) = (\mathrm{P})$, and a related 1-parameter family of mappings T_σ, each T_σ mapping a convex subset $\mathcal{C}$ of a Banach space $\mathcal{B}$ into itself, with the properties:

(i) $T_0(x) \equiv x_0,\ x \in \mathcal{C}$, where $x_0 \in \mathcal{C}$. (i.e. $T_0 \equiv$ constant)

(ii) the map $T : (x, \sigma) \mapsto T_\sigma(x)$ is a compact continuous map from $\mathcal{C} \times [0, 1]$ into $\mathcal{C}$

(iii) a fixed point of T_σ is a solution of the problem (P_σ);

(2) By analyzing solutions x of the non-linear problems (P_σ), one proves there is a fixed constant M such that

$$(*) \qquad \|x\| < M$$

whenever x solves one of the problems (P_σ), $0 \le \sigma \le 1$;

(3) Apply the Leray-Schauder fixed point theorem to give a fixed point of T_1 (and hence a solution of $(\mathrm{P}_1) = (\mathrm{P})$.

In practice establishing the estimate $(*)$ is usually the most difficult step in the implementation of this procedure; certainly this is so in the applications to quasilinear equations discussed here and in the next lecture.

Before discussing such applications, we give a proof of the Leray-Schauder fixed point theorem. As a preliminary, we need the following infinite dimensional analogue of the Brouwer fixed point theorem:

Theorem 2 (Schauder). *If $B = \{x \in \mathcal{B} : \|x\| \le 1\}$ and if $T : B \to B$ is compact and continuous, then T has a fixed point.*

Proof: Let K denote the closure of the convex hull of $T(B)$. Then K is compact (by the general Banach space fact that the closure of the convex hull of a compact set is compact). Then $T : K \to K$ by restriction. For $\varepsilon > 0$ arbitrary, choose a finite cover of K by open balls $\{B_\varepsilon(x_j)\}_{j=1,\ldots,N}$ with radius ε and centres $x_j \in K$, and define

$$J_\varepsilon = \frac{\sum_{i=1}^N \operatorname{dist}(x, K \setminus B_\varepsilon(x_i))x_i}{\sum_{j=1}^N \operatorname{dist}(x, K \setminus B_\varepsilon(x_j))}\,.$$

Notice that J_ε maps K into the convex hull K_N of $\{x_j\}_{j=1,\ldots,N}$, and

$$(*) \qquad \|J_\varepsilon(x) - x\| \le \varepsilon \quad \forall\ x \in K.$$

Then $T_\varepsilon \equiv J_\varepsilon \circ T|K_N$ is a continuous map of K_N into itself. Since K_N is a convex subset of the finite dimensional subspace of $\mathcal{B}$ which is spanned by $\{x_j\}_{j=1,\ldots,N}$, we know that K_N is homeomorphic to a ball in some Euclidean space $\mathbb{R}^n (n \le N)$. Hence by the Brouwer fixed point theorem T_ε has a fixed point x_ε. But then by $(*)$

$$\|T(x_\varepsilon) - x_\varepsilon\| < \varepsilon$$

and using the compactness of K it is then easy to show that a convergent sequence x_{ε_j} $(\varepsilon_j \downarrow 0)$ converges to x with $T(x) = x$.

Corollary 1. *If $T : B \to \mathcal{B}$ is continuous and compact, and if $T(\partial B) \subset B$, then T has a fixed point in B.*

Proof: Define $T_* = R \circ T$, where R is the radial retraction of $\mathcal{B}$ onto the unit ball B (that is, $R(y) = y$ if $\|y\| \le 1$ and $R(y) = y/\|y\|$ if $\|y\| > 1$). Then T_* has a fixed point $y \in B$ by the above theorem; thus

$$(*) \qquad T_*(y) = y, \quad y \in B.$$

If $\|y\| = 1$, then $R \circ T(y) = T(y)$ (because $T(\partial B) \subset B$), so $(*)$ gives $T(y) = y$ in this case. On the other hand, if $\|y\| < 1$ then $\|T(y)\| < 1$, otherwise $(*)$ would give $\|y\| = \|R(T(y))\| = 1$, and hence again $(*)$ give $T(y) = y$.

Proof of Theorem 1: First we show that the theorem can be reduced to the case $\mathcal{C} = \mathcal{B}$. In fact given any closed convex set $\mathcal{C} \subset \mathcal{B}$, we can find a continuous $Q : \mathcal{B} \to \mathcal{C}$ with $Q(x) \equiv x,\ x \in \mathcal{C}$. (e.g. by Dugundji's extension theorem—see Theorem 7.2, pp.44–45 of "Nonlinear Functional Analysis," by Klaus Deimling, Springer 1985). Then we can apply Theorem 1 with all of $\mathcal{B}$ in place of $\mathcal{C}$ and with $\widetilde{T}_\sigma = T_\sigma \circ Q : \mathcal{B} \to \mathcal{B}$ in place of T_σ, and notice that, since $T_\sigma : \mathcal{C} \to \mathcal{C}$, a point x is a fixed point of $\widetilde{T}_\sigma$ if and only if $x \in \mathcal{C}$ and x is a fixed point of T_σ.

We can assume without loss of generality (by replacing $T(x, \sigma)$ by $T(x + x_0, \sigma) - x_0$ and $\mathcal{C}$ by $\mathcal{C} - x_0$) that $x_0 = 0 \in \mathcal{C}$. Furthermore we can also assume $M = 1$ in the inequality $(*)$ (otherwise replace T by $\widetilde{T}$, defined by $\widetilde{T}(x, \sigma) = M^{-1}T(Mx, \sigma)$).

For $\varepsilon \in (0,1)$, let $Q_\varepsilon : B \to \mathcal{B} \times [0,1]$ be the continuous map defined by

$$Q_\varepsilon(x) = \begin{cases} (\frac{x}{\|x\|}, \frac{1-\|x\|}{\varepsilon}), & \text{if } 1-\varepsilon \le \|x\| \le 1 \\ (\frac{x}{1-\varepsilon}, 1), & \text{if } \|x\| \le 1-\varepsilon, \end{cases}$$

and define

$$\widetilde{T} = T \circ Q_\varepsilon : B \to \mathcal{B}.$$

Observe that then $\widetilde{T}$ is continuous and compact and, by construction, $\widetilde{T}(\partial B) = \{0\}$, so we can apply Theorem 2 to give $x_\varepsilon \in B$ such that

$$\widetilde{T}(x_\varepsilon) = x_\varepsilon.$$

We claim that $\|x_\varepsilon\|$ is bounded away from 1 as $\varepsilon \downarrow 0$; indeed otherwise there is a sequence $\varepsilon_j \downarrow 0$ with $\|x_{\varepsilon_j}\| \to 1$; since $T_{\varepsilon_j}(x_{\varepsilon_j}) = T(Q_{\varepsilon_j}(x_{\varepsilon_j}))$ and since T is compact, we can select a convergent subsequence $x_{\varepsilon_{j'}}$ which converges to $x \in \partial B$. Evidently then (taking another subsequence if necessary)

$$Q_{\varepsilon_{j'}}(x_{\varepsilon_{j'}}) \to (x, \sigma)$$

for some $\sigma \in [0,1]$, and hence

$$T(x,\sigma) = x,$$

which contradicts the hypothesis $(*)$ of Theorem 1 (with $M = 1$). Thus $\|x_{\varepsilon_j}\|$ is bounded away from 1, and hence $Q_{\varepsilon_j}(x_{\varepsilon_j}) = (1-\varepsilon_j)^{-1}x_{\varepsilon_j}$ for all sufficiently large j and (using compactness of T again) there is a subsequence $x_{\varepsilon_{j'}}$ converging to $x \in B$ with $T(x,1) = x$, as required.

We now illustrate the use of the Leray-Schauder method with a simple class of non-linear problems. We in fact consider the following class of problems which generalize the method of sub/super-solutions considered in the last lecture to the case when the right side also depends on Du. Thus we consider problems

$$(\ddagger) \qquad \begin{cases} Lu = f(x,u,Du) \text{ in } \Omega \\ u = \psi \text{ on } \partial\Omega, \end{cases}$$

where Ω is a bounded connected $C^{2,\mu}$ domain, ψ is a given $C^{2,\mu}(\overline{\Omega})$ function, and $Lu = \sum_{i,j=1}^n a_{ij}(x) D_i D_j u$, with $a_{ij} = a_{ji}$ and

$$\text{(i)} \qquad \lambda|\xi|^2 \le \sum_{i,j=1}^n a_{ij}\xi_i\xi_j, \qquad |a_{ij}|_{0,\mu,\Omega} \le \Lambda, \quad x \in \overline{\Omega},\ \xi \in \mathbb{R}^n,$$

where λ, Λ are positive constants. f is a given continuous function on $\overline{\Omega} \times \mathbb{R} \times \mathbb{R}^n$; we impose restrictions on f in (iii), (iv) below (and the additional condition (v) in case $n \ge 3$).

Let $M > 0$ be given. As in the previous lecture we need to assume the existence of sub and super-solutions: we assume that there exist $\varphi^\pm \in C^{2,\mu}(\overline{\Omega})$ such that $-M \le \varphi^- \le \varphi^+ \le M$ and

$$\text{(ii)} \qquad \begin{cases} L\varphi^+ \le f(x,\varphi^+,D\varphi^+) \text{ in } \Omega \\ \varphi^+ \ge \psi \text{ on } \partial\Omega. \end{cases} \qquad \begin{cases} L\varphi^- \ge f(x,\varphi^-,D\varphi^-) \text{ in } \Omega \\ \varphi^- \le \psi \text{ on } \partial\Omega. \end{cases}$$

Concerning the function f we assume that $f : \overline{\Omega} \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ and there is a constant $\Lambda > 0$ with

$$\text{(iii)} \qquad |f(x,z,p)| \le \Lambda(1+|p|)^2, \quad |D_p f(x,z,p)| \le \Lambda(1+|p|)$$

$$\text{(iv)} \qquad |f(x_2,z_2,p) - f(x_1,z_1,p)| \le \Lambda(|x_1-x_2|^\mu + |z_1-z_2|)(1+|p|)^2$$

for all $x, x_1, x_2 \in \overline{\Omega}$, $z, z_1, z_2 \in [-M, M]$ and all $p \in \mathbb{R}^n$. In case $n \ge 3$, we also need to assume the additional condition that for each $M > 0$ there is a function $\delta_M : [0,\infty) \to [0,\infty)$ such that $\lim_{t\to\infty} \delta_M(t) = 0$ and such that

$$\text{(v)} \qquad f(x,z_2,p) - f(x,z_1,p) \ge -\delta_M(|p|)(1+|p|)^2$$

for all $x \in \overline{\Omega}$, $p \in \mathbb{R}^n$ and $-M \le z_1 \le z_2 \le M$. Notice that this is trivially satisfied if either $f(x,z,p)$ is an increasing function of z or if $|f(x,z,p)| \le \delta_M(|p|)(1+|p|)^2$ for some such function δ_M and all $(x,z,p) \in \overline{\Omega} \times [-M, M] \times \mathbb{R}^n$.

Then we have the following generalization of the method of sub and super-solutions:

Theorem 3. *If $n = 2$ and if* (i)–(iv) *above hold, then there is $u \in C^{2,\mu}(\overline{\Omega})$ satisfying the problem* $(\ddagger)$. *The same is true for $n \ge 3$ under the additional hypothesis* (v).

For the proof we shall use the following two lemmas together with the Leray-Schauder theorem.

Lemma 1. *Suppose $|f(x,z,p)| \le \Lambda(1+|p|)^2$ for all $(x,z,p) \in \overline{\Omega} \times [-M, M] \times \mathbb{R}^n$, and suppose also the Lipschitz condition: $\sup_{(x,z)\in\Omega\times[-M,M], |p|,|q|<K, p\ne q} |p-q|^{-1}|f(x,z,p) - f(x,z,q)| < \infty$ for each $K > 0$, and*

$$\lambda|\xi|^2 \le \sum_{i,j=1}^n a_{ij}(x)\xi^i\xi^j \le \Lambda|\xi|^2 \ \ \forall x \in \Omega,\ \xi \in \mathbb{R}^n,$$

and suppose $u \in C^2(\Omega) \cap C^1(\overline{\Omega})$ is a solution of problem (‡) with $|u| \le M$ and $|\psi|_{C^2} \le \Lambda$. Then there is a constant C, depending only on $\Omega, \lambda, \Lambda, M$, such that

$$\sup_{\partial\Omega} |Du| \le C.$$

Proof: The proof uses a "local barriers" construction near points $y \in \partial\Omega$. The technique here can be modified to work also with the more general quasilinear problems to be considered in the next lecture.

Suppose $y \in \partial\Omega$, choose any ball $B_\rho(x_0) \subset \mathbb{R}^n \setminus \Omega$ such that $y \in \partial B_\rho(x_0)$, and define

$$\Psi(r) = \theta(\log(r/\rho - 1 + \varepsilon) - \log\varepsilon),\ r \ge \rho,$$

where $\theta, \varepsilon > 0$ are constants to be chosen. Our aim is to use Ψ to construct a barrier function on the region

$$U = \Omega \cap \{x : \rho < |x - x_0| < (1 + \varepsilon^{1/2})\rho\},$$

By direct computation

$$\Psi'(r) = \theta\rho^{-1}(r/\rho - 1 + \varepsilon)^{-1}, \quad \Psi''(r) = -\theta\rho^{-2}(r/\rho - 1 + \varepsilon)^{-2},$$

and hence, with $r = |x - x_0|$,

$$D_i[\Psi(r)] = \theta\rho^{-1}(r/\rho - 1 + \varepsilon)^{-1}x^i/r$$

and

$$D_iD_j[\Psi(r)] = -\theta\rho^{-2}(r/\rho - 1 + \varepsilon)^{-2}(r^{-2}x^ix^j) + \theta r^{-1}\rho^{-1}(r/\rho - 1 + \varepsilon)(\delta_{ij} - r^{-2}x^ix^j).$$

Using the ellipticity condition $\lambda|\xi|^2 \le \sum a_{ij}\xi^i\xi^j \le \Lambda|\xi|^2$, it then follows by direct computation that

$$L(\Psi(r)) \le -\theta\lambda\rho^{-2}(r/\rho - 1 + \varepsilon)^{-2} + n\Lambda\theta\rho^{-1}(r/\rho - 1 + \varepsilon)^{-1}.$$

Now let $v(x) = \psi(x) + \Psi(r)$. Since $|f(x, u, Dv)| \le \Lambda(1 + |D\psi| + |D\Psi(r)|)^2$, we then easily check that

$$L(v) - f(x, u(x), Dv) \le L\psi - \theta\lambda\rho^{-2}(r/\rho - 1 + \varepsilon)^{-2} + n\Lambda\theta\rho^{-1}(r/\rho - 1 + \varepsilon)^{-1} + 2\Lambda(1 + \Lambda^2)(1 + \theta^2\rho^{-2}(r/\rho - 1 + \varepsilon)^{-2}),$$

so selecting $\theta = (4\Lambda(1 + \Lambda^2))^{-1}\lambda$ we have

$$\begin{aligned} L(v) - f(x, u(x), Dv) &\le -\tfrac{1}{2}\lambda\rho^{-2}(r/\rho - 1 + \varepsilon)^{-2} + n\Lambda\theta\rho^{-1}(r/\rho - 1 + \varepsilon)^{-1} \\ &\qquad + 2\Lambda(1 + \Lambda^2) + \Lambda \\ &\le -\tfrac{1}{2}\lambda(\rho^2\varepsilon)^{-1} + n\Lambda(4\rho\sqrt{\varepsilon})^{-1} + 2\Lambda(1 + \Lambda^2) + \Lambda < 0 \end{aligned}$$

for $1 \le r/\rho \le 1 + \varepsilon^{1/2}$ and for $\varepsilon > 0$ sufficiently small, depending only on $\rho, \Lambda, M, \lambda$.

Thus we have $Lv \le f(x, u(x), Dv)$ on the domain $U \equiv \Omega \cap \{x : \rho < |x - x_0| < (1 + \varepsilon^{1/2})\rho\}$. Note also that $\Psi \ge 0$ everywhere on $\overline{U}$, and $\Psi = 0$ at the point $y \in \partial U$. In addition we have by direct computation that $v|\Omega \cap \partial B_{(1+\varepsilon^{1/2})\rho}(x_0) \ge \theta\log(1 + \varepsilon^{-1/2}) - M$, which we can arrange to be greater than M by choosing ε suitably. Thus we have then that $u \le v$ everywhere on ∂U. On the other hand, since $L(v - u) = f(x, u, Du) - f(x, u, Dv) = B(x).D(u - v)$ for a suitable bounded function B, by the maximum principle on U we then have $u \le v$ everywhere in $\overline{U}$. It then directly follows (since $v(y) = \psi(y) = u(y)$ by construction of v), that the directional derivative $D_\eta u(y)$ of u in direction of the inward pointing unit normal to $\partial\Omega$ at y is $\le$ the corresponding derivative of v, which in turn is bounded by a constant depending only on $M, \psi, \rho, \Lambda, \lambda$.

Since all the hypotheses (with the same constant Λ) hold with $-u$ in place of u and with $-f(x, -z, -p)$ in place of $f(x, z, p)$, we actually deduce then that $|D_\eta u(y)| \le C(M, \psi, \rho, \Lambda, \lambda)$, and since all the derivatives of u tangent to $\partial\Omega$ are bounded by M (because $u = \psi$ on $\partial\Omega$), we then have $|Du(y)| \le C$, where C depends only on M, Λ, λ, and ρ, and since we can take a fixed $\rho > 0$ which works for all points $y \in \partial\Omega$ (because the domain is given to be C^2), the proof is complete.

The second lemma is:

Lemma 2. *Suppose that u is any $C^{2,\mu}(\overline{\Omega})$ solution of the problem (‡), suppose that $|u|_{0,\Omega} \le M$, where M is a given constant, and suppose that* (i), (iii),(iv) *hold in case $n = 2$ and* (i),(iii),(iv),(v) *hold in case $n \ge 3$. Then*

$$|u|_{2,\mu,\Omega} \le C,$$

where C depends only on Λ, λ, Ω, and M in case $n = 2$ and also on the function δ_M of (v) *in case $n \ge 3$.*

Proof: The proof is another example of the "method of scaling." We are first going to prove that $|Du|$ is bounded by a fixed constant depending only on $M, \lambda, \Lambda, \Omega$. Indeed by Lemma 1 we already have such a bound on $\partial\Omega$:

$$(*) \qquad \sup_{\partial\Omega} |Du| \le C_0, \quad C_0 = C_0(M, \lambda, \Lambda, \Omega) > 0.$$

We begin by recalling the Schauder estimates of Lecture 12, namely if $B_\rho^\Omega(y) = \Omega \cap B_\rho(y)$, then there is $R_0 = R_0(\Omega) > 0$ such that for all $\rho \le R_0$, $y \in \overline{\Omega}$, and $\theta \in (0,1)$ the following estimate holds:

$$(1) \quad \begin{aligned} &\rho^{2+\mu}[D^2u]_{\mu, B_{\theta\rho}^\Omega(y)} + \rho^2|D^2u|_{0,B_{\theta\rho}^\Omega(y)} \\ &\quad \le C\,(|u|_{0,B_\rho^\Omega(y)} + \rho^2|f(\cdot,u(\cdot),Du(\cdot))|_{0,B_\rho^\Omega(y)} \\ &\qquad\qquad + \rho^{2+\mu}[f(\cdot,u(\cdot),Du(\cdot))]_{\mu,B_\rho^\Omega(y)}), \end{aligned}$$

where C depends only on Ω, θ, Λ (and not on ρ). On the other hand since $|u| \le M$ and since we have the structural conditions (iii), (iv), it is straightforward to check (see Exercise 15.3) that

$$(2) \quad \begin{aligned} &\rho^2|f(\cdot,u,Du)|_{0,B_\rho^\Omega(y)} + \rho^{2+\mu}[f(\cdot,u,Du)]_{\mu,B_\rho^\Omega(y)} \le \\ &\quad C\,(\rho^2(1+|Du|_{0,B_\rho^\Omega(y)})^2 + \rho^3(1+|Du|_{0,B_\rho^\Omega(y)})^3 \\ &\qquad\qquad + \rho^3(1+|Du|_{0,B_\rho^\Omega(y)})|D^2u|_{0,B_\rho^\Omega(y)}), \end{aligned}$$

where C depends only on M, Λ, Ω.

Now we agree to pick $\rho \in (0,1]$ small enough so that

$$(3) \qquad \rho|Du|_{0,B_\rho^\Omega(y)} \le 1,$$

and note that then (2) implies

(4)

$$\rho^2|f(\cdot,u,Du)|_{0,B_\rho^\Omega(y)} + \rho^{2+\mu}[f(\cdot,u,Du)]_{\mu,B_\rho^\Omega(y)} \le C\,(1+\rho^2|D^2u|_{0,B_\rho^\Omega(y)}),$$

with $C = C(M, \Omega, \Lambda)$. The interpolation inequality

$$\rho^2|D^2u|_{0,B_\rho^\Omega(y)} \le \varepsilon\rho^{2+\mu}[D^2u]_{\mu,B_\rho^\Omega(y)} + C|u|_{0,B_\rho^\Omega(y)},$$

in combination with (1) and (4), then gives

$$\rho^{2+\mu}[D^2u]_{\mu,B_{\theta\rho}^\Omega(y)} + \rho^2|D^2u|_{0,B_{\theta\rho}^\Omega(y)} \le C\,(1+\varepsilon\rho^{2+\mu}[D^2u]_{\mu,B_\rho^\Omega(y)}),$$

with $C = C(\varepsilon, \Omega, M, \Lambda)$, provided $\rho \in (0,1]$ is small enough to ensure (3). But then by the absorption lemma (Lemma 2 of Lecture 6) we deduce that in fact

$$(5) \qquad \rho^{2+\mu}[D^2u]_{\mu,B_{\theta\rho}^\Omega(y)} + \rho^2|D^2u|_{0,B_{\theta\rho}^\Omega(y)} \le C$$

for any $\rho \in (0,1]$ as in (3), with C a fixed constant determined by $M, \lambda, \Lambda, \Omega$, θ, and μ.

Now let $0 < \delta \le \min\{1, R_0, C_0^{-1}\}$ (C_0 as in $(*)$) be given, assume that $|Du|_{0,\Omega} > 1/\delta$, and let $\rho^{-1} = |Du|_{0,\Omega}$, so that in particular (3) holds. Since $|Du|_{0,\Omega} > C_0$, C_0 as in $(*)$, we see that $\max_\Omega |Du|$ is not attained on $\partial\Omega$, so we can select a point $x_0 \in \Omega$ with $|Du(x_0)| = |Du|_{0,\Omega}$. We now make the change of variable $x \mapsto \rho^{-1}(x - x_0)$, and let $v(x) = u(x_0 + \rho x)$ on the domain $\widetilde{\Omega} \equiv \{\rho^{-1}(x-x_0) : x \in \Omega\}$. Then we have

$$(6) \qquad |Dv(0)| = |Dv|_{0,\widetilde{\Omega}} = 1, \text{ and } \sup_{\partial\widetilde{\Omega}} |Dv| \le \delta M.$$

Also, since $\rho^{|\beta|}D^\beta u(x) = (D^\beta v)(\rho^{-1}(x-x_0))$ we deduce from (5), with $\theta = 1/2$ that

$$[D^2v]_{\mu,B_{1/2}^{\widetilde{\Omega}}(y)} + |D^2v|_{0,B_{1/2}^{\widetilde{\Omega}}(y)} \le C \quad \forall y \in \widetilde{\Omega},$$

where C is a fixed constant depending only on μ, M, Λ, Ω. Actually, since any ball of radius 1 with center in $\widetilde{\Omega}$ can be covered by a fixed number (depending only on n) of balls of radius $1/2$ with centers in $\widetilde{\Omega}$, we conclude that in fact

$$(7) \qquad [D^2v]_{\mu,B_1^{\widetilde{\Omega}}(y)} + |D^2v|_{0,B_1^{\widetilde{\Omega}}(y)} \le C \quad \forall y \in \widetilde{\Omega}$$

whenever $y \in \widetilde{\Omega}$.

Now suppose the result of Lemma 2 is false. Then for k sufficiently large we can apply the above with $1/k, x_k \in \Omega, \rho_k \le 1/k, a_{ij}^{(k)}, f_k$ in place of $\delta, x_0, \rho, a_{ij}, f$ respectively, where $a_{ij}^{(k)}, f_k$ satisfy conditions (i)–(iv) with fixed constants λ, Λ independent of k and also condition (v) with fixed function δ_M in case $n \ge 3$. Thus we get a sequence $v_k \in C^{2,\mu}(\Omega_k)$, where $\Omega_k \equiv \{\rho_k^{-1}(x-x_k) : x \in \Omega\}$ and where

$$(8) \qquad |Dv_k(0)| = |Dv_k|_{0,\Omega_k} = 1, \ |v_k|_{0,\Omega_k} \le M, \ \sup_{\partial\Omega_k} |Dv_k| \le k^{-1}C_0 \to 0$$

with C_0 as in $(*)$, and with

$$(9) \qquad [D^2v_k]_{\mu,\Omega_k\cap B_1(y)} + |D^2v_k|_{0,\Omega_k\cap B_1(y)} \le C, \quad \forall y \in \Omega_k,$$

where C is a constant independent of k.

Notice that v_k satisfies the equation

$$\widetilde{L}_k v_k = \widetilde{f}_k(x, v_k, Dv_k),$$

where $\widetilde{L}_k w = \sum_{i,j=1}^n a_{ij}^{(k)} D_iD_j w$ and

$$\widetilde{a}_{ij}^{(k)}(x) = a_{ij}^{(k)}(x_k + \rho_k x), \quad \widetilde{f}_k(x,z,p) = \rho_k^2 f_k(x_k + \rho_k x, z, \rho_k^{-1}p),$$

and using the conditions (i)–(v) for $a_{ij}^{(k)}$, f_k we deduce that $[\widetilde{a}_{ij}^k]_{\mu,\Omega_k\cap B_1(y)} \to$

0, and

(10)
$$\begin{cases} |\widetilde{f}_k(x,z,p)| \le \Lambda(\rho_k + |p|)^2 \\ |\widetilde{f}_k(x_2,z_2,p_2) - \widetilde{f}_k(x_1,z_1,p_1)| \le \Lambda\big(\rho_k^\mu(\rho_k + |p_1| + |p_2|)^2|x_2 - x_1|^\mu + \\ \qquad (\rho_k + |p_1| + |p_2|)^2|z_2 - z_1| + (\rho_k + |p_1| + |p_2|)|p_1 - p_2|\big) \end{cases}$$

for (x,z,p), (x_1,z_1,p_1), $(x_2,z_2,p_2) \in \widetilde{\Omega} \times [-M,M] \times \mathbb{R}^n$. (See Exercise 15.3 below.) In case $n \ge 3$, the additional hypothesis (v) implies that for $-M \le z_1 \le z_2 \le M$

(11) $$\widetilde{f}_k(x,z_2,p) - \widetilde{f}_k(x,z_1,p) \ge -\varepsilon_k(\rho_k + |p|)^2, \varepsilon_k \downarrow 0.$$

In view of the fact that each Ω_k is obtained from Ω by a translation composed with the homothety $x \mapsto \rho_k^{-1}x$, we conclude first that either the domain $\Omega_k \to \mathbb{R}^n$ in the sense that for each $R > 0$ there is k_R such that $\Omega_k \supset B_R(0)$ for all $k \ge k_R$, or else Ω_k tends to a half-space in the $C^{2,\mu}$-sense described in Lecture 12; we denote the limit domain by Ω_∞. (Thus Ω_∞ is either $\mathbb{R}^n$ or a half-space in $\mathbb{R}^n$.) Now in view of (8)–(11) we can apply the Arzela-Ascoli lemma to give a subsequence such that $v_k \to v$ locally in the C^2-norm on Ω_∞, where $v \in C^{2,\mu}(\overline{\Omega}_\infty)$ satisfies an equation of the form

(12) $$L_\infty v = f_\infty(v, Dv),$$

where $L_\infty w = \sum_{i,j=1}^n a_{ij}^{(\infty)} D_i D_j v$ with (a_{ij}^∞) a constant positive definite symmetric matrix and where $f_\infty(z,p)$ satisfies

(13)
$$\begin{cases} |f_\infty(z,p)| \le \Lambda |p|^2 \\ |f_\infty(z_1,p_1) - f_\infty(z_2,p_2)| \le \\ \qquad \Lambda\big((|p_1| + |p_2|)^2|z_1 - z_2| + (|p_1| + |p_2|)|p_1 - p_2|\big) \end{cases}$$

for z, z_1, $z_2 \in [-M,M]$ and p, p_1, $p_2 \in \mathbb{R}^n$. Also by (8) we know that v satisfies

(14) $$|Dv(0)| = |Dv|_{0,\Omega_\infty}, Dv = 0 \text{ on } \partial\Omega_\infty, |v|_{0,\Omega_\infty} \le M$$

Notice that after an orthogonal transformation of coordinates the equation (12) takes the form

(15) $$\Delta v = f_\infty(v, Dv)$$

where f_∞ satisfies (13) still. If Ω_∞ is a half-space we can use even reflection (since $|Dv| = 0$ on the boundary of the half-space) to deduce that we have an equation of the form (15) on all of $\mathbb{R}^n$, and (by (8)) v is not constant, but has bounded C^1-norm on $\mathbb{R}^n$. In case $n = 2$ the equation (15) has the form $\Delta v + q|Dv|^2 = 0$, with q bounded and with v bounded, hence by the maximum principle result of problem 15.6 below we deduce that $v \equiv$ const., contradicting (14), hence $(*)$ is established in case $n = 2$. On the other hand if $n \ge 3$ then the additional fact (11) implies that $f_\infty(z,p)$ is non-decreasing in z, and so choosing a unit vector $a \in \mathbb{R}^n$ such that $D_a v(0) = |Dv(0)|$, by differentiating the equation (15) we see that $w = D_a v$ is a $C^{1,\mu}(\mathbb{R}^n)$ weak solution of an equation of the form $\Delta w + b \cdot Dw + cw = 0$, with $c \le 0$. Hence, by the generalization of the Hopf maximum principle described in problem 15.5 below we again have $D_a v \equiv$ const. $\ne 0$, contradicting the boundedness of v, so again $(*)$ is established.

The fact that $|u|_{2,\mu,\Omega} \le C$ then follows from $(*)$ and (4) (keeping in mind that $B_\rho(y)$ is an arbitrary ball with $y \in \overline{\Omega}$ and $\rho < R$).

Next we discuss existence of solutions of (‡) in the special case when

$(*)$ $$f_z(x,z,p) \ge \varepsilon, \quad (x,z,p) \in \Omega \times \mathbb{R} \times \mathbb{R}^n,$$

where $\varepsilon > 0$ is any fixed constant. Then we have the following theorem, which we shall use in the proof of the main theorem (Theorem 3) in the case when $\varepsilon = 1$, but which is of some importance in itself:

Theorem 4. *Let $M = \varepsilon^{-1}\sup_\Omega |f(x,0,0)| + \max_{\partial\Omega} |\psi|$, and suppose that* (i), (iii), (iv), *and* $(*)$ *hold, and in case $n \ge 3$ assume in addition that* (v) *holds. Then there is a unique $C^{2,\mu}(\overline{\Omega})$ solution of* (‡).

Notes: (1) In this case, as for Lemma 2, we do not need to assume the existence of $\varphi^\pm$ as in condition (ii).

(2) Observe that Theorem 4 fails in case $\varepsilon = 0$ in $(*)$ (so in particular Theorem 3 fails without the condition (ii)), even when L is just the Laplacian operator in dimension 2 and $f(x,z,p)$ has no dependence on z. For example we claim the equation

$$\Delta u = 1 + |x|^2|Du|^2$$

(i.e. $L = \Delta$ and $f(x,z,p) \equiv 1 + |x|^2|p|^2$) has no $C^2(\overline{\Omega})$ solution on the disk $\Omega = \{x \in \mathbb{R}^2 : |x| < R\}$ with boundary data ψ which is constant on the boundary $|x| = R$ if $R^2 \ge \pi$. To see this first note that if u is a solution with such constant boundary data then so is $\widetilde{u} = u \circ Q$, where Q is any orthogonal transformation of $\mathbb{R}^2$. Then $v = u - \widetilde{u}$ satisfies the equation $\Delta v -$

$|x|^2(Du+D\widetilde{u})\cdot Dv = 0$ in Ω with $v=0$ on $\partial\Omega$, hence $v\equiv 0$ by the maximum principle. Thus $u=u(r)$, $r=|x|$. On the other hand one can explicitly find all such radially symmetric solutions: the equation in case $u=u(r)$ is just $(ru')'=r(1+(ru')^2)$, and since u is C^1 we must have $u'(0)=0$, so the ODE integrates to give

$$\arctan(ru') = r^2/2, \quad 0\le r\le R,$$

and hence $\{\arctan(ru') : r\in[0,R]\}$ is the closed interval $[0,R^2/2]$, which is impossible if $R^2\ge\pi$, because the arctan function has range equal to the open interval $(-\pi/2,\pi/2)$.

Proof of Theorem 4: For $\sigma\in[0,1]$, consider the problems

$$\begin{cases} Lu-\varepsilon u = \sigma\widetilde{f}(x,u,Du) \text{ in } \Omega \\ u=\sigma\psi \text{ on } \partial\Omega, \end{cases}$$

where $\widetilde{f}(x,z,p) = f(x,z,p)-\varepsilon z$. (Notice that then $\widetilde{f}(x,z,p)$ is then still increasing in z by $(*)$.)

We define operators $T_\sigma : C^{1,\mu}(\overline{\Omega})\to C^{2,\mu}(\overline{\Omega})\subset C^{1,\mu}(\overline{\Omega})$ by setting $T_\sigma(v)=u$, where u is the $C^{2,\mu}(\overline{\Omega})$ solution of the linear problem

$$\begin{cases} Lu-\varepsilon u = \sigma\widetilde{f}(x,v,Dv) \text{ in } \Omega \\ \qquad u=\sigma\psi \text{ on } \partial\Omega. \end{cases}$$

(By Lecture 12 such a u exists.) Next we claim that each $T(v,\sigma)=T_\sigma(v)$ is a continuous compact operator of $(v,\sigma)\in C^{1,\mu}(\overline{\Omega})\times[0,1]$ into $C^{2,\mu}(\overline{\Omega})\subset C^{1,\mu}(\overline{\Omega})$; we will check this in a more general context in the next lecture—suffice it to say for the moment that it is a direct consequence of the Schauder estimates of Lecture 12 and the fact that closed balls in $C^{2,\mu}(\overline{\Omega})$ are compact in $C^{1,\mu}(\overline{\Omega})$ by virtue of the Arzela-Ascoli lemma. Notice in particular $T(v,0)$ is the identically zero function on $\overline{\Omega}$, so $T_0=0$.

Now according to Lemma 2, we have $|u|_{2,\mu,\Omega}<C$, whenever $\sigma\in[0,1]$ and u is a solution of equation $T_\sigma(u)=u$, with C depending only on λ,Λ,Ω, and M, where M is any upper bound for $|u|$ on Ω. On the other hand $T_\sigma u=u$ means that

$$\begin{cases} Lu-\varepsilon u = \sigma\widetilde{f}(x,u,Du) \text{ in } \Omega \\ \qquad u=\sigma\psi \text{ on } \partial\Omega. \end{cases}$$

and an easy application of the maximum principle gives

$$\sup|u|\le\varepsilon^{-1}\sup_\Omega|f(x,0,0)|+\max_{\partial\Omega}|\psi|.$$

(See problem 15.8; use the result of that problem with σf in place of f and $\delta=\sigma\varepsilon$.)

Thus Lemma 2 is applicable with $M=\varepsilon^{-1}\sup_\Omega|f(x,0,0)|+\max_{\partial\Omega}|\psi|$ and hence we have, for all $\sigma\in[0,1]$, $|u|_{2,\mu,\Omega}\le C$, with C depending only on $\lambda,\Lambda,f,\Omega,\psi$, whenever u is a solution of $Lu-u=\sigma\widetilde{f}(x,u,Du)$ with $u=\sigma\psi$ on $\partial\Omega$. Thus all the conditions for the Leray-Schauder theorem hold with $\mathcal{C}=\mathcal{B}=C^{1,\mu}(\overline{\Omega})$, and hence we have a solution of the above problem with $\sigma=1$. This completes the proof of Theorem 4.

Proof of Theorem 3: Let $M>0$ be such that

$$-M\le\varphi^-\le\varphi^+\le M,$$

and define $\widetilde{f}(x,z,p)=f(x,z,p)-\beta z(1+|p|^2)$, where $\beta>0$ (depending on M and Λ) is chosen large enough to ensure that

$$D_z\widetilde{f}(x,z,p)\le 0, \quad (x,z,p)\in\overline{\Omega}\times[-M,M]\times\mathbb{R}^n. \tag{2}$$

Notice that this can be done because of (iii), (iv). With such a β and $\widetilde{f}$, and for a given $v\in C^{1,\mu}(\overline{\Omega})$ and $\sigma\in[0,1]$, consider the problem

$$\begin{cases} Lu=\beta u(1+|Du|^2)+\widetilde{f}(x,\varphi^-+\sigma(v-\varphi^-),Du) \text{ in } \Omega \\ u=\psi \text{ on } \partial\Omega. \end{cases} \tag{3}$$

Since the right side in (3) is has the form $\hat{f}(x,u,Du)$ with $\hat{f}_z(x,z,p)\ge 1$, we can apply Theorem 4 (with $\varepsilon=1$ in the condition $(*)$ there), and hence a $C^{2,\mu}(\overline{\Omega})$ solution of (3) exists, and (by the maximum principle) it is unique. We can therefore define the operators $T_\sigma : C^{1,\mu}(\overline{\Omega})\to C^{2,\mu}(\overline{\Omega})\subset C^{1,\mu}(\overline{\Omega})$ by setting $T_\sigma(v)$ to be equal to the unique $C^{2,\mu}(\overline{\Omega})$ solution of this problem. Furthermore (see the general discussion of the next lecture) $T : C^{1,\mu}(\overline{\Omega})\times[0,1]\to C^{1,\mu}(\overline{\Omega})$ defined by $T(v,\sigma)=T_\sigma(v)$ is a continuous compact operator.

Now set $\mathcal{B}=\{v\in C^{1,\mu}(\overline{\Omega}) : v=0 \text{ on } \partial\Omega\}$, and $\mathcal{C}=\{v\in\mathcal{B} : \varphi^-\le v\le\varphi^+ \text{ in } \overline{\Omega}\}$. We claim that $T:\mathcal{C}\times[0,1]\to\mathcal{C}$. The argument for this is based on the maximum principle as follows:

Suppose $v\in\mathcal{C}$, and let $\sigma\in[0,1]$ and $u=T_\sigma(v)$. Then (3) holds for u, so

$$\begin{aligned} Lu=\beta u(1+|Du|^2)+\widetilde{f}(x,\varphi^-+\sigma(v-\varphi^-),Du)\le \\ \beta u(1+|Du|^2)+\widetilde{f}(x,\varphi^-,Du), \end{aligned} \tag{4}$$

by virtue of the facts that $\widetilde{f}(x,z,p)$ is decreasing in z and $v\ge\varphi^-$ because $v\in\mathcal{C}$. In view of (4) and the fact that $L\varphi^-\ge f(x,\varphi^-,D\varphi^-)\equiv\beta\varphi^-(1+$

$|D\varphi^-|^2) + \widetilde{f}(x, \varphi^-, D\varphi^-)$, we then have

$$L(\varphi^- - u) - \beta(1 + |Du|^2)(\varphi^- - u) - \beta\varphi^-(Du + D\varphi^-)\cdot D(\varphi^- - u) \geq \widetilde{f}(x, \varphi^-, D\varphi^-) - \widetilde{f}(x, \varphi^-, Du),$$

and hence we have an inequality of the form

$$L(\varphi^- - u) + b\cdot D(\varphi^- - u) + c(\varphi^- - u) \geq 0,$$

where b, c are bounded and $c \leq 0$. Also $\varphi^- - u \leq 0$ on $\partial\Omega$, hence by the weak maximum principle we have $\varphi^- - u \leq 0$ everywhere in Ω.

Using $\varphi^- \leq v \leq \varphi^+$, and again using the fact that $f(x, z, p)$ is decreasing in z, we also have

$$Lu - \beta u(1 + |Du|^2) = \widetilde{f}(x, \varphi^- + \sigma(v - \varphi^-), Du) = \widetilde{f}(x, (1-\sigma)\varphi^- + \sigma v, Du) \geq \widetilde{f}(x, \varphi^+, Du),$$

so a similar maximum principle argument establishes $u \leq \varphi^+$ everywhere in Ω.

Thus we have $T : \mathcal{C} \times [0,1] \to \mathcal{C}$ as claimed.

Next note that if $\sigma \in [0,1]$ and if $u \in \mathcal{C}$ satisfies $T_\sigma(u) = u$, then $Lu = \beta u(1 + |Du|^2) + \widetilde{f}(x, \varphi^- + \sigma(u - \varphi^-), Du)$ and (since $u, \varphi^- \in \mathcal{C}$ and hence $|\varphi^- + \sigma(u - \varphi^-)| \leq \sigma|u| + (1-\sigma)|\varphi^-| \leq \sigma M + (1-\sigma)M = M$) we can apply Lemmas 1,2 (with $\beta z(1 + |p|^2) + \widetilde{f}(x, \varphi^- + \sigma(z - \varphi^-), p)$ in place of $f(x, z, p)$) to give $|u|_{2,\mu,\Omega} \leq C$ and hence

$$|u|_{1,\mu,\Omega} \leq C,$$

where C depends only on $\Lambda, M, \lambda, \psi, \Omega$. Also $T_0(v) \equiv u_0$, where u_0 is the unique solution of $Lu = \beta u(1 + |Du|^2) + \widetilde{f}(x, \varphi^-, Du)$ in Ω with $u = \psi$ on $\partial\Omega$. Thus all the conditions for application of the Leray-Schauder Theorem (Theorem 1) are satisfied, hence there is a $u \in \mathcal{C}$ such that $T_1 u = u$. Thus $u \in C^{2,\mu}(\overline{\Omega})$ and $Lu = \beta u(1 + |Du|^2) + \widetilde{f}(x, u, Du) \equiv f(x, u, Du)$, so u is the required solution.

LECTURE 15 PROBLEMS

15.1 If $\mathcal{B}$ is a Banach space and if $\mathcal{C}$ is an open convex subset of $\mathcal{B}$ containing 0, prove that for each $x \in \mathcal{B} \setminus \mathcal{C}$ there is a unique $\sigma(x) \in (0,1]$ such that $\sigma(x)x \in \partial\mathcal{C}$. Prove also that $\sigma(x)$ is continuous on $\mathcal{B} \setminus \mathcal{C}$.
Hint: If $x \in \mathcal{B}\setminus\{0\}$ and if $\rho > 0$, prove that the set $\{ty + (1-t)x : y \in B_\rho(0),\ t \in (0,1)\}$ is an open set containing the line segment $\ell = \{tx : t \in (0,1)\}$. Here $B_\rho(0)$ is the open ball center 0 and radius ρ.

15.2 Check the claim that $u \leq v$ on U in the proof of Lemma 1.

15.3 Check the inequalities (2) and (10) in the proof of Lemma 2.

15.4 If Ω is a bounded $C^{2,\mu}$ domain in $\mathbb{R}^2$ and if $\psi \in C^{2,\mu}(\overline{\Omega})$ prove that there is a $C^{2,\mu}(\overline{\Omega})$ solution of the following problems:

(a) $$\begin{cases} \Delta u = (u^3 - 3u^2)|Du|^2 + 1 \text{ in } \Omega \\ u = \psi \text{ on } \partial\Omega. \end{cases}$$

(b) $$\begin{cases} \Delta u = f(x, u, Du) \text{ in } \Omega \\ u = \psi \text{ on } \partial\Omega, \end{cases}$$

where in (b) f satisfies (iii), (iv) and there exist $v_1,\ v_2 \in \mathbb{R}^2$ such that $f(x, z, v_1) \to +\infty$ as $z \to +\infty$ and $f(x, z, v_2) \to -\infty$ as $z \to -\infty$, uniformly for $x \in \overline{\Omega}$.

15.5 Prove that the strong maximum principle is valid for C^1 weak solutions of equations of the form $\Delta u + b\cdot Du + cu = 0$, with b, c bounded and $c \leq 0$.

Hint: Show that the proof of the Hopf boundary point lemma applies essentially as before, except that at the last step we use the weak maximum principle for weak solutions in place of the weak maximum principle for classical solutions.

15.6 Suppose u is a bounded $C^1(\mathbb{R}^2)$ weak solution of the equation $\Delta u + q|Du|^2 = 0$, where q is bounded on $\mathbb{R}^2$, and let β be such that $\sup|q|, \sup|u| \leq \beta$.

(i) Prove $\int_{B_R(0)} |Du|^2\zeta^2 \leq C\int_{B_R(0)} |D\zeta|^2$ for each $\zeta \in C_c^\infty(B_R(0))$, $C = C(\beta)$.

Hint: The weak form of the equation is $\int_{\mathbb{R}^2} Du\cdot D\zeta - q|Du|^2\zeta = 0$ for any $\zeta \in C_c^\infty(\mathbb{R}^2)$; replace ζ by $e^{(\beta+1)u}\zeta^2$.

(ii) Prove $u \equiv$ const. Hint: Use the "logarithmic cut-off trick" in (i); i.e. use the Lipschitz functions $\zeta_R(x) = (\log R)^{-1}\max\{\log(\min\{(R/r)^2, R\}), 0\}$, $r = |x|$, for $R > 1$ and let $R \to \infty$.

15.7 If $\varphi^\pm$ are as in the above lecture, and φ^+ is not identically equal to φ^-, prove that $\varphi^+ > \varphi^-$ everywhere in Ω and $D_\eta \varphi^+ > D_\eta \varphi^-$ at each boundary point where $\varphi^+ = \varphi^-$.
Hint: The strong maximum principle holds without any condition on the sign of "c" in case the maximum value is zero.

15.8 If $u \in C^2(\Omega) \cap C^0(\overline{\Omega})$, with Ω bounded and $u = \psi$ on $\partial\Omega$, is a solution of $Lu - \delta u = f(x, u, Du)$, with f as in (iii), (iv) above, $\delta > 0$ and with $D_z f(x, z, p) \geq 0$, prove that $\sup |u| \leq \delta^{-1} \sup_\Omega |f(x, 0, 0)| + \max_{\partial\Omega} |\psi|$.

Hint: The equation can be written $Lu - (f(x, u, Du) - f(x, 0, 0)) = (\delta u + f(x, 0, 0))$. Show that then $\sup |u| > \delta^{-1} \sup |f(x, 0, 0)| + \max |\psi|$ contradicts the maximum principle for either u or $-u$.

Lecture 16

Quasilinear Problems

A serious defect of the discussion so far is that the classes of problems described do not include (except for some relatively trivial special cases) equations which arise as Euler-Lagrange equations of a given functional. More specifically, suppose we have a bounded domain $\Omega \subset \mathbb{R}^n$ and a functional $\mathcal{F}(u) = \int_\Omega F(x, u, Du)$, where F is a given C^2 function of the variables $(x, z, p) \in \overline{\Omega} \times \mathbb{R} \times \mathbb{R}^n$. Then (see the discussion of Lecture 1), the Euler-Lagrange equation is

$$\sum_{i,j=1}^n F_{p_i p_j}(x, u, Du) D_i D_j u = f(x, u, Du),$$

where $f(x, z, p) = F_z(x, z, p) - \sum_{j=1}^n (p_j F_{p_j z}(x, z, p) - F_{p_j x^j}(x, z, p))$. Notice that this has the general form

$$Q(u) = f(x, u, Du),$$

where $Q(u) = \sum_{i,j=1}^n a_{ij}(x, u, Du) D_i D_j u$; i.e. an operator which is linear in the second derivatives, with coefficients given functions of (x, u, Du). Such operators are called quasilinear operators, and are said to be elliptic if the matrix $(a_{ij}(x, z, p))$ is positive definite at each point of $\overline{\Omega} \times \mathbb{R} \times \mathbb{R}^n$. Notice that for the special class arising as Euler-Lagrange operators of a functional $\mathcal{F}$ as described above, the condition of ellipticity is equivalent to the requirement that the matrix $(F_{p_i p_j}(x, z, p))$ be positive definite for each (x, z, p), or in other words that $F(x, z, p)$ be locally uniformly convex in the variable p.

The operator is called uniformly elliptic if there are fixed constants $\lambda, \Lambda > 0$ such that

$$\lambda|\xi|^2 \leq \sum_{i,j=1}^n a_{ij}(x, z, p)\xi_i \xi_j \leq \Lambda|\xi|^2, \quad (x, z, p) \in \overline{\Omega} \times \mathbb{R} \times \mathbb{R}^n, \ \xi \in \mathbb{R}^n.$$

In many of the more interesting applications, such uniform ellipticity will not hold. For example the minimal surface operator mentioned in Lecture 1 has coefficients $a_{ij}(x,z,p) = \delta_{ij} - (1+|p|^2)^{-1}p_i p_j$ and one easily checks that the minimum eigenvalue of the matrix of these coefficients is $(1+|p|^2)^{-1}$ and the maximum eigenvalue is 1, so that the operator is not uniformly elliptic.

We want to show how the general Leray-Schauder procedure described in the previous lecture applies to quasilinear elliptic problems, i.e. problems involving operators of the form $(*)$ above.

We in fact consider the quasilinear Dirichlet problem

$$(\mathrm{P}) \qquad \begin{cases} Q(u) = f(x,u,Du) \text{ in } \Omega \\ u = \psi \text{ on } \partial\Omega, \end{cases}$$

where Ω is a bounded $C^{2,\mu}$ domain and $\psi \in C^{2,\mu}(\overline{\Omega})$, and where the operator $Q(u)$ is as above, and we assume $a_{ij}(x,z,p)$, $f(x,z,p)$ are Hölder continuous with respect to the x variable and locally Lipschitz with respect to the z and p variables. Specifically we assume that for each $K>0$ there is a constant Λ_K such that

$$(*) \quad \begin{aligned} &|g(x_2,z_2,p_2) - g(x_1,z_1,p_1)| \le \\ &\quad \Lambda_K(|x_1-x_2|^\mu + |z_1-z_2| + |p_1-p_2|), \quad |z_1|,|z_2|,|p_1|,|p_2| \le K,\ x \in \overline{\Omega} \end{aligned}$$

for each of the choices $g = a_{ij}$, $g = f$, and we also assume that for each K there is $\lambda_K > 0$ with

$$(**) \qquad \lambda_K|\xi|^2 \le \textstyle\sum_{i,j} a_{ij}\xi_i\xi_j \le \Lambda_K|\xi|^2, \quad x \in \overline{\Omega},\ |z|,|p| < K;$$

thus the equation in (P) is elliptic but not necessarily uniformly elliptic.

We can use the Leray-Schauder theory as in the previous lecture to discuss existence for this problem under certain restrictions on the coefficient functions a_{ij} and on the given function f. We begin with a simple scheme which gives some of the important quasilinear existence results.

The general idea here is to set up a map $T : C^{1,\mu(\overline{\Omega})\times[0,1]} \to C^{1,\mu}(\overline{\Omega})$, and corresponding maps $T_\sigma(v) = T(v,\sigma)$, by defining $T_\sigma(v)$ to be the solution u of the linear problem

$$(\ddagger) \qquad \begin{cases} \sum_{i,j=1}^n a_{ij}(x,v,Dv)D_iD_ju = \sigma f(x,v,Dv) \text{ in } \Omega \\ u = \sigma\psi \text{ on } \partial\Omega, \end{cases}$$

where a_{ij}, f are functions of the variables $(x,z,p) \in \overline{\Omega}\times\mathbb{R}\times\mathbb{R}^n$ which are locally Hölder continuous in the x variables, continuous in the σ variables, and and which have bounded second derivatives with respect to the variables (z,p) on $\mathbb{R}\times\mathbb{R}^n$.

Notice that the conditions $(*)$ guarantee

$$|a_{ij}(x,v,Dv)|_{0,\mu,\Omega}, |f(x,v,Dv)|_{0,\mu,\Omega} \le C(\Lambda_K,\Omega)$$

whenever $|v|_{1,\mu,\Omega} \le K$, and hence, by the Schauder theory of Lecture 12 and the Fredholm Alternative (FA) described in the supplement to Lecture 12, and the maximum principle (which eliminates the first alternative in (FA)), such u exists and

$$(\ddagger\ddagger) \qquad |u|_{2,\mu,\Omega} \le C(|f(\cdot,v,Dv)|_{0,\mu,\Omega} + |\psi|_{2,\mu,\Omega}) \le C,$$

where $C = C(K,\Omega)$. This shows in particular that T takes bounded subsets of $C^{1,\mu}(\overline{\Omega})\times[0,1]$ into bounded subsets of $C^{2,\mu}(\overline{\Omega})$, and such sets are evidently compact in in $C^{1,\mu}(\overline{\Omega})$ (in fact they are compact in $C^2(\overline{\Omega})$ by Arzela-Ascoli), so T is a compact operator. It is also continuous, because if $v_k \to v$ in $C^{1,\mu}(\overline{\Omega})$ and $\sigma_k \to \sigma$ in $[0,1]$, then $u_k = T(v_k,\sigma_k)$ is bounded in $C^{2,\mu}(\overline{\Omega})$ by $(\ddagger\ddagger)$, and hence has a subsequence u_{k_j} which converges in $C^2(\overline{\Omega})$ to a $C^{2,\mu}(\overline{\Omega})$ solution u of the equation $(\ddagger)$; i.e. $u = T_\sigma(v)$ (unique, independent of the sequences v_k,σ_k and subsequence u_{k_j} chosen), and hence the passage to a subsequence was not necessary, and the continuity of T is established.

Hence the Leray-Schauder theorem gives the required fixed point of T (and hence the required $C^{2,\mu}(\overline{\Omega})$ solution of the problem (P)) provided only that any $u \in C^{2,\mu}(\overline{\Omega})$ which satisfies (P_σ) for some $\sigma > 0$ obeys the *a-priori* estimate

$$(\dagger) \qquad |u|_{1,\mu,\Omega} < M,$$

for some fixed constant M. Checking $(\dagger)$ in a given case is the hard part of the Leray–Schauder procedure.

For the moment we want to point out that it is possible to reduce the problem of proving an *a-priori* bound like $(\dagger)$ to the formally simpler problem of showing that there is a fixed constant K such that

$$(\dagger\dagger) \qquad |u|_{0,\Omega} + |Du|_{0,\Omega} \le K$$

whenever u is a $C^{2,\mu}(\overline{\Omega})$ solution of one of the problems (P_σ), $\sigma \in [0,1]$. This follows because of the following general theorem:

Theorem 1. *Suppose $u \in C^{2,\mu}(\overline{\Omega})$ satisfies the equation*

$$\textstyle\sum_{i,j=1}^n a_{ij}(x,u,Du)D_iD_ju = f(x,u,Du)$$

on the bounded $C^{2,\mu}$ domain Ω, suppose $u = \psi$ on $\partial\Omega$, where $\psi \in C^{2,\mu}(\overline{\Omega})$, and suppose that for each K there are $\lambda_K, \Lambda_K > 0$ such that $(*)$, $(**)$ hold. Then

$$|u|_{0,\Omega} + |Du|_{0,\Omega} + |\psi|_{2,\mu,\Omega} \le K \Rightarrow |u|_{1,\mu,\Omega} \le C, \quad C = C(\Omega,\mu,K,\lambda_K,\Lambda_K).$$

We give the proof of this below in the case $n = 2$; the proof of Theorem 1 for the case $n \ge 3$ will be deferred until we have developed De Giorgi-Nash theory. Notice that, by applying this result to the solutions u of $T_\sigma(u) = u$, we get, as we claimed above, the following:

Corollary. *Under the conditions imposed on $a_{ij}(x,z,p)$, $f(x,z,p)$ in the discussion following* (P) *above, we have* $(\dagger\dagger) \Rightarrow (\dagger)$*, and hence (by the Leray-Schauder Theorem),* $(\dagger\dagger) \Rightarrow$ *there is a $u \in C^{2,\mu}(\overline{\Omega})$ satisfying problem* (P).

The principal ingredient needed in the proof of Theorem 1 in case $n = 2$ is given in the following theorem. Notice that for this result no smoothness on a_{ij}, f is needed.

Theorem 2 (C. B. Morrey). *If $n = 2$, if u is any $C^2(\overline{\Omega})$ solution of an equation of the form*

$$\textstyle\sum_{i,j=1}^2 a_{ij}D_iD_ju = f \text{ on } \Omega,$$

where $\sum_{i,j=1}^2 a_{ij}\xi_i\xi_j \ge |\xi|^2$, $|a_{ij}| \le L_1$, and $|f| \le L_2$ (L_1, L_2 any fixed constants), and if $\partial\Omega \in C^2$ and $u = 0$ on $\partial\Omega$, then

$$|u|_{1,\beta,\Omega} \le C,$$

where C depends only on Ω, L_1, and L_2, and $\beta \in (0,1)$ depends only on L_1, Ω.

Now, as a first step in the proof of Theorem 2, we make the observation that each of the partial derivatives D_1u, D_2u of u must be a weak solution of a divergence-form elliptic equation, as follows.

Just for the moment, assume that $u \in C^3$ and a_{ij}, f are C^1. Write $a_{11} = a$, $a_{12} = a_{21} = b, a_{22} = c$, and note that then the equation can be written in the form

$$u_{11} + 2a^{-1}bu_{12} + a^{-1}cu_{22} = a^{-1}f,$$

where the subscripts denote partial derivatives. Then by differentiation with respect to the x^2 variable, we obtain

$$D_1u_{21} + 2D_2(a^{-1}bu_{21}) + D_2(a^{-1}cu_{22}) = D_2(a^{-1}f),$$

and (see Exercise 16.1) one can easily check that this holds in the weak sense, even if u is only C^2 and a_{ij}, f are merely bounded. Thus we have an equation for $v = D_2u$ of the form

$$(*) \qquad \int_\Omega \textstyle\sum_{i,j=1}^2 \alpha_{ij}D_ivD_j\zeta = \int_\Omega \sum_{j=1}^2 f_jD_j\zeta$$

for each $\zeta \in C_c^1(\Omega)$, where $\alpha_{11} = 1$, $\alpha_{12} = 2a_{11}^{-1}a_{12}$, $\alpha_{21} = 0$, $\alpha_{22} = a_{11}^{-1}a_{22}$, $f_1 = a_{11}^{-1}f$, and $f_2 = 0$. Notice that then we have $\sum\alpha_{ij}\xi_i\xi_j = a_{11}^{-1}\sum a_{ij}\xi_i\xi_j$, so since $a_{11} \ge 1$ we have

$$(**) \qquad \textstyle\sum_{i,j=1}^2 \alpha_{ij}\xi_i\xi_j \ge L_1^{-1}|\xi|^2 \ \forall\xi \in \mathbb{R}^2, \quad |\alpha_{ij}| \le L_1, \quad |f_j| \le L_2.$$

(Of course $v = D_1u$ satisfies a similar equation.)

Concerning the solutions of such equations we have the following result, which we will show directly implies Theorem 2.

Theorem 3. *Suppose $n = 2$, $B_R(0) \subset \Omega$, and $v \in H^1(B_R(0))$ satisfies an equation of the form $(*)$ above, where $(**)$ holds. Then*

$$|v(x) - v(y)| \le C\Big(\frac{|x-y|}{R}\Big)^\beta \big(R^{-1}\|v\|_{L^2(B_R(0))} + L_2R\big), \quad x, y \in B_{R/2}(0),$$

where $C > 0$ and $\beta \in (0,1)$ depend only on L_1.

If the hypothesis $B_R(0) \subset \Omega$ is replaced by the hypotheses that $B_R(0) \cap \Omega = B_R^+ \equiv \{x \in B_R(0) : x^2 > 0\}$ and $v = 0$ on $T \equiv \{x \in B_R(0) : x^2 = 0\}$ in the sense that $\varphi v \in H_0^1(B_R^+)\ \forall\varphi \in C_c^\infty(B_R(0))$, then the conclusion continues to hold with B_R^+, $B_{R/2}^+$ in place of B_R, $B_{R/2}$ respectively.

Before we begin the proof, we need to establish two lemmas, both of which are valid in any dimension $n \ge 2$:

Lemma 1 (Poincaré Inequality). *For any connected and bounded Lipschitz domain Ω, there is a constant C depending only on Ω such that*

$$\int_\Omega |u - \overline{u}|^2\,dx \le C\int_\Omega |Du|^2\,dx, \quad \forall u \in H^1(\Omega),$$

where $\overline{u} = |\Omega|^{-1}\int_\Omega u(x)\,dx$, with $|\Omega|$ denoting the Lebesgue measure of Ω.

Proof: Recall Rellich's theorem for Lipschitz domains—if $\{u_k\} \subset H^1(\Omega)$ and if $\|u_k\|_{1,\Omega} \le K$, with K independent of k, then there is a subsequence $u_{k'}$ converging strongly in $L^2(\Omega)$ and weakly in $H^1(\Omega)$ to a $H^1(\Omega)$ function u.

Now if the lemma is false, there is a sequence $u_k \in H^1(\Omega)$ with $\int_\Omega u_k = 0$, $\int_\Omega |Du_k|^2 \to 0$, and $\int_\Omega u_k^2 = 1\ \forall k$. Using the above Rellich theorem there is a subsequence converging weakly in $H^1(\Omega)$ and strongly in $L^2(\Omega)$ to an $H^1(\Omega)$ function u with $\int_\Omega u = 0$, $\int_\Omega u^2 = 1$ and with $\int_\Omega |Du|^2 = 0$, so that $u =$ constant. (Notice that we used the fact that Ω is connected here.) However $u =$ constant contradicts the fact that $\int_\Omega u = 0$ and $\int_\Omega u^2 \neq 0$.

We also need the following lemma, due to C. B. Morrey.

Lemma 2 (C. B. Morrey.) *Suppose $\Omega \subset \mathbb{R}^n$ is convex and open with diameter $= R$, $u \in H^1(\Omega)$, and*

$$(*) \qquad \int_{B_\sigma(y)\cap\Omega} |Du| \leq \beta\sigma^{n-1+\mu},$$

whenever $y \in \overline{\Omega}$ and $0 < \sigma \leq R$, where $\mu \in (0,1]$ and $\beta > 0$ are constants. Then

$$|u(x) - u(y)| \leq C\beta|x-y|^\mu, \quad x, y \in \Omega,$$

where C depends only on Ω.

Remark 1. Since (by Hölder's inequality)

$$\int_{B_\sigma(y)\cap\Omega} |Du| \leq (\omega_n\rho^n)^{1-1/p} \|Du\|_{L^p(B_\sigma(y)\cap\Omega)} \ \forall\, p \geq 1,$$

the hypothesis $(*)$ is satisfied with $C(\Omega, p)\beta$ in place of β, provided $\exists\, p \geq 1$ with

$$\int_{B_\sigma(y)\cap\Omega} |Du|^p \leq \beta^p\sigma^{n-p+p\mu}, \quad \forall\, y \in \overline{\Omega},\ 0 < \sigma \leq R.$$

(The case $p = 2$ is especially important—it is the case we are going to use in the proof of Theorem 3.)

Proof of Lemma 2: Recall the spherical integration formula for convex domains: If $f \in C^0(\overline{\Omega})$, Ω convex, if $y \in \Omega$, and if, for each $\omega \in S^{n-1}$, $\rho(\omega)$ denotes the maximum value of $\rho > 0$ such that $y + \rho\omega \in \overline{\Omega}$, then

$$(1) \qquad \int_\Omega f(x)\,dx = \int_{S^{n-1}} \int_0^{\rho(\omega)} f(y + r\omega) r^{n-1}\, dr d\omega.$$

Assume for the moment that $u \in C^1(\overline{\Omega})$, and note that if $z \in \Omega$, if $0 < \sigma \leq R/2$, and if $x, y \in \Omega \cap B_\sigma(z)$, then by integration along the segment joining y to x we obtain

$$u(y) - u(x) = \int_0^{|y-x|} \omega \cdot Du(y + \tau\omega)\, d\tau,$$

where $\omega = |y-x|^{-1}(y-x)$. Then letting $\sigma(\omega)$ denote the largest $\rho > 0$ such that $y + \rho\omega \in \overline{B}_\sigma(z) \cap \overline{\Omega}$, and integrating with respect to x over $\Omega \cap B_\sigma(z)$,

we obtain

$$\begin{aligned}
|\Omega \cap B_\sigma(z)||u(y) - u_{z,\sigma}| &\leq \int_{\Omega\cap B_\sigma(z)} \int_0^{\sigma(\omega)} |Du(y+\tau\omega)|\, d\tau dx \\
&= \int_{S^{n-1}} \int_0^{\sigma(\omega)} \int_0^{\sigma(\omega)} |Du(y+\tau\omega)|\, d\tau r^{n-1}\, dr d\omega \\
&\leq \frac{(2\sigma)^n}{n} \int_{S^{n-1}} \int_0^{\sigma(\omega)} |Du(y+\tau\omega)|\, d\tau d\omega \\
&= \frac{(2\sigma)^n}{n} \int_{\Omega\cap B_\sigma(z)} \frac{|Du(x)|}{|y-x|^{n-1}}\, dx,
\end{aligned}$$

and this is valid for all $y \in \Omega \cap B_\sigma(z)$. On the other hand, using the given inequality $(*)$ for u, we deduce that

$$\int_{\Omega\cap B_\sigma(z)} \frac{|Du(x)|}{|y-x|^{n-1}}\, dx \leq \int_{\Omega\cap B_{2\sigma}(y)} \frac{|Du(x)|}{|y-x|^{n-1}}\, dx \leq C\beta\sigma^\mu,$$

where $C = C(\mu)$. Since $|B_\sigma(z) \cap \Omega| \geq C\sigma^n$ for all $0 < \sigma \leq R$ (where C depends only on Ω), we thus deduce that

$$|u(y) - u_{z,\sigma}| \leq C\beta\sigma^\mu, \quad y \in \Omega \cap B_\sigma(z),$$

where $u_{z,\sigma} = |\Omega \cap B_\sigma(z)|^{-1} \int_{\Omega\cap B_\sigma(z)} u$. By the triangle inequality this gives

$$|u(y) - u(x)| \leq 2C\beta\sigma^\mu, \quad x, y \in \Omega \cap B_\sigma(z), \quad \sigma < R/2,$$

which, in view of the arbitrariness of z, is the required result.

Proof of Theorem 3: We begin by substituting $\zeta = (v - \lambda)\varphi^2$ in $(*)$, where λ is a constant to be chosen and $\varphi \in C_c^\infty(\Omega)$. This gives

$$\int_\Omega \sum_{i,j=1}^2 \varphi^2 \alpha_{ij} D_i v D_j v = \\ -2\int_\Omega \sum_{i,j=1}^2 \varphi\alpha_{ij} D_i v D_j \varphi + \int_\Omega \sum_{j=1}^n f_j D_j\big((v-\lambda)\varphi^2\big),$$

and hence, by $(**)$,

$$\int_\Omega |Dv|^2\varphi^2 \leq 4L_1 \int_\Omega \varphi|Dv||D\varphi||v-\lambda| + 2L_2 \int_\Omega (\varphi|v-\lambda||D\varphi| + \varphi^2|Dv|).$$

By virtue of the Schwarz inequality this gives

$$\int_\Omega |Dv|^2\varphi^2 \leq 64(1 + L_1^2) \int_\Omega |D\varphi|^2(v-\lambda)^2 + 16L_2^2 \int_\Omega \varphi^2.$$

Now suppose that $B_\rho(y)$ is an arbitrary ball contained in $B_R(0)$, and choose φ to be a cut-off function relative to the balls $B_\rho(y)$, $B_{\rho/2}(y)$. In fact let us assume that $\varphi \equiv 1$ in $B_{\rho/2}(y)$, $\varphi \equiv 0$ outside $B_\rho(y)$, $0 \le \varphi \le 1$ everywhere, and $|D\varphi| \le 3/\rho$ everywhere.

Then (1) gives

$$(2) \qquad \int_{B_{\rho/2}(y)} |Dv|^2 \le C_1\rho^{-2}\int_{B_\rho(y)\setminus B_{\rho/2}(y)} (v-\lambda)^2 + C_2L_2^2\rho^2,$$

where C_1 depends only on L_1 and C_2 is an absolute constant.

Now notice that by change of scale (see Exercise 16.2 below), the Poincaré inequality of Lemma 1 above gives

$$\int_{B_\rho(y)\setminus B_{\rho/2}(y)} (v-\lambda)^2 \le C\rho^2\int_{B_\rho(y)\setminus B_{\rho/2}(y)} |Dv|^2,$$

where $\lambda = |B_\rho(y)\setminus B_{\rho/2}(y)|^{-1}\int_{B_\rho(y)\setminus B_{\rho/2}(y)} v$ and where C is a fixed constant independent of ρ. Thus, taking $\lambda = |B_\rho(y)\setminus B_{\rho/2}(y)|^{-1}\int_{B_\rho(y)\setminus B_{\rho/2}(y)} v$ in (2),

$$(3) \qquad \int_{B_{\rho/2}} |Dv|^2 \le C_1\int_{B_\rho(y)\setminus B_{\rho/2}(y)} |Dv|^2 + C_2L_2^2\rho^2,$$

where C_1 depends on L_1 and C_2 is an absolute constant.

Adding $C_1\int_{B_{\rho/2}(y)} |Dv|^2$ to each side of (3) we obtain

$$(4) \qquad \int_{B_{\rho/2}(y)} |Dv|^2 \le \theta_0\int_{B_\rho(y)} |Dv|^2 + C_2L_2^2\rho^2,$$

where $\theta_0 = \frac{C_1}{1+C_1} \in (0,1)$ is a fixed constant (depending only on L_1). Notice that this implies, with $\theta = \max\{\theta_0, 1/2\}$

$$F(\rho/2) \le \theta F(\rho),$$

where $F(\sigma) = \int_{B_\sigma(y)} |Dv|^2 + CL_2^2\sigma^2$, with C depending only on L_1 and L_2R. By iteration, this readily leads to

$$F(\rho/2^j) \le \theta^j F(\rho/2), \quad j \ge 1,$$

and, since any $\sigma \in (0,\rho)$ satisfies $\rho/2^j \le \sigma < \rho/2^{j-1}$ for some $j \ge 1$, this gives

$$F(\sigma) \le C(\sigma/\rho)^\beta F(\rho/2), \quad 0 < \sigma \le \rho/2,$$

where $\beta \in (0,1)$ is chosen to ensure that $2^{-\beta} = \theta$. Going back to (2) to estimate the integral on the right, we obtain finally

$$\int_{B_\sigma(y)} |Dv|^2 \le C(\sigma/\rho)^\beta\Big(\int_{B_\rho(y)} v^2 + L_2^2\rho^2\Big),$$

provided only that $B_\rho(y) \subset B_R(0)$.

Using Lemma 2 (or more correctly the remark following Lemma 2 with $p = 2$), we then conclude the required Hölder estimate for v on $B_{R/2}(0)$.

We prove the second part of the theorem from the first by using a reflection argument as follows. We let $\overline{v}$ be defined on all of B_R by odd reflection; that is, $\overline{v}(x^1,x^2) = -v(x^1,-x^2)$ for $x^2 < 0$ and $x \in B_R$ and $\overline{v} = v$ on B_R^+. We readily check that $\overline{v} \in H^1(B_R)$ and satisfies an equation of the form $(*)$, $(**)$ with B_R in place of Ω, with coefficients $\overline{a}_{ij}$ given by odd reflection if exactly one of i,j is 2 and by even reflection otherwise, and, in place of f_j, the functions $\overline{f_j}$ on B_R given by odd reflection for $j = 1$ and even reflection for $j = 2$. (See Exercise 15.3 below.) Thus the second part of the theorem follows from the first part.

We can now prove Theorem 2.

Proof of Theorem 2: First note that if $B_R(x_0) \subset \Omega$, then as shown in the discussion prior to Theorem 3, each of D_1u and D_2u satisfy an equation of the type considered in Theorem 3, hence we deduce

$$(1) \quad |Du(x_1)-Du(x_2)| \le C(1+\sup_{B_R(x_0)}|Du|)|x_1-x_2|^\beta, \quad \forall x_1,x_2 \in B_{R/2}(x_0).$$

Now the domain is C^2, hence for any $x_0 \in \partial\Omega$ we can find a C^2 transformation Ψ which flattens the boundary near x_0; thus there is $R > 0$ such that Ψ is a one-to-one C^2 map $B_R(x_0)$ onto an open set $U \subset \mathbb{R}^2$ such that $\Psi(x_0) = 0$, $\Psi(\Omega \cap B_{R(x_0)}) = U^+ \equiv \{x \in U : x^2 > 0\} \supset B_R^+$, and $\Psi(B_R(x_0) \cap \partial\Omega) = T \equiv \{x \in U : x^2 = 0\}$, and such that Ψ^{-1} is of class C^2. Notice that in terms of the coordinates $y = \Psi(x)$ the function $\widetilde{u}(y) \equiv u(x)$ satisfies an equation of the form

$$(2) \quad \textstyle\sum_{i,j=1}^2 \widetilde{a}_{ij}D_{y^i}\widetilde{u}D_{y^j}\widetilde{u} = -\sum_{i,j,k=1}^2 a_{ij}(x)D_{x^i}D_{x^j}\Psi^k(x)D_{y^k}\widetilde{u} + \widetilde{f} \text{ on } B_R^+,$$

where $\widetilde{a}_{ij}(y) = \sum_{k,\ell=1}^2 (D_i\Psi^k(x))(D_j\Psi^\ell(x))a_{k\ell}(x)$, and where $\widetilde{f}(y) = f(x)$. Notice also that $D_{y^1}\widetilde{u} = 0$ on $T = U \cap \{x : x^2 = 0\}$. But, applying the discussion preceding Theorem 2 to the solution $\widetilde{u}$ of the equation (2), we then deduce that $v = D_1\widetilde{u}$ satisfies an equation of the same form as that considered in Theorem 3, with $\widetilde{L}_1, \widetilde{L}_2$ in place of L_1, L_2, where $\widetilde{L}_j$ are determined by L_1, L_2, Ψ, and L, where L is any upper bound for $\sup_\Omega |Du|$. Since $D_1\widetilde{u}$ vanishes on T, the second part of Theorem 3 applies to give

$$|D_1\widetilde{u}(y_1) - D_1\widetilde{u}(y_2)| \le C|y_1-y_2|^\beta(1+\sup|D\widetilde{u}|), \quad y_1,y_2 \in B_{R/2}^+.$$

Notice also that the proof of Theorem 3 establishes that

$$(3) \qquad \int_{B_\sigma(y)\cap B_R^+} \textstyle\sum_{j=1}^2 |D_j D_1 \widetilde{u}|^2 \le C\sigma^{2\beta},$$

whenever $0 < \sigma < R/2$ and $y \in B_{R/2}^+$, where $C > 0$ and $\beta \in (0,1)$ depend only on L_1, L_2, Ψ and L, where L is any upper bound for $\sup_\Omega |Du|$. But by equation (2) we have

$$D_1 D_1 \widetilde{u} = \widetilde{a}_{11}^{-1}\big(\textstyle\sum_{(i,j)\neq(1,1)} \widetilde{a}_{ij} D_i D_j \widetilde{u} - \sum_{i,j=1}^2 a_{ij}(x) D_{x^i} D_{x^j} \Psi^k(x) D_k \widetilde{u} + \widetilde{f}\big),$$

and hence (3) implies

$$\int_{B_\sigma(y)\cap B_R^+} |D^2\widetilde{u}|^2 \le C'\sigma^{2\beta},$$

whenever $y \in B_{R/2}^+$ and $0 < \sigma < R/2$, where C' depends on the same quantities as C in (2).

Then by the Morrey lemma (Lemma 2 above) we have

$$|D\widetilde{u}(y_1) - D\widetilde{u}(y_2)| \le C|y_1 - y_2|^\beta, \quad y_1, y_2 \in B_{R/2}^+,$$

where C depends only on L_1, L_2, L, and Ψ. Using the inverse transformation Ψ^{-1}, we have thus shown that

$$(3) \qquad |u(x_1) - u(x_2)| \le C|x_1 - x_2|^\beta, \quad x_1, x_2 \in \Omega \cap B_{\theta R}(x_0),$$

where $\theta \in (0,1)$ depends only on Ψ, and C depends only on L_1, L_2, L, and Ψ.

Now by virtue of an elementary covering argument, using the compactness of $\partial\Omega$, we deduce that we can select R, θ, β, C such that (3) is valid for any $x_0 \in \partial\Omega$. The proof is thus completed by combining (1) and (3).

We can now complete the proof of Theorem 1 (that the estimate $(**)$ implies the estimate $(*)$) in case $n = 2$. In fact, from Theorem 2 we see immediately that $(**) \Rightarrow |u|_{1,\beta,\Omega} \le M_1$, where M_1 and β depend only on the constant K in $(**)$ and the functions a_{ij} and f. But then the cofficients $a_{ij}(x, u(x), Du(x))$ in the equation $Qu = \sigma f(x, u, Du)$ are $C^{0,\beta}(\overline{\Omega})$ functions of x, and also the $f(x, u(x), Du(x))$ is a $C^{0,\beta}(\overline{\Omega})$ function of x, and the $C^{0,\beta}$-norms of these functions depend only on M_1 and the functions f, a_{ij}. Also the minimum eigenvalue of the matrix $(a_{ij}(x, u(x), Du(x)))$ is bounded below (uniformly for x in $\overline{\Omega}$) by a fixed constant depending only on K and the coefficient functions a_{ij}. Thus the global Schauder estimates of Lecture 12 imply that the $C^{2,\beta}(\overline{\Omega})$-norm of u in bounded above by a constant depending only on K and

the functions a_{ij}, f. Thus in particular the $C^{1,\mu}(\overline{\Omega})$-norm of u is bounded above by a constant depending on the same quantities. Thus Theorem 1 is established in case $n = 2$.

As we mentioned above, we defer the proof of Theorem 1 in the case $n \ge 3$ until after we have established De Giorgi-Nash theory.

LECTURE 16 PROBLEMS

16.1 Give the proof of the claim made in the discussion following the statement of Theorem 2 that $v = D_2 u$ is a weak solution of $D_1 v_1 + 2D_2(a^{-1}bv_1) + D_2(a^{-1}cv_2) = D_2(a^{-1}f)$ assuming that $u \in C^2$.

16.2 By scaling, check the claim that the Poincaré inequality $\int_{B_1} |u - \overline{u}|^2 \le C \int_{B_1} |Du|^2$ for $u \in H^1(B_1)$ implies $\int_{B_\rho} |u - \overline{u}|^2 \le C\rho^{-2} \int_{B_\rho} |Du|^2$ for $u \in H^1(B_\rho)$.

16.3 Check the claim about $\overline{v}$ (the odd reflection of v) made in the last paragraph of the proof of Theorem 3.

16.4 Consider the problem

$$(\ddagger) \qquad \begin{cases} \sum_{i,j=1}^n a_{ij}(Du) D_i D_j u = 0 \text{ in } \Omega \\ u = \psi \text{ on } \partial\Omega, \end{cases}$$

where Ω is a bounded $C^{2,\mu}$ domain in $\mathbb{R}^n$, $a_{ij} \in C^2(\mathbb{R}^n)$, $\psi \in C^{2,\mu}(\overline{\Omega})$, and $\gamma|\xi|^2 \le \sum a_{ij}(p)\xi^i\xi^j \le \Lambda|\xi|^2$ for all $p \in \mathbb{R}^n$ and all $\xi \in \mathbb{R}^n$. (That is, the equation is uniformly elliptic.)

(i) Prove that any $C^{2,\mu}(\Omega)$ solution of $(\ddagger)$ satisfies $|u|_{0,\Omega} + \sup_{\partial\Omega} |Du| \le C$, where C depends only on Ω, γ, Λ, and M, with M any upper bound for $|\psi|_{2,\mu,\Omega}$.
Hint: Show that Lemma 1 of Lecture 15 can be applied.

(ii) Show that $(\ddagger)$ has a $C^{2,\mu}(\overline{\Omega})$ solution.
Hint: Show that each partial derivative $v = D_\ell u$ is a weak solution of an equation of the form $\sum D_i(a_{ij}(Du)D_j v) + \sum b_j D_j v = 0$, where b_j are bounded on $\overline{\Omega}$.

16.5 If the hypotheses are as in 16.4, except that the hypothesis of uniform ellipticity is replaced by the hypotheses that $(a_{ij}(p))$ is positive definite for each $p \in \mathbb{R}^n$ and that Ω is uniformly convex (so that there is $R > 0$ such that for each $y \in \partial\Omega$ there is a ball B of radius R with $\overline{\Omega} \subset \overline{B}$ and $y \in \partial B$).

(i) Prove that for each $y \in \partial\Omega$ there are affine (i.e. linear plus constant) functions $\varphi_y^{\pm}$ with

$$\varphi_y^{\pm}(y) = \psi(y),\ \varphi_y^- \le \psi \le \varphi^+ \text{ on } \overline{\Omega}, \text{ and } \sup_{y\in\partial\Omega,\, x\in\overline{\Omega}} |D\varphi_y^{\pm}| \le C,$$

where C depends only on Λ_0 = any upper bound for $|\psi|_{C^2(\overline{\Omega})}$ and on R.

(ii) Prove that problem $(\ddagger)$ above has a $C^{2,\mu}(\overline{\Omega})$ solution.

Lecture 17

De Giorgi, Nash, Moser Theory

We want to look at equations of the form

$$(*) \qquad \sum_{|\alpha|,|\beta|\le 1}(-1)^{|\beta|-1}D^{\beta}(a_{\alpha\beta}D^{\alpha}u) = 0,$$

in a ball $B_R(x_0) \subset \mathbb{R}^n$, where no continuity assumptions are made concerning the coefficients $a_{\alpha\beta}$. We do assume the ellipticity condition

$$\text{(i)} \qquad \sum_{|\alpha|=|\beta|=1} a_{\alpha\beta}\xi^{\alpha}\xi^{\beta} \ge |\xi|^2, \quad \forall\, \xi \in \mathbb{R}^n,\ x \in B_R(x_0);$$

(Notice that, using the alternative notation $a^{ij} = a_{e_i e_j}$, this can be written

$$\sum_{i,j=1}^{n} a^{ij}\xi_i\xi_j \ge |\xi|^2, \quad \xi \in \mathbb{R}^n,\ x \in B_R(x_0).)$$

We also assume that the "top-order" coefficients a^{ij} are bounded and measurable; specifically, we assume

$$\text{(ii)} \qquad \sum_{|\alpha|=|\beta|=1}^{n} |a_{\alpha\beta}| \le \Lambda \quad \text{on } B_R(x_0).$$

Concerning the lower-order coefficients (i.e. $a_{\alpha\beta}$ with $|\alpha|+|\beta| \le 1$), we need to at least to assume L^p bounds for suitable p. In fact we assume

$$a_{\alpha\beta} \in L^q(B_R(x_0)) \text{ if } |\alpha|+|\beta| = 1, \quad a_{00} \in L^{q/2}(B_R(x_0))$$

for some $q > n$, and

$$\text{(iii)} \qquad R^{1-n/q}\sum_{|\alpha|+|\beta|=1}\|a_{\alpha\beta}\|_{L^q(B_R(x_0))} + R^{2-2n/q}\|a_{00}\|^{1/2}_{L^{q/2}(B_R(x_0))} \le \Gamma.$$

(Of course we allow the possibility $q = \infty$ here, which corresponds to the important case when all the cofficients are bounded and measurable.)

For our present discussion we are interested in non-negative weak subsolutions of the equation $*$; that is we assume that $u \geq 0$, $u \in W^{1,2}(B_R(x_0))$, and

$$(**) \qquad \int_{B_R(x_0)} \sum_{|\alpha|,|\beta|\leq 1} a_{\alpha\beta} D^\beta u D^\alpha \zeta \leq 0 \ \forall \zeta \in W_0^{1,2}(B_R(x_0)) \text{ with } \zeta \geq 0.$$

Then we have

Theorem 1. *If $u \geq 0$, $u \in W^{1,2}(B_R(x_0))$, and if* (i), (ii), (iii) *and* $(**)$ *all hold, then u is locally bounded in $B_R(x_0)$, and for each $\theta \in (0,1)$, $p > 0$,*

$$\sup_{B_{\theta R}(x_0)} u \leq C\left(R^{-n} \int_{B_R(x_0)} u^p\right)^{1/p}$$

where $C = C(n, \theta, p, \Lambda, \Gamma)$.

Before we begin the proof, we need a couple of preliminaries concerning L^p and $W^{1,p}$ functions. Let Ω be any bounded domain in $\mathbb{R}^n$.

Lemma 1. *If measure of Ω is finite, and if $u \in L^1(\Omega)$, then*

$$\sup_\Omega |u| = \lim_{p\to\infty} \|u\|_{L^p(\Omega)}.$$

Remark: Of course, here and subsequently, $\sup_\Omega |u|$ means essential supremum of $|u|$ over Ω.

Proof: Evidently $\|u\|_{L^p} \leq |\Omega|^{1/p} \sup_\Omega |u|$, and hence $\limsup_{p\to\infty} \|u\|_{L^p} \leq \sup_\Omega |u|$.

On the other hand if $\lambda < \sup_\Omega |u|$, then $|u| \geq \lambda$ on a set Ω_λ of positive measure, and hence $\|u\|_{L^p} \geq \lambda |\Omega_\lambda|^{1/p}$, so that $\liminf_{p\to\infty} \|u\|_{L^p} \geq \lambda$.

Lemma 2 (Sobolev inequality.) *If $u \in W_0^{1,p}(\Omega)$ and if $1 \leq p < n$, then*

$$\left(\int_\Omega |u|^{np/(n-p)}\right)^{(n-p)/n} \leq C(n,p) \int_\Omega |Du|^p.$$

Proof: By the usual approximation procedure, it is enough to establish the inequality for $u \in C_c^\infty(\Omega)$. Also, it is enough to establish the case $p = 1$: we in fact prove

$$(1) \qquad \left(\int_\Omega |u|^{n/(n-1)}\right)^{(n-1)/n} \leq \int_\Omega |Du|,$$

because once this is established we get the general case by replacing the function u by $|u|^{(n-1)p/(n-p)}$ and using the Hölder inequality on the right. (Notice that $D|u| = (\operatorname{sgn} u) Du$—see Lemma 2 of Lecture 12.)

Since we can take $u \in C_c^\infty(\Omega)$, it is enough to prove (1) with $\Omega = \mathbb{R}^n$. This is easily checked as follows:

For $n = 2$

$$(2) \qquad \int_{\mathbb{R}} |u(\xi^1,\xi^2)|^2 \, d\xi^1 \leq \sup_{\xi^1 \in \mathbb{R}} |u(\xi^1,\xi^2)| \int_{\mathbb{R}} |u(\xi^1,\xi^2)| \, d\xi^1,$$

while for $n > 2$ by Hölder's inequality we have

$$(3) \qquad \int_{\mathbb{R}^{n-1}} |u(\xi',\xi^n)|^{n/(n-1)} \, d\xi' \leq \\ \left(\int_{\mathbb{R}^{n-1}} |u(\xi',\xi^n)|^{(n-1)/(n-2)} \, d\xi'\right)^{(n-2)/(n-1)} \left(\int_{\mathbb{R}^{n-1}} |u(\xi',\xi^n)| \, d\xi'\right)^{1/(n-1)}.$$

Now in case $n = 2$ we can use (2) together with two applications of the inequality

$$(4) \qquad \sup |h(t)| \leq \int_{-\infty}^{\infty} |h'(t)| \, dt \text{ for } h \in C_c^1(\mathbb{R})$$

to give

$$\int_{\mathbb{R}} |u(\xi^1,\xi^2)|^2 \, d\xi^1 \leq \left(\int_{-\infty}^{\infty} |D_1 u(\xi^1,\xi^2)| \, d\xi^1\right) \left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |D_2 u(\xi^1,t)| \, d\xi^1 dt\right),$$

and we get the required inequality by integrating over $\mathbb{R}$ with respect to ξ^2.

For $n \geq 3$ we note that if the required inequality holds with $n-1$ in place of n then (3) implies

$$\int_{\mathbb{R}^{n-1}} |u(\xi',\xi^n)|^{n/(n-1)} \, d\xi' \leq \int_{\mathbb{R}^{n-1}} |D_{\xi'} u(\xi',\xi^n)| \, d\xi' \left(\int_{\mathbb{R}^{n-1}} |u(\xi',\xi^n)| \, d\xi'\right)^{1/(n-1)},$$

and using (4) again in the second integral on the right, we have

$$\int_{\mathbb{R}^{n-1}} |u(\xi',\xi^n)|^{n/(n-1)} \, d\xi \leq \left(\int_{\mathbb{R}^{n-1}} |D_{\xi'} u(\xi',\xi^n)| \, d\xi'\right) \left(\int_{\mathbb{R}^n} |D_n u(\xi)| \, d\xi\right)^{1/(n-1)},$$

and hence we get (1) by integrating with respect to ξ^n. Thus (1) holds for all $n \geq 2$ by mathematical induction.

Remark: We are going to use the Sobolev inequality with $p = 2$ in case $n \geq 3$. Notice that $p = 2$ is not a legitimate choice for $n = 2$; in this case we shall instead use the following corollary.

Corollary 1. *If $n = 2$ and $u \in W_0^{1,2}(\Omega)$, then*

$$\left(\int_\Omega |u|^{2\kappa}\right)^{1/\kappa} \le C|\Omega|^{1/\kappa}\int_\Omega |Du|^2 \quad \forall\, \kappa \ge 2,\ C = C(\kappa).$$

Proof of Theorem 1: We take parameters $k \ge 1,\ \varepsilon > 0,\ p \ge 1,\ \alpha, \beta > 0$ with $\alpha > 1 + \beta$, and we substitute $\zeta = u(u_k + \varepsilon)^{2p-2}\varphi^{2\alpha p - 2\beta}$ in place of ζ in $(**)$, where $u_k = \min\{u, k\}$ and φ is a cut-off function of the form

$$\varphi \equiv 1 \text{ on } B_{\theta R},\ \varphi \equiv 0 \text{ outside } B_R$$
$$0 \le \varphi \le 1, \quad |D\varphi| \le \frac{2}{(1-\theta)R} \qquad \text{everywhere.}$$

Notice that (by Lemma 2 of Lecture 12) we have $u_k \in W^{1,\infty}(B_R(x_0))$ and $D_j u_k = D_j u$ a.e. on the set where $u < k$ and $D_j u_k = 0$ a.e. on the set where $u \ge k$. In particular one easily checks that the above choice of ζ is in $W_0^{1,2}(B_R(x_0))$, and hence is a legitimate choice in the inequality $(**)$. Furthermore, direct calculation shows that

$$\begin{aligned}(1)\qquad D_j\big(u(u_k+\varepsilon)^{2p-2}\varphi^{2\alpha p-2\beta}\big) &= (D_j u)(u_k+\varepsilon)^{2p-2}\varphi^{2\alpha p-2\beta} + \\ &\big((2p-2)u(u_k+\varepsilon)^{2p-3}D_j u_k\big)\varphi^{2\alpha p-2\beta} \\ &+ (2\alpha p - 2\beta)u(u_k+\varepsilon)^{2p-2}\varphi^{2\alpha p-2\beta-1}D_j\varphi,\end{aligned}$$

and hence since $p \ge 1$ and since (i) holds we get

$$\begin{aligned}\textstyle\sum_{i,j=1}^n a^{ij} D_i u D_j\big(u(u_k+\varepsilon)^{2p-2}\varphi^{2\alpha p-2\beta}\big) &\ge \\ |Du|^2(u_k+\varepsilon)^{2p-2}\varphi^{2\alpha p-2\beta}& \\ + (2\alpha p - 2\beta)\textstyle\sum_{i,j=1}^n u(u_k+\varepsilon)^{2p-2}\varphi^{2\alpha p-2\beta-1}a_{ij}D_i u D_j\varphi.&\end{aligned}$$

Hence the inequality $(**)$ gives

$$\begin{aligned}\int_{B_R}|Du|^2(u_k+\varepsilon)^{2p-2}\varphi^{2\alpha p-2\beta} &\le \frac{Cp\Lambda}{(1-\theta)R}\int_{B_R}|Du|u(u_k+\varepsilon)^{2p-2}\varphi^{2\alpha p-2\beta-1} \\ &+ \int_{B_R}\big(\textstyle\sum_{|\beta|=1}a_{0\beta}D^\beta u + a_{00}u\big)\,u(u_k+\varepsilon)^{2p-2}\varphi^{2\alpha p-2\beta} + \\ &\textstyle\sum_{|\alpha|=1}a_{\alpha 0}u D^\alpha\big(u(u_k+\varepsilon)^{2p-2}\varphi^{2\alpha p-2\beta}\big)\end{aligned}$$

with $C = C(\alpha, \beta, n)$. Notice that since $Du_k = 0$ for $u > k$ and $\varphi \le 1$ we can then estimate

$$\begin{aligned}\int_{B_R}|Du|^2(u_k+\varepsilon)^{2p-2}\varphi^{2\alpha p-2\beta} &\le Cp\int_{B_R}\Big(\big(\Lambda R^{-1} + \textstyle\sum_{|\alpha|+|\beta|=1}|a_{\alpha\beta}|\big)|Du| \\ &+ \big(|a_{00}| + R^{-1}\textstyle\sum_{|\alpha|=1}|a_{\alpha 0}|\big)u\Big)\big(u(u_k+\varepsilon)^{2p-2}\varphi^{2\alpha p-2\beta-1}\big),\end{aligned}$$

where $C = C(\alpha, \beta, n, \theta)$, and by using the Cauchy-Schwartz inequality we then conclude that

$$\begin{aligned}\int_{B_R}|Du|^2(u_k+\varepsilon)^{2p-2}\varphi^{2\alpha p-2\beta} &\le Cp^2\int_{B_R}\Big(\Lambda^2R^{-2} + \textstyle\sum_{|\alpha|+|\beta|=1}a_{\alpha\beta}^2 + \\ &R^{-1}\textstyle\sum_{|\alpha|=1}|a_{\alpha 0}| + |a_{00}|\Big)u^2(u_k+\varepsilon)^{2p-2}\varphi^{2\alpha p-2\beta-2},\end{aligned}$$

where again C depends only on n, α, β. Notice that by direct computation this gives

$$\begin{aligned}\int_{B_R}|D\big(u\,(u_k+\varepsilon)^{p-1}\varphi^{\alpha p-\beta}\big)|^2 &\le Cp^2\int_{B_R}\Big((1+\Lambda^2)R^{-2} + \\ &\textstyle\sum_{|\alpha|+|\beta|=1}a_{\alpha\beta}^2 + R^{-1}\sum_{|\alpha|=1}|a_{\alpha 0}| + |a_{00}|\Big)u^2(u_k+\varepsilon)^{2p-2}\varphi^{2\alpha p-2\beta-2},\end{aligned}$$

Now it is convenient at this point to suppose that $R = 1$; this involves no loss of generality, because we can change scale by the transformation $x \mapsto R^{-1}x$. Notice that the assumptions are stated in "scale-invariant" form—that is the assumptions remain unchanged except that $R = 1$ everywhere.

Then the above inequality gives

$$(2)\quad \int_{B_1}|D\big(u\,(u_k+\varepsilon)^{p-1}\varphi^{\alpha p-\beta}\big)|^2 \le Cp^2\int_{B_1}F\,u^2(u_k+\varepsilon)^{2p-2}\varphi^{2\alpha p-2\beta-2},$$

where $F = 1 + \Lambda^2 + \sum_{|\alpha|+|\beta|=1}a_{\alpha\beta}^2 + |a_{00}|$. Now notice that the assumption (iii) implies that

$$\|F\|_{L^{q/2}(B_1)} \le C$$

with $C = C(n, \Lambda, \Gamma)$, and hence using the Hölder inequality and the Sobolev inequality (Lemma 2 above) in (2) we conclude that

$$\begin{aligned}\Big(\int_{B_1}\big(u\,(u_k+\varepsilon)^{p-1}\varphi^{\alpha p}\big)^{2\kappa}\varphi^{-2\beta\kappa}\Big)^{1/\kappa} &\le \\ Cp^2\Big(\int_{B_1}\big(u(u_k+\varepsilon)^{p-1}\varphi^{\alpha p}\big)^{2\mu}\varphi^{-2(\beta+1)\mu}\Big)^{1/\mu},&\end{aligned}$$

where $C = C(n, \alpha, \beta, \Lambda, \Gamma, q)$, and where $\mu = q/(q-2)$ and $\kappa = n/(n-2)$ when $n \ge 3$ and $\kappa > q/(q-2)$ when $n = 2$. (Notice that in case $n = 2$ we have used Corollary 1 above.)

Now we let $\varepsilon \downarrow 0$ and $k \uparrow \infty$, so that the above inequality (by the monotone convergence theorem) implies

$$(3)\qquad \Big(\int_{B_1}(u\varphi^\alpha)^{2p\kappa}\varphi^{-2\beta\kappa}\Big)^{1/\kappa} \le Cp^2\Big(\int_{B_1}(u\varphi^\alpha)^{2p\mu}\varphi^{-2(\beta+1)\mu}\Big)^{1/\mu}.$$

Now we choose β: we in fact select β such that $(\beta+1)\mu = \beta\kappa$; that is, we take $\beta = \mu/(\kappa-\mu)$, and at the same time we choose $\alpha = \beta + 1$. Then (3) gives

$$\text{(4)} \qquad \Big(\int_{B_1} (u\varphi^\alpha)^{2p\kappa}\, dv\Big)^{1/\kappa} \le C\, p^2 \Big(\int_{B_1} (u\varphi^\alpha)^{2p\mu}\, dv\Big)^{1/\mu},$$

where $dv = \varphi^{-2\beta\kappa} dx$.

Notice that this can be written in the form

$$\text{(5)} \qquad \Psi(\gamma p) \le (C\,p^2)^{\mu/p}\Psi(p), \quad p \ge \mu, \ \gamma = \kappa/\mu > 1,$$

where

$$\Psi(p) = \int_{B_1} (u\varphi^\alpha)^{2p}\, dv)^{1/p},$$

where $C = C(n, q, \Lambda, \Gamma, \theta)$; in particular C does not depend on p.

Now iterating (5) with $p_0 = \mu, \gamma p_0, \gamma^2 p_0, \ldots, \gamma^{k-1} p_0$, we conclude that

$$\Psi(p_0\gamma^k) \le C^{1+\gamma^{-1}+\gamma^{-2}+\cdots+\gamma^{-k}}\, \gamma^{p_0 \sum_{j=1}^k j\gamma^{-j}}\, \Psi(p_0), \quad k \ge 1.$$

Then, taking account of the facts that $\varphi \le 1$, and hence that $dv \ge dx$, and that $\varphi \equiv 1$ on B_θ, we conclude from Lemma 1 that

$$\sup_{B_\theta} u^2 \le C\Big(\int_{B_1} (u\varphi^\alpha)^{2p_0}\, dv\Big)^{1/p_0},$$

and since $\alpha \ge \beta\kappa$, this gives

$$\text{(6)} \qquad \sup_{B_\theta} u^2 \le C\Big(\int_{B_1} u^{2p_0}\, dx\Big)^{1/p_0},$$

which is the required result with $p = p_0 \ge \mu$.

To get the required result for arbitrary $p > 0$, we just have to note that by (4) (with $p = 1$) and the Hölder inequality (with respect to the measure $dv = \varphi^{-2\kappa\beta}\, dx$) we have, for any $\varepsilon \in (0,1)$,

$$\Big(\int_{B_1} (u\varphi^\alpha)^{2\kappa}\, dv\Big)^{1/\kappa} \le C$$
$$\Big(\int_{B_1} (u\varphi^\alpha)^{2\kappa}\, dv\Big)^{(1-\varepsilon)/\kappa}\Big(\int_{B_1} (u\varphi^\alpha)^{2\varepsilon\mu\kappa/(\kappa-\mu(1-\varepsilon))}\, dv\Big)^{(\kappa-(1-\varepsilon)\mu)/(\mu\kappa)}.$$

Then since $u \in L^{2\kappa}_{\text{loc}}(\Omega)$ (by Problem 17.2 below), we deduce that

$$\Big(\int_{B_1} (u\varphi^\alpha)^{2\kappa}\, dv\Big)^{1/\kappa} \le C\Big(\int_{B_1} (u\varphi^\alpha)^{2\varepsilon\mu\kappa/(\kappa-\mu(1-\varepsilon))}\, dv\Big)^{(\kappa-(1-\varepsilon)\mu)/(\varepsilon\mu\kappa)}$$

We are thus able to estimate $(\int_{B_\theta} u^{2\kappa}\, dx)^{1/\kappa}$ in terms of $(\int_{B_1} u^{2p}\, dx)^{1/p}$ for arbitrarily small $p > 0$, provided we choose $\alpha > 0$ large enough to ensure $\alpha p \ge 1 + \beta$, with $\beta = \kappa\mu/(\kappa-\mu)$. This completes the proof.

We now turn our attention to non-negative weak supersolutions. That is, we consider $u \in W^{1,2}(B_R(x_0))$ with $u \ge 0$ a.e. and

$$(**) \qquad \int_{B_R(x_0)} \sum_{|a|,|\beta|\le 1} a_{\alpha\beta} D^\beta u\, D^\alpha \zeta \ge 0$$

for each $\zeta \in W^{1,2}_0(B_R(x_0))$ with $\zeta \ge 0$ on Ω. The main result here is the following:

Theorem 2. *Suppose $u \ge 0$ is a weak supersolution of $Lu = 0$ in $B_R(x_0)$ (i.e. $u \in W^{1,2}(B_R(x_0))$ satisfies $(**)$), let $\theta \in (0,1)$, and suppose conditions* (i), (ii), (iii) *hold. Then, for any $0 < p < n/(n-2)$,*

$$\inf_{B_{\theta R}(x_0)} u \ge C\big(R^{-n}\int_{B_{\theta R}(x_0)} u^p\, dx\big)^{1/p},$$

where $C = C(n, \Lambda, \Gamma, \theta)$.

The proof is achieved by combining the following two lemmas:

Lemma 3. *Under the above hypotheses, for each $p > 0$, $\theta \in (0,1)$*

$$\sup_{B_{\theta^2 R}(x_0)} u^{-p} \le CR^{-n}\int_{B_{\theta R}(x_0)} u^{-p}\, dx,$$

where $C = C(n, \theta, \Lambda, \Gamma, p)$.

Lemma 4. *Under the same hypotheses, for any $0 < p < n/(n-2)$,*

$$\big(\int_{B_{\theta R}(x_0)} u^p\big)\big(\int_{B_{\theta R}(x_0)} u^{-p}\big) \le CR^{2n},$$

where $C = C(n, \theta, \Lambda, \Gamma, p)$.

Before we begin the proof of these lemmas, we make some general remarks about the inequality $*$. We let $\gamma < 1$ be a constant, and replace ζ in $*$ by $(u+\varepsilon)^{\gamma-1}\zeta$, where $\varepsilon > 0$ and $\zeta \in C^\infty_c(\Omega)$. This gives

$$\int_\Omega \sum_{|\alpha|=|\beta|=1}\big((1-\gamma)(u+\varepsilon)^{\gamma-2} a_{\alpha\beta} D^\beta u\, D^\alpha u - (u+\varepsilon)^{\gamma-1} a_{\alpha\beta} D^\beta u\, D^\alpha \zeta\big) \le$$
$$\int_\Omega \big(\sum_{|\alpha|=1}\big((u+\varepsilon)^{\gamma-1}\zeta\tilde a_{0\alpha} D^\alpha u + \tilde a_{\alpha 0}(u+\varepsilon)^\gamma D^\alpha\zeta\big) + \tilde a_{00}\zeta(u+\varepsilon)^\gamma\big),$$

where

$$\widetilde{a}_{0\alpha} = a_{0\alpha} - (1-\gamma)a_{\alpha 0}\frac{u}{u+\varepsilon}$$
$$\widetilde{a}_{\alpha 0} = \frac{u}{u+\varepsilon}a_{\alpha 0}, \quad |\alpha| = 1, \quad \widetilde{a}_{00} = \frac{u}{u+\varepsilon}a_{00}$$

Thus in case $\gamma = 0$, we get

(a) $$\int_\Omega \sum_{|\alpha|=|\beta|=1}(\zeta a_{\alpha\beta} D^\alpha w D^\beta w - a_{\alpha\beta} D^\beta w D^\alpha \zeta) \le \int_\Omega (\sum_{|\alpha|=1}(\zeta \widetilde{a}_{0\alpha} D^\alpha w + \widetilde{a}_{\alpha 0} D^\alpha \zeta) + \widetilde{a}_{00}\zeta),$$

with $w = \log(u+\varepsilon)$, while in case $\gamma \neq 0$ (i.e. $0 < \gamma < 1$ or $\gamma < 0$), we get

(b) $$\int_\Omega \sum_{|\alpha|=|\beta|=1}((1-\gamma)\gamma^{-2}\zeta a_{\alpha\beta} D^\alpha w D^\beta w - \gamma^{-1} a_{\alpha\beta} D^\beta w D^\alpha \zeta) \le \int_\Omega (\sum_{|\alpha|=1}(\gamma^{-1}\zeta \widetilde{a}_{0\alpha} D^\alpha w + \widetilde{a}_{\alpha 0} D^\alpha \zeta) + \widetilde{a}_{00} w\zeta),$$

with $w = (u+\varepsilon)^\gamma$.

Remark: Notice that the coefficients $\widetilde{a}_{\alpha\beta}$ satisfy the same bounds (iii) as $a_{\alpha\beta}$, with $C\Lambda$ in place of Λ, where $C = C(\gamma)$.

Proof of Lemma 3: Notice that (b) above tells us that $w = (u+\varepsilon)^\gamma$, for $\gamma < 0$ is a non-negative subsolution of an equation of the same form as $*$, and, by the remark above, the conditions (i), (ii), (iii) all hold with $C\Lambda$ in place of Λ. Hence by the result of Theorem 1 above we have the required result after letting $\varepsilon \downarrow 0$.

Proof of Lemma 4: First note that by (a) above we get, after replacing ζ by ζ^2 and using the conditions (i), (ii), that

$$\int_\Omega \zeta^2 |Dw|^2 \le C\int_\Omega (\zeta|Dw||D\zeta| + \zeta^2|a_{00}| + \zeta(\sum_{|\alpha|+|\beta|=1}|a_{\alpha\beta}|)(|D\zeta| + \zeta|Dw|)),$$

and so by using the Cauchy-Schwartz inequality we obtain

$$\int_\Omega \zeta^2 |Dw|^2 \le C\int_\Omega (|D\zeta|^2 + \zeta^2(|a_{00}| + \sum_{|\alpha|+|\beta|=1} a_{\alpha\beta}^2)),$$

which in view of the bounds (iii) of above gives us

(1) $$\int_\Omega \zeta^2 |Dw|^2 \le C\int_\Omega (|D\zeta|^2 + F\,\zeta^2),$$

where

(2) $$R^{1-q/n}\|F\|_{L^q(B_{\theta^{-1}R}(x_0))} \le C\Lambda,$$

with $C = C(n)$.

Then taking ζ, φ to be a cut-off functions, like φ used in the proof of Theorem 1 above, but with $\theta^{-2}R, \theta^{-1}$ respectively in place of R) we obtain in particular that

(3) $$\int_{B_{\theta^{-1}R}} |Dw|^2 \le CR^{n-2},$$

where $C = C(n,\theta,\Lambda)$, and replacing ζ in (1) by $(w-\lambda)^{2q-2}\varphi^{2\alpha q - 2\beta}$, where $q \ge 2$ and α, β are to be chosen with $\alpha \ge 1 + \beta$,

$$\int_\Omega v^{2q-2}|Dv|^2\varphi^{2\alpha q-2\beta} \le C\int_\Omega (q^2 v^{2q-4}|Dv|^2\varphi^{2\alpha q-2\beta} + F\,v^{2q-2}\varphi^{2\alpha q-2\beta-2}),$$

where $C = C(n,\Lambda,\theta,\alpha,\beta)$ and $v = w - \lambda (\equiv \log(u+\varepsilon) - \lambda)$. By Young's inequality we have

$$Cq^2 v^{2q-4} \le \frac{1}{2}v^{2q-2} + (C')^q q^{2q},$$

and hence the above inequality implies

(4) $$\int_\Omega v^{2q-2}|Dv|^2\varphi^{2\alpha q-2\beta} \le C^q q^{2q}\int_\Omega |Dv|^2\varphi^{2\alpha q-2\beta} + C\int_\Omega F\,v^{2q-2}\varphi^{2\alpha q-2\beta-2}.$$

Using (2) and the fact that $v^{2q-2} \le 1 + v^{2q}$, this implies

(5) $$\int_\Omega v^{2q-2}|Dv|^2\varphi^{2\alpha q-2\beta} \le C^q q^{2q} R^{n-2} + C\int_\Omega F\,v^{2q}\varphi^{2\alpha q-2\beta-2}.$$

Hence

(6) $$R^{2-n}\int_\Omega |D(|v|^q\varphi^{\alpha q-\beta})|^2 \le C^q q^{2q} + C\int_\Omega F\,v^{2q}\varphi^{2\alpha q-2\beta-2},$$

so that (Cf. the proof of Theorem 1 above) by the Sobolev inequality and the choices $\alpha = \kappa\beta$, $\beta = \mu/(\kappa-\mu)$, $\mu = q/(q-2)$ we have

(7) $$\Big(R^{-n}\int_\Omega (|v|\varphi^\alpha)^{2q\kappa}d\eta\Big)^{1/\kappa} \le C^q q^{2q} + Cq^2 R^{-n}\Big(\int_\Omega (|v|\varphi^\alpha)^{2q\mu}d\eta\Big)^{1/\mu},$$

where $d\eta = \varphi^{-2\kappa\beta}\,dx$ and $\kappa = n/(n-2)$ for $n \ge 3$ and $\kappa > q/(q-2)$ in case $n = 2$.

Setting $q = \kappa^\nu$, and using the inequality $(a+b)^\varepsilon \le a^\varepsilon + b^\varepsilon$ (valid for any $a, b \ge 0$, $\varepsilon \in (0,1)$), we then obtain

$$\Psi(\nu+1) \le C\kappa^{2\nu} + C^{\nu/\kappa^\nu}\Psi(\nu), \quad \nu \ge \nu_0, \tag{8}$$

where

$$\Psi(\nu) = \big(R^{-n}\int_\Omega (|v|\varphi^\alpha)^{2\kappa^\nu}\,d\eta\big)^{1/\kappa^\nu}$$

Iterating the inequality in (8) we get (since $\sum_{j=1}^{\nu}\kappa^{2j} \le C\kappa^{2\nu}$ and $\sum_{j=1}^{\infty} j\kappa^{-j} < \infty$) that

$$\Psi(\nu) \le C\kappa^{2\nu} + C\Psi(\nu_0). \tag{9}$$

and where $\nu_0 \ge 1$ is such that $\kappa^{\nu_0} \ge 2$.

But now by (7) (with $q = 2$) in combination with Hölder's inequality (with respect to the measure $d\eta = \varphi^{-2\kappa\beta}\,dx$) we also have

$$\begin{aligned}\Psi(\nu_0) &\le \big(C_1 + C_2R^{-n}\int_\Omega v^2\varphi^{2\alpha-2\kappa\beta}\,dx\big)\\ &\le \big(C_1 + C_2R^{-n}\int_{B_{\theta^{-1}R}(x_0)} v^2\,dx\big).\end{aligned} \tag{10}$$

For suitable choice of λ we can use the Poincare inequality to deduce

$$R^{-n}\int_{B_{\theta^{-1}R}(x_0)} v^2 \le CR^{2-n}\int_{B_{\theta^{-1}R}(x_0)} |Dw|^2 \le C \tag{11}$$

by (3) above.

Hence (9), (10), (11) imply that

$$\Psi(\nu) \le C\kappa^{2\nu}, \quad \nu = 1, 2, \ldots,$$

and by Hölder's inequality we thus have

$$\big(R^{-n}\int_{B_R(x_0)} |v|^j\,dx\big)^{1/j} \le Cj, \quad j = 1, 2, \ldots.$$

Taking j^{th} powers and using the fact that $\sum_{j=1}^{\infty} x^j/j^j \ge \exp(c_1x) - c_2$ for suitable constants $c_1, c_2 > 0$, we then have (after summing over j) that

$$R^{-n}\int_{B_R(x_0)} e^{c_1|\log(u+\varepsilon)-\lambda|}\,dx \le C$$

for suitable $c_1 > 0$ depending only on n, Λ, θ. This of course implies

$$\int_{B_R(x_0)} e^{c_1(\log(u+\varepsilon)-\lambda)}\,dx,\ \int_{B_R(x_0)} e^{-c_1(\log(u+\varepsilon)-\lambda)}\,dx \le CR^n,$$

and hence

$$\big(\int_{B_R(x_0)} (u+\varepsilon)^{c_1}\,dx\big)\big(\int_{B_R(x_0)} (u+\varepsilon)^{-c_1}\,dx\big) \le C^2R^{2n}, \tag{12}$$

so Lemma 2 is established for sufficiently small $p > 0$.

To finish the proof, we have to show that we can go from small p to any $p < n/(n-2)$. For this we again use the identity (b) above, this time with $\gamma \in (0,1)$. Indeed using (b) together with the Cauchy-Schwartz inequality we obtain (Cf. (1) above)

$$\int_\Omega \zeta^2|Dw|^2 \le C\int_\Omega (w^2|D\zeta|^2 + F\,w^2\zeta^2),$$

and hence

$$\int_\Omega |D(w\zeta)|^2 \le C\int_\Omega (w^2|D\zeta|^2 + F\,w^2\zeta^2).$$

Choosing $\zeta = \varphi^{2\alpha-2\beta}$ as before and applying the Hölder and Sobolev inequalities as before (Cf. the argument going from (6) to (7) above), we conclude

$$\big(\int_\Omega (w\varphi^\alpha)^{2\kappa}\,d\eta\big)^{1/\kappa} \le C\big(\int_\Omega (w\varphi^\alpha)^{2\mu}\,d\eta\big)^{1/\mu},$$

where as before $d\eta = \varphi^{-2\kappa\beta}\,dx$, $\mu = q/(q-2)$, and $\kappa = n/(n-2)$ for $n \ge 3$ and $\kappa > q/(q-2)$ for $n = 2$.

By virtue of the Hölder inequality this gives, for each $\varepsilon > 0$,

$$\big(\int_\Omega (w\varphi^\alpha)^{2\kappa}\big)^{1/\kappa} \le C\big(\int_\Omega (w\varphi^\alpha)^{2\varepsilon\kappa\mu/(\kappa-(1-\varepsilon)\mu)}\big)^{(\kappa-(1-\varepsilon)\mu)/(\varepsilon\kappa\mu)},$$

where $C = C(n, \gamma, \Lambda, \Gamma, \theta)$.

The required inequality for all $p < n/(n-2)$ now evidently follows from this and (12), provided we select α so that $\alpha\varepsilon = \kappa\beta + 1$.

LECTURE 17 PROBLEMS

17.1 Give a (reasonably elementary) proof of Theorem 1 in case $n = 1$, when Ω is an interval $(a, b) \subset \mathbb{R}$; assume the coefficients $a_{\alpha\beta}$ are all bounded if you wish.

17.2 By using cut-off functions and the Sobolev inequality, prove the inclusion $W^{1,2}(\Omega) \subset L^p_{\text{loc}}(\Omega)$, where $p = 2n/(n-2)$ in case $n \geq 3$ and $p \geq 1$ is arbitrary in case $n = 2$.

17.3 Discuss the dependence of the constant C in Theorem 1 on the constant Λ; specifically, show that $C \leq C_0 \Lambda^Q$ for suitable constants C_0, Q depending only on n, θ, p.

Lecture 18

Applications of De Giorgi Nash Moser Theory

We continue our discussion of the operator

$$Lu = \textstyle\sum_{|\alpha|,|\beta|\leq 1} (-1)^{|\beta|} D^\beta (a_{\alpha\beta} D^\alpha u),$$

where the coefficients are subject to conditions (i), (ii), (iii) of the previous lecture. We begin by noting the following important consequence of Theorems 1 and 2 of the previous lecture.

Theorem 1 (Harnack's inequality.) *If $u \geq 0$, $u \in W^{1,2}(\Omega)$ and $Lu = 0$ weakly in Ω, then there is $C = C(n, \theta, \Lambda, \Gamma)$ such that*

$$\sup_{B_{\theta R}(x_0)} u \leq C \inf_{B_{\theta R}(x_0)} u$$

for any $\theta \in (0, 1)$ and any ball $B_R(x_0) \subset \Omega$.

Proof: Since u is a weak solution, both Theorems 1 and 2 of Lecture 17 are applicable. That is, using the theorems with $p = 1$, we have

$$\sup_{B_{\theta R}(x_0)} u \leq CR^{-n} \int_{B_{\theta' R}(x_0)} u\, dx \leq C \inf_{B_{\theta R}(x_0)} u,$$

where $\theta' = (1 + \theta)/2$, and where C depends only on $n, \theta, \Lambda, \Gamma$.

Another principal consequence of the previous two lectures is the following bound on weak solutions of the inhomogeneous equation

$$(\ddagger) \qquad Lu = \textstyle\sum_{|\beta|\leq 1} (-1)^{|\beta|} D^\beta f_\beta,$$

where f_β are given functions with $f_\beta \in L^q(\Omega)$ for $|\beta| = 1$ and $f_0 \in L^{q/2}(\Omega)$ for some $q > n$; without loss of generality we assume this q agrees with that in (iii) of Lecture 17.

Theorem 2. *If $u \in W^{1,2}(\Omega)$ is a weak solution of the equation (‡) above, where (i), (ii),(iii) of Lecture 17 hold, and where $f_\beta \in L^q(\Omega)$ for $|\beta| = 1$ and $f_0 \in L^{q/2}(\Omega)$, with $q > n$, as in (iii) of Lecture 17, then for each $p > 0$*

$$\sup_{B_{\theta R}(x_0)} |u| \le C_1\big(R^{-n}\int_{B_R(x_0)} |u|^p\,dx\big)^{1/p} + C_2\big(\textstyle\sum_{|\beta|=1} R^{1-n/q}\|f_\beta\|_{L^q(B_R(x_0))} + R^{2-2n/q}\|f_0\|_{L^{q/2}(B_R(x_0))}\big),$$

where $C_j = C_j(n, \theta, \Lambda, \Gamma, p)$, $j = 1, 2$.

We emphasize that there is no sign restriction on u here.

Proof: Replacing u by δu, where

$$\delta = \Big(\varepsilon + R^{1-n/q}\textstyle\sum_{|\beta|=1}\|f_\beta\|_{L^q(B_R)} + R^{2-2n/q}\|f_0\|_{L^{q/2}(B_R)}\Big)^{-1},$$

we see that (since we can let $\varepsilon \downarrow 0$) it is enough to prove

(1) $$\sup_{B_{\theta R}(x_0)} |u| \le C_1\Big(R^{-n}\int_{B_R} |u|^p\,dx\Big)^{1/p} + C_2,$$

where C_1, C_2 depend only on $n, \theta, \Lambda, \Gamma R^{1-n/q}$, under the assumption that

(2) $$\textstyle\sum_{|\beta|=1} R^{1-n/q}\|f_\beta\|_{L^q(B_R)} + R^{2-2n/q}\|f_0\|_{L^{q/2}(B_R)} \le 1.$$

Now we choose a $C^2(\mathbb{R})$ function φ satisfying the conditions

(3) $$\begin{cases} \varphi(t) \equiv |t| \text{ for } |t| \ge 1 \\ \varphi''(t) \ge 0,\ \varphi(t) \equiv \varphi(-t),\ \varphi(t) \ge 1/2\ \forall\, t \in \mathbb{R} \end{cases}$$

Now u satisfies

$$\int_\Omega \textstyle\sum_{|\alpha|,|\beta|\le 1} a_{\alpha\beta} D^\alpha u D^\beta \zeta = \int_\Omega \sum_{|\beta|\le 1} f_\beta D^\beta \zeta$$

for each $\zeta \in C_c^\infty(\Omega)$. Replacing ζ by $\zeta\varphi'(u)$, we get

$$\int_\Omega \big(\zeta\varphi''(u)\textstyle\sum_{|\beta|=1,\,|\alpha|\le 1} a_{\alpha\beta} D^\alpha u D^\beta u + \varphi'(u)\sum_{|\alpha|,|\beta|\le 1} a_{\alpha\beta} D^\alpha u D^\beta \zeta\big)\,dx \le \int_\Omega \Big(\sum_{|\beta|=1}\varphi''(u) f_\beta D^\beta u\zeta + \sum_{|\beta|\le 1}\varphi'(u) f_\beta D^\beta \zeta\Big),$$

and using the ellipticity condition (i) and the Cauchy-Schwartz inequality to give

$$\big|\textstyle\sum_{|\beta|=1} a_{0\beta} u D^\beta u\big| \le \frac{1}{2}\big(|Du|^2 + \sum_{|\beta|=1} u^2 a_{0\beta}^2\big),$$
$$\big|\textstyle\sum_{|\beta|=1} f_\beta D^\beta u\big| \le \frac{1}{2}\big(|Du|^2 + \sum_{|\beta|=1} f_\beta^2\big),$$

we deduce

$$\int_\Omega \textstyle\sum_{|\alpha|,|\beta|\le 1} \widetilde{a}_{\alpha\beta} D^\alpha w D^\beta \zeta \le 0$$

if $\zeta \ge 0$, where $w = \varphi(u)$ and

$$\begin{aligned} \widetilde{a}_{\alpha\beta} &= a_{\alpha\beta} \text{ if } |\beta| = 1, |\alpha| \le 1 \\ \widetilde{a}_{\alpha 0} &= \frac{u\varphi'(u)}{\varphi(u)} a_{\alpha 0} - \frac{\varphi'(u)}{\varphi(u)} f_\alpha - a_{0\alpha}\frac{\varphi'' u}{\varphi}, \quad |\alpha| = 1 \\ \widetilde{a}_{00} &= \frac{u\varphi'(u)}{\varphi(u)} a_{00} - \frac{\varphi'(u)}{\varphi(u)} f_0 - \tfrac{1}{2}\textstyle\sum_{|\beta|=1}(f_\beta^2 + u^2 a_{0\beta}^2)\frac{\varphi''}{\varphi} \end{aligned}$$

This has exactly the form of the inequality considered in Lecture 17, with (i), (ii), (iii) holding with $C\Lambda$, $C(\Gamma)$ in place of Λ, Γ respectively.

Thus the required inequality (1) now follows from Lecture 17, because $|u| \le w \le 1 + |u|$ everywhere in Ω.

The following theorem gives the main Hölder continuity result for weak solutions of the equation $Lu = \sum_{|\beta|\le 1}(-1)^{|\beta|} D^\beta f_\beta$, with f_β as in Theorem 2.

Theorem 3. *Suppose $u \in W^{1,2}(\Omega)$, is a weak solution of the equation $Lu = \sum_{|\beta|\le 1} D^\beta f_\beta$ in Ω, where the coefficients $a_{\alpha\beta}$ and the functions f_β satisfy the restrictions of Theorem 2. If $\overline{B}_R(x_0) \subset \Omega$ and if $\theta \in (0, 1)$, then (after redefinition of u on a set of measure zero)*

$$|u(x) - u(y)| \le C\Big(\sup_{B_R(x_0)} |u| + R^{1-n/q}\textstyle\sum_{|\beta|=1}\|f_\beta\|_{L^q(B_R(x_0))} + R^{2-2n/q}\|f_0\|_{L^{q/2}(B_R(x_0))}\Big)\Big(\frac{|x-y|}{R}\Big)^\mu$$

for all $x, y \in B_{\theta R}(x_0)$, where $\mu \in (0, 1)$ depends on $n, \theta, \Lambda, \Gamma R^{1-n/q}$, and where C depends also on θ.

This continues to be true for points $x, y \in \Omega \cap B_{\theta R}(x_0)$, if the hypothesis $\overline{B}_R(x_0) \subset \Omega$ is replaced by the hypotheses that $x_0 \in \partial\Omega$, $u|\partial\Omega \cap B_R(x_0)$ is Lipschitz with Lipschitz constant $\le \Gamma$, and if there is a coordinate transformation $\Phi : B_R(x_0) \to U$, with $\Phi^{-1} : U \to B_R(x_0)$ such that

$$\Phi(B_R(x_0) \cap \Omega) = U_+ \equiv \{(x^1, \ldots, x^n) \in U : x^n > 0\}$$
$$\Phi(B_R(x_0) \cap \partial\Omega) = T \equiv \{(x^1, \ldots, x^n) \in U : x^n = 0\},$$

and Lip Φ, Lip $\Phi^{-1} \le \Lambda$.

Proof: We already know by Theorem 2 that u is bounded on $B_R(x_0)$. Also we can assume $\sup_{B_R(x_0)} |u| \ne 0$, otherwise there is nothing to prove. By multiplying the equation by the factor

$$\Big(\sup_{B_R(x_0)} |u| + R^{1-n/q}\textstyle\sum_{|b|=1}\|f_\beta\|_{L^q(B_R(x_0))} + R^{2-2n/q}\|f_0\|_{L^{q/2}(B_R(x_0))}\Big)^{-1}$$

we reduce to proving

(1) $$\sup{}_{B_\rho(y)} u - \inf{}_{B_\rho(y)} u \le C\left(\frac{\rho}{R}\right)^{\beta},$$

where sup, inf mean essential sup and essential inf respectively and where $C > 0$ and $\beta \in (0,1)$ depend only on $n, \Gamma R^{1-n/q}, \Lambda, \theta$, subject to the assumptions

(2) $$\sup_{B_R(x_0)} |u| \le 1, \quad R^{1-n/q}\textstyle\sum_{|b|=1}\|f_\beta\|_{L^q(B_R(x_0))}) + R^{2-2n/q}\|f_0\|_{L^{q/2}(B_R(x_0))} \le 1.$$

((1) evidently implies the required continuity estimate for points $x, y \in B_{\theta R}(x_0)$ with $|x-y| \le (1-\theta)R$; for points x, y with $|x-y| \ge (1-\theta)R$ the result trivially holds, provided we take the constant C sufficiently large.)

To prove (1), we take fixed $\rho_0 \le (1-\theta)R$, $y \in B_{\theta R}(x_0)$, and for $\rho \le \rho_0$ we define

$$M_\rho = \sup{}_{B_\rho(y)} u, \quad m_\rho = \inf{}_{B_\rho(y)} u.$$

Note that M_ρ, m_ρ are finite by Theorem 2 .

Let $v = M_\rho - u + (\rho/R)^{1-n/q}$, and notice that, by direct computation,

$$\int_\Omega \textstyle\sum_{|\alpha|=1,|\beta|\le 1} a_{\alpha\beta} D^\alpha v D^\beta \zeta = -\displaystyle\int_\Omega \Big(u\textstyle\sum_{|\beta|\le 1} a_{0\beta} D^\beta\zeta + \sum_{|\beta|\le 1} f_\beta D^\beta \zeta\Big),$$

so that we can write

$$\int_\Omega \textstyle\sum_{|\alpha|,|\beta|\le 1} \widetilde{a}_{\alpha\beta} D^\alpha v D^\beta \zeta = 0,$$

where

$$\widetilde{a}_{\alpha\beta} = a_{\alpha\beta} \text{ in case } |\alpha| = 1, \quad \widetilde{a}_{0\beta} = v^{-1}(ua_{0\beta} - f_\beta), \quad |\beta| = 1,$$
$$\widetilde{a}_{00} = v^{-1}(ua_{00} - f_0),$$

so that (since $v \ge (\rho/R)^{1-n/q}$ and since we are assuming (1)), we have

$$\rho^{1-n/q}\|\widetilde{a}_{0\beta}\|_{L^q(\Omega)} + \rho^{2-2n/q}\|\widetilde{a}_{00}\| \le C(1+\Gamma),$$

where $C = C(n)$, and hence we can apply the result of Lecture 17 to v on $B_\rho(y)$ in order to give

$$\inf_{B_{\rho/2}(y)} v \ge C \inf_{B_{\rho/2}(y)} v\,dx,$$

so that

(3) $$\int_{B_{\rho/2}(y)} (M_\rho - u)\,dx \le C\left(M_\rho - M_{\rho/2} + (\rho/R)^{1-n/q}\right).$$

Similarly, starting with $v = u - m_\rho + (\rho/R)^{1-n/q}$, we get

(4) $$\int_{B_{\rho/2}(y)} (u - m_\rho)\,dx \le C\left(m_{\rho/2} - m_\rho + (\rho/R)^{1-n/q}\right).$$

By adding (3) and (4) we then conclude

$$M_\rho - m_\rho \le C(M_\rho - m_\rho) - C(M_{\rho/2} - m_{\rho/2}) + C(\rho/R)^{1-n/q},$$

where $C = C(n, \Lambda, \Gamma R^{1-n/q}) > 1$, or in other words

$$M_{\rho/2} - m_{\rho/2} \le \mu(M_\rho - m_\rho) + C(\rho/R)^{1-n/q},$$

where $\mu = 1 - C^{-1} \in (0,1)$ does not depend on R; indeed μ evidently depends only on the parameters $n, \Lambda, \Gamma R^{1-n/q}$).

By iteration, this gives

$$M_{\rho/2^\nu} - m_{\rho/2^\nu} \le \mu^\nu(M_\rho - m_\rho) + C(\rho/R)^{1-n/q}\left(\textstyle\sum_{j=1}^{\nu} 2^{-j}\right),$$

so that for any pair of points $x, y \in B_{\theta R}(x_0)$ with $|x-y| = \rho \le \rho_0 \equiv (1-\theta)R$ we have, since $\theta = 2^{-\beta}$ for some $\beta > 0$, and, since every $r \in (0, \rho_0)$ satisfies $2^{-\nu}\rho_0 \le r \le 2^{-(\nu-1)}\rho_0$ for some $\nu \ge 1$,

$$|u(x) - u(y)| \le 2^\alpha\left(\frac{\rho}{R}\right)^\alpha + C\left(\frac{r}{R}\right)r^{1-n/q}.$$

This evidently implies the required inequality (1).

Finally we want to discuss the extension of the above interior boundedness and continuity results to the boundary, assuming that Ω is a Lipschitz domain and that u satisfies Lipschitz Dirichlet data on $\partial\Omega$.

Specifically, assume $x_0 \in \partial\Omega$ and that there is $R > 0$ and Lipschitz maps $\Phi : B_R(x_0) \to U$, $\Phi^{-1} : U \to B_\rho(x_0)$ with

$$\Phi(B_R(x_0) \cap \Omega) = U_+ \equiv \{(x^1, \dots, x^n) \in U : x^n > 0\}$$
$$\Phi(B_R(x_0) \cap \partial\Omega) = T \equiv \{(x^1, \dots, x^n) \in U : x^n = 0\},$$

and Lip Φ, Lip $\Phi^{-1} \le \Lambda$.

By making the change of variable $y = \Phi(x)$ in the weak form of the equation, we obtain (in place of the equation for u on $B_{R(x_0)} \cap \Omega$) an equation for $\widetilde{u} \equiv u \circ \Phi^{-1}$ of the same form on U_+, with $C\Lambda$, $C\ G$ in place of Λ, Γ respectively.

Now we assume that $u = \varphi$ on $B_R(x_0) \cap \partial\Omega$ (in the sense that $(u-\varphi)\zeta \in W_0^{1,2}(B_R(x_0) \cap \Omega)$ whenever $\zeta \in C_c^\infty(B_R(x_0))$), so that $\widetilde{u} = \widetilde{\psi} \equiv \psi \circ \Phi^{-1}$ on T in the sense that $(\widetilde{u} - \widetilde{\psi})\zeta \in W_0^{1,2}(U_+)$ whenever $\zeta \in C_c^\infty(U)$.

Then on $\widetilde{U} \equiv U_+ \cup T \cup \{(x^1, \ldots, -x^n) : (x^1, \ldots, x^n) \in U_+\}$, we define $v_O(x^1, \ldots, -x^n) = -v(x^1, \ldots, x^n)$ (odd reflection) and $v_E(x^1, \ldots, -x^n) = v(x^1, \ldots, x^n)$ (even reflection). Then one readily checks that, on $\widetilde{U}$, the odd reflection $(\widetilde{u} - \widetilde{\psi})_O$ satisfies an equation of the same form as that for $\widetilde{u}$ on U_+ with coefficients given by

$$\widetilde{a}_{\alpha\beta} = (a_{\alpha\beta})_O \text{ in case exactly one of } \alpha,\ \beta = e_n$$
$$\widetilde{a}_{\alpha\beta} = (a_{\alpha\beta})_E \text{ otherwise.}$$

In checking that the weak form of the equation holds for test functions $\zeta \in C_c^\infty(\widetilde{U})$ we need to note that any such test function can be written in the form $\zeta = \zeta_1 + \zeta_2$, where ζ_1 is the even part of ζ with respect to the variable x^n and ζ_2 is the odd part. Notice that by definition the odd part vanishes on T.

Now we want to prove Theorem 1 of Lecture 16.

Thus we consider solutions of the problem

$$\text{(P)} \qquad \begin{cases} \sum_{i,j=1}^n a_{ij}(x,u,Du) D_i D_j u = f \text{ on } \Omega \\ u = \psi \text{ on } \partial\Omega \end{cases}$$

where f is a given bounded function on Ω, and where the coefficient functions a_{ij} are $C^1(\overline{\Omega}\times\mathbb{R}\times\mathbb{R}^n)$. In fact we assume the following bounds on $\Omega \cap B_R(x_0) \times [-K,K] \times \{p \in \mathbb{R}^n : |p| \le K\}$, where K is a given positive constant and $B_R(x_0)$ is a given ball in $\mathbb{R}^n$:

$$(*) \qquad \begin{cases} \sum_{i,j=1}^n (|a_{ij}(x,z,p)| + |Da_{ij}(x,z,p)|) \le \Gamma \\ R^2|f(x)| \le \Gamma, \end{cases}$$

where $\Gamma = \Gamma(K)$ is a constant and $|Da(x,z,p)|$ is defined by

$$|Da(x,z,p)| = \sum_{j=1}^n (R|D_{x^j}a(x,z,p)| + |D_{p^j}a(x,z,p)|) + R|p||D_z a(x,z,p)|$$

for any function $a = a(x,z,p)$. Of course we also as usual assume the ellipticity condition that, for each $K > 0$,

$$(**) \qquad \sum_{i,j=1}^n a_{ij}(x,z,p)\xi_i\xi_j \ge \lambda|\xi|^2,$$

for all $(x,z,p) \in \Omega \times [-K,K] \times \{p \in \mathbb{R}^n : |p| \le K\}$, where $\lambda = \lambda(K) > 0$.

Theorem 1. **($C^{1,\beta}$ estimate, interior.)** *If $B_R(x_0) \subset \Omega$, if $\theta > 0$, if $*$, $(**)$ both hold in Ω, and if u is a C^2 solution of (P) with $|u|_{0,B_R(x_0)} + R|Du|_{0,B_R(x_0)} \le K$ (where $K > 0$ is an arbitrary constant), then*

$$|u|_{1,\beta,B_{\theta R}(x_0)} \le C,$$

where $C = C(n,\theta,K,\Gamma(K)/\lambda(K))$ and $\beta = \beta(n,\theta,K,\Gamma(K)/\lambda(K)) \in (0,1)$.

Theorem 2. **($C^{1,\beta}$ estimate, boundary.)** *Suppose the assumptions are as in Theorem 1, except that the assumption $B_R(x_0) \subset \Omega$ is replaced by the requirements that $u = 0$ on $\partial\Omega \cap B_R(x_0)$, that $x_0 \in \partial\Omega$, and that $\partial\Omega$ is C^2 near Ω in the following sense: There is a diffeomorphism $\Phi : B_R(x_0) \to U$ (U open in $\mathbb{R}^n$) such that Φ, Φ^{-1} have C^2 norm $\le \Gamma_1$ and Φ "flattens" the boundary in the usual sense that*

$$\Phi(\Omega \cap B_R(x_0)) = U_+ \text{ and } \Phi(\partial\Omega \cap B_R(x_0)) = T,$$

where $U_+ = \{x \in U : x^n > 0\}$ and $T = \{x \in U : x^n = 0\}$. Then

$$|u|_{1,\beta,\Omega\cap B_{\theta R}(x_0)} \le C,$$

where $C = C(\theta,K,\Gamma(K)/\lambda(K),\Gamma_1)$.

Proof of Theorem 1: We first differentiate the equation with respect to x^ℓ, where $\ell \in \{1,\ldots,n\}$, asssuming for the moment that $u \in C^3(\Omega)$. This gives

$$(1) \qquad \sum_{i,j=1}^n \big(a_{ij}u_{ij\ell} + a_{ijp_k}u_{k\ell}u_{ij} + a_{ijz}u_\ell u_{ij} + a_{ijx^\ell}u_{ij}\big) + D_\ell f = 0,$$

where all the coefficient functions are evaluated at $(x,z,p) = (x,u,Du)$ and where subscripts (as in a_{ijp_k}) denote partial differentiation with respect to the indicated variables; we are also using the notation

$$u_i = D_i u, \quad u_{ij} = D_i D_j u, \quad u_{ij\ell} = D_i D_j D_\ell u.$$

We now write (1) in the divergence form

(2)
$$\sum_{i,j=1}^n \big(D_i(a_{ij}u_{j\ell}) + \sum_{k=1}^n a_{ijp_k}(u_{k\ell}u_{ij} - u_{ki}u_{j\ell}) + a_{ijz}(u_\ell u_{ij} - u_i u_{j\ell}) + a_{ijx^\ell}u_{ij} - a_{ijx^i}u_{j\ell}\big) + D_\ell f = 0.$$

On the other hand if we multiply through by $2u_\ell$ and sum over $\ell = 1,\ldots,n$, then we obtain

(3) $$\sum_{i,j=1}^n \big(D_i(a_{ij}v_j) - 2\sum_{\ell=1}^n a_{ij}u_{i\ell}u_{j\ell} + \sum_{k=1}^n a_{ijp_k}(u_{ij}v_k - v_ju_{ki})\big) + \sum_{i,j=1}^n \big(2a_{ijz}(vu_{ij} - u_i\sum_k u_ku_{kj}) + 2a_{ijx^\ell}u_\ell u_{ij} - 2a_{ijx^i}\sum_k u_ku_{kj}\big) + 2\sum_{\ell=1}^n u_\ell D_\ell f = 0,$$

where $v = |Du|^2$.

We will now multiply through by a constant γ in (2) and add the result to (3). With $w = \gamma D_\ell u + v$, since $\sum_\ell u_\ell D_\ell f = D_\ell(u_\ell f) - u_{\ell\ell}f$, we get an equation of the general form

(4)
$$\sum_{i,j=1}^n \big(D_i(a_{ij}D_jw) - 2\sum_{\ell=1}^n a_{ij}u_{i\ell}u_{j\ell} + \sum_{\ell=1}^n b_{ijk}w_ku_{ij} + b_{ij}u_{ij}\big) + \sum_{j=1}^n D_jf^j = 0,$$

where

(5) $$\sum_{i,j,k}|b_{ijk}| + \sum_{i,j}|b_{ij}| \le C, \quad C = C(K,\Gamma,\gamma),$$

provided $|u|_0 + |Du|_0 \le K$, and where

(6) $$f^j = (2u_j + \gamma\delta_{j\ell})f,$$

so that $|f^j| \le 3K\Gamma$. Integrating by parts in (4), we then see

(7)
$$\int_\Omega \sum_{i,j=1}^n \big(a_{ij}D_jwD_i\zeta + 2\zeta\sum_\ell a_{ij}u_{i\ell}u_{j\ell}\big) = + \int_\Omega \zeta\big(\sum_{i,j,k=1}^n b_{ijk}w_ku_{ij} + \sum_{i,j=1}^n b_{ij}u_{ij} - \sum_j f^jD_j\zeta\big)$$

for any $\zeta \in C_c^\infty(B_R(x_0))$. Taking a sequence u_k of C^3 functions which converge to u locally in Ω with respect to the C^2 norm (and noting that then

$$\sum_{i,j}a_{ij}(x,u_k,Du_k)D_iD_ju_k = f_k$$

with f_k converging to f uniformly on compact subsets of Ω), we then see that the identity (7) holds if u is merely C^2 on Ω. But by the Cauchy-Schwartz inequality we have

(8) $$\begin{cases} |\sum_{i,j,k=1}^n b_{ijk}w_ku_{ij}| \le \varepsilon\sum_{i,j=1}^n u_{ij}^2 + \varepsilon^{-1}C|Dw|^2, \\ |\sum_{i,j=1}^n b_{ij}u_{ij}| \le \varepsilon\sum_{i,j=1}^n u_{ij}^2 + \varepsilon^{-1}C, \end{cases}$$

for any $\varepsilon > 0$ and (by the ellipticity assumption $(**)$) we have

(9) $$\sum_{i,j,\ell=1}^n a_{ij}u_{i\ell}u_{j\ell} \ge \lambda\sum_{i,j}u_{ij}^2.$$

Hence taking $\varepsilon > 0$ small enough and using (8), (9) we obtain from (6) that

(10) $$\int_\Omega \sum_{i,j=1}^n a_{ij}D_jwD_i\zeta \le C\int_\Omega (1+|Dw|^2)\zeta + \int_\Omega \sum_{j=1}^n f^jD_j\zeta$$

for all $\zeta \in C_c^\infty(\Omega)$ with $\zeta \ge 0$, where $C = C(K,\Gamma,\lambda,\gamma) > 0$. Replacing ζ by $e^{\beta w}\zeta$ in this inequality, with $\beta > 0$, we then have

(11)
$$\int_\Omega e^{\beta w}\sum_{i,j=1}^n a_{ij}D_jwD_i\zeta \le C\int_\Omega \big((1+|Dw|^2) - \lambda\beta|Dw|^2\big)\zeta + \int_\Omega e^{\beta w}f^jD_j\zeta,$$

for all $\zeta \ge 0,\ \zeta \in C_c^\infty(\Omega)$.

Then, selecting $\beta > C/\lambda$, we have

(12) $$\int_\Omega e^{\beta w}\sum_{i,j=1}^n a_{ij}D_jwD_i\zeta \le \int_\Omega (C + e^{\beta w}f^jD_j\zeta),$$

for all $\zeta \ge 0,\ \zeta \in C_c^\infty(\Omega)$.

Now take fixed $\rho_0 > 0$ with $B_{\rho_0}(x_0) \subset \Omega$, take $\rho \le \rho_0$, $Q_{\ell,\gamma} = \sup_{B_\rho(x_0)} w$ ($\equiv \sup_{B_\rho}(\gamma D_\ell u + v)$), and $\widetilde{w} = Q_{\ell,\gamma} - w + \rho/\rho_0$. Then notice that

$$\int_\Omega \sum \widetilde{a}_{ij}D_j\widetilde{w}D_i\zeta \ge \int_\Omega (C\zeta + \sum \hat{f}_jD_j\zeta),$$

where $\hat{f} = \widetilde{w}^{-1}C$ and $\hat{f}_j = \widetilde{w}^{-1}\widetilde{f}_j$, so that

$$|\hat{f}| \le C\rho_0/\rho \le C(\rho_0/\rho)^2, \qquad |\hat{f}_j| \le C\rho_0/\rho.$$

We can then apply Theorem 2 of Lecture 17 (with $q = \infty$, $p = 1$, and with ρ in place of R) in order to conclude

$$C^{-1}\rho^{-n}\int_{B_{\rho/2}(x_0)} \widetilde{w} \le \inf_{B_{\rho/2}(x_0)} \widetilde{w},$$

where $C = C(\lambda,\Gamma,K,\gamma,n)$. In particular, using the definition of $\widetilde{w}$, we conclude that

(13) $$C^{-1}\rho^{-1}\int_{B_{\rho/2}(x_0)} (Q_{\ell,\gamma} - \gamma D_\ell u - v) \le \inf_{B_{\rho/2}(x_0)} (Q_{\ell,\gamma} - \gamma D_\ell u - v) + \rho/\rho_0,$$

where $C = C(\Gamma,K,\lambda,\gamma,n)$. Now for each $\ell = 1,\dots,n$ define

$$w_\ell^\pm = \pm 10nMD_\ell u + v,$$
$$M = \sup_{B_\rho(x_0)} |Du|$$
$$W_\ell^\pm = \sup_{B_\rho(x_0)} w_\ell^\pm.$$

(Notice that thus $W_\ell^\pm = Q_{\ell,\gamma}$ with $\gamma = \pm 10nM$ respectively.) Using (13) with $\gamma = \pm 10nM$ respectively, and adding the resultant inequalities, we conclude

$$(14)\quad C^{-1}\rho^{-n}\int_{B_{\rho/2}(x_0)} (W_\ell^+ + W_\ell^- - 2v) \le \inf_{B_{\rho/2}(x_0)} (W_\ell^+ - w_\ell^+) + \inf_{B_{\rho/2}(x_0)} (W_\ell^- - w_\ell^-) + 2\rho/\rho_0.$$

Now notice that

$$10nM \operatorname{osc}_{B_\rho(x_0)} D_\ell u - \operatorname{osc}_{B_\rho(x_0)} v \le \operatorname{osc}_{B_\rho(x_0)} w_\ell^\pm \le 10nM \operatorname{osc}_{B_\rho(x_0)} D_\ell u + \operatorname{osc}_{B_\rho(x_0)} v,$$

and since

$$(15)\quad \begin{aligned} \operatorname{osc}_{B_\rho(x_0)} v &\equiv \sup_{B_\rho(x_0)} v - \inf_{B_\rho(x_0)} v \\ &= \sup_{x,y\in B_\rho(x_0)} (|Du|^2(x) - |Du|^2(y)) \\ &= \sup_{x,y\in B_\rho(x_0)} (Du(x) + Du(y))\cdot(Du(x) - Du(y)) \\ &\le 2nM \max{}_j \operatorname{osc}_{B_\rho(x_0)} D_j u. \end{aligned}$$

Thus choosing ℓ (depending on ρ) such that

$$(16)\quad \operatorname{osc}_{B_\rho(x_0)} D_\ell u = \max{}_j \operatorname{osc}_{B_\rho(x_0)} D_j u,$$

we conclude that

$$(17)\quad 8nM \operatorname{osc}_{B_\rho(x_0)} D_\ell u \le \operatorname{osc}_{B_\rho(x_0)} w_\ell^\pm \le 12nM \operatorname{osc}_{B_\rho(x_0)} D_\ell u.$$

Notice also that (16) implies

$$(18)\quad \begin{aligned} \inf_{B_\rho(x_0)} (W_\ell^+ + W_\ell^- - 2v) &\equiv W_\ell^+ + W_\ell^- - 2\sup_{B_\rho(x_0)} v \\ &\ge 10nM \sup_{B_\rho(x_0)} D_\ell u - 10nM \sup_{B_\rho(x_0)} D_\ell u + 2\inf_{B_\rho(x_0)} v - 2\sup_{B_{\rho/2}(x_0)} v \\ &\ge 10nM \operatorname{osc} D_\ell u - 2\operatorname{osc} v \\ &\ge 6nM \operatorname{osc} D_\ell u \ge \tfrac{1}{2}\operatorname{osc}_{B_\rho(x_0)} w_\ell^\pm \end{aligned}$$

Using (17) and (18) in (14) (and using also the fact that $\sup_{B_\rho} f - \inf_{B_{\rho/2}} f \le \operatorname{osc}_{B_\rho} f - \operatorname{osc}_{B_{\rho/2}} f$ for any function f), we then have

$$\begin{aligned} &C^{-1}(\operatorname{osc}_{B_\rho(x_0)} w_\ell^+ + \operatorname{osc}_{B_\rho(x_0)} w_\ell^-) \le C^{-1}\operatorname{osc}_{B_\rho(x_0)} D_\ell u \\ &\quad \le \operatorname{osc}_{B_\rho(x_0)} w_\ell^+ + \operatorname{osc}_{B_\rho(x_0)} w_\ell^- - \operatorname{osc}_{B_{\rho/2}(x_0)} w_\ell^+ - \operatorname{osc}_{B_{\rho/2}(x_0)} w_\ell^- + 2\rho/\rho_0, \end{aligned}$$

so that

$$\omega_\ell(\rho/2) \le \theta\omega_\ell(\rho),$$

where

$$\omega_j(\rho) = \operatorname{osc}_{B_\rho(x_0)} w_j^+ + \operatorname{osc}_{B_\rho(x_0)} w_j^-, \quad j = 1,\ldots,n.$$

Now notice that, by (16) and (17), we have

$$\omega_\ell(\rho) \ge C^{-1}\max{}_j \omega_j(\rho),$$

and of course $\omega_j(\rho/2) \le \omega_j(\rho)$ and hence

$$\begin{aligned} \textstyle\sum_{j=1}^n \omega_j(\rho/2) &= \textstyle\sum_{j\ne\ell}\omega_j(\rho/2) + \omega_\ell(\rho/2) \\ &\le \textstyle\sum_{j\ne\ell}\omega_j(\rho) + \theta\omega_\ell\rho + C\rho/\rho_0 \\ &= \textstyle\sum_{j=1}^n \omega_j(\rho) - (1-\theta)\omega_\ell(\rho) + c\rho/\rho_0 \\ &\le (1 - \frac{1-\theta}{nC})\textstyle\sum_{j=1}^n \omega_j(\rho) + C\frac{\rho}{\rho_0}. \end{aligned}$$

Thus

$$(19)\quad \textstyle\sum_{j=1}^n \omega_j(\rho/2) \le \mu\sum_{j=1}^n \omega_j(\rho) + C\rho/\rho_0, \quad \rho\in(0,\rho_0)$$

where $\mu = 1 - \frac{1}{nC} + \frac{\theta}{nC} \in (\frac{1}{2}, 1)$ depends only on $\Gamma, \lambda, n, M, \rho_0$. By iterating (19) we obtain, for any integer $\nu > 0$,

$$\begin{aligned} \textstyle\sum_{j=1}^n \omega_j(2^{-\nu}\rho_0) &\le \mu^j \textstyle\sum_{j=1}^n \omega_j(\rho_0) + C\sum_{k=0}^{\nu} \mu^{\nu-k}2^{-k} \\ &\le (\textstyle\sum_{j=1}^n \omega_j(\rho_0) + 2C)\mu^\nu. \end{aligned}$$

For any $\rho \in (0,\rho_0)$, we can find an integer ν such that $2^{-\nu-1}\rho_0 \le \rho \le 2^{-\nu}\rho_0$, and hence, taking $\beta \in (0,1)$ such that $2^{-\beta} = \mu$, this gives

$$\textstyle\sum_{j=1}^n \omega_j(\rho) \le C(\rho/\rho_0)^\beta, \qquad \rho\in(0,\rho_0),$$

where $C > 0$ and $\beta \in (0,1)$ depend only on Γ, λ, n, M. But then using the left inequality in (17) we deduce

$$*\operatorname{osc}_{B_\rho(x_0)} D_j u \le C(\rho/\rho_0)^\beta, \qquad \rho\in(0,\rho_0),\ j = 1,\ldots,n.$$

In view of the arbitrariness of ρ, x_0, this evidently directly implies the required Hölder estimate for $D_j u$.

Proof of Theorem 2: We begin by making the coordinate transformation $y = \Phi(x)$, $x \in \Omega \cap B_R(x_0)$, so that $u(x) \equiv \widetilde{u}(y)$, where, on U_+, $\widetilde{u}$ satisfies an equation of the same form as the equation for u on $\Omega \cap B_R(x_0)$. It follows that there is no loss of generality in assuming that $x_0 = 0$ and

$$\Omega \cap B_R(x_0) = B_R^+(0) \equiv \{x \in B_R(0) : x^n > 0\}.$$

We shall henceforth make this assumption. Also if $\ell \in \{1, \dots, n-1\}$, then $D_{y^\ell}\widetilde{u} = D_{y^\ell}\widetilde{\psi}$, which is a given C^1 function on T, hence we can repeat the above argument with $\ell \in \{1, \dots, n-1\}$; notice that in particular in place of (4) we get a similar identity with sums with respect to ℓ over $1, \dots, n-1$ only. However

$$\sum_{i,j=1}^n \sum_{\ell=1}^{n-1} a_{ij} u_{i\ell} u_{j\ell} \ge \lambda \sum_{j=1}^n \sum_{\ell=1}^{n-1} u_{j\ell}^2 \ge C^{-1} \sum_{j,\ell}^n u_{j\ell}^2 - C \tag{20}$$

for suitable constant C, by virtue of the fact that we can use the equation to express u_{nn} in terms of the other derivatives thus:

$$u_{nn} = -\sum_{(i,j)\neq(n,n)} a_{nn}^{-1} a_{ij} u_{ij} + a_{nn}^{-1} f.$$

Hence the same argument as before applies to again give an inequality like (12) for any $\ell = 1, \dots, n-1$. Furthermore, by a reflection argument as in the proof of the boundary estimate of Theorem 3 above, the odd relection of $D_\ell u$ satisfies an inequality like (12) on the whole ball $B_R(0)$. The rest of the argument then carries through, showing that

$$|D_\ell u(x) - D_\ell u(y)| \le C\Big(\frac{|x-y|}{R}\Big)^\beta, \quad x, y \in \Omega \cap B_{\theta R}(x_0),\ \ell = 1, \dots, n-1. \tag{21}$$

Thus it remains only to check that a similar estimate holds with $\ell = n$. However (12) implies that

$$\int_{\Omega \cap B_{\rho/2}(x_0)} |DD_\ell u|^2 \le C\rho^{-2} \int_{\Omega \cap B_\rho(x_0)} (\sup_{B_\rho(x_0)} D_\ell u - D_\ell u(x))^2\, dx,$$

for each $\ell = 1, \dots, n-1$. But then using (21) we have

$$\int_{\Omega \cap B_{\rho/2}(x_0)} |DD_\ell u|^2 \le C\rho^{n-2+2\beta}, \quad \rho \in (0, \rho_0).$$

But, using (20) again, we then have by Morrey's lemma (see Lecture 16) that there is also a Hölder estimate for $D_n u$, as required.

Lecture 19

The Minimal Surface & Mean Curvature Equations

The minimal surface equation (abbreviated MSE), is the Euler-Lagrange equation of the area functional

$$\mathcal{A}(u) = \int_\Omega \sqrt{1 + |Du|^2}\, dx.$$

Notice that the area functional is has the general form of a functional $\mathcal{F}(u) = \int_\Omega f(Du)\, dx$, in the special case when

$$f(p) = \sqrt{1 + |p|^2}, \quad p \in \mathbb{R}^n,$$

in which case

$$f_{p_i}(p) = \frac{p_i}{\sqrt{1+|p|^2}}, \quad f_{p_i p_j} = (1+|p|^2)^{-1/2}\Big(\delta_{ij} - \frac{p_i p_j}{1+|p|^2}\Big),$$

so $(f_{p_i p_j}(p))$ is positive for each p. Indeed (see Problem 16.4), the eigenvalues of $((1+|p|^2)^{1/2} f_{p_i p_j}(p))$ are 1, with multiplicity $n-1$ and $(1+|p|^2)^{-1}$ with multiplicity 1.

The Euler-Lagrange equation for the area functional (i.e. the minimal surface equation) is

$$\sum_{i=1}^n D_i\Big(\frac{D_i u}{\sqrt{1+|Du|^2}}\Big) = 0,$$

or, equivalently,

$$(1+|Du|^2)^{-1/2} \sum_{i,j=1}^n \Big(\delta_{ij} - \frac{D_i u D_j u}{1+|Du|^2}\Big) D_i D_j u = 0.$$

The operator on the left is called the minimal surface operator, and will be denoted $\mathcal{M}$.

The prescribed mean curvature equation is the equation

$$\mathcal{M}(u) = H(x, u),$$

where $H(x, z)$ is a given function on $\Omega \times \mathbb{R}$. Notice that this equation is the Euler-Lagrange equation of the functional

$$\mathcal{F}(u) = \mathcal{A}(u) + \int_\Omega \int_0^{u(x)} H(x,t)\,dt\,dx \equiv \int_\Omega \Big(\sqrt{1+|Du|^2} + \int_0^{u(x)} H(x,t)\,dt\,dx.$$

First we want to discuss the application of the Leray-Schauder theory developed in Lecture 15 to the sovability of the Dirichlet problem for the minimal surface equation and the mean curvature equation. Indeed we already saw (see problems 16.XX, above) that the Dirichlet problem for the minimal surface equation is solvable on a smooth uniformly convex domain with smooth boundary data. Here we show that actually a more precise result holds.

So let Ω be a bounded $C^{2,\mu}$ domain in $\mathbb{R}^n$, and let ψ be a given $C^{2,\mu}(\overline{\Omega})$ function. The Dirichlet problem is

$$(*) \qquad \begin{cases} \mathcal{M}(u) = 0 \text{ on } \Omega \\ \quad u = \psi \text{ on } \partial\Omega \end{cases}$$

Recall that any solution of this problem actually minimizes the area functional relative to all $C^1(\overline{\Omega})$ functions which agree with u on $\partial\Omega$. Thus a solution of the Dirichlet problem $(*)$ solves the geometric variational problem of finding the the graph which is the surface of least area relative to all comparison graphs over $\overline{\Omega}$ with boundary given by $\Gamma = \{(x, \varphi(x)) : x \in \partial\Omega\}$. Such graphs are thus special examples of minimal surfaces, the study of which is an important topic in differential geometry; the minimal surface operator itself (and closely related operators) arise naturally in a number of physically and geometrically motivated problems; for instance see examples 11, 12, 13, 14 of Lecture 1.

The Dirichlet problem discussed above, over general $C^{2,\mu}$ domains, provides a classic example of the way in which the geometry of the domain enters into consideration in the study of existence and qualitative behaviour of solutions of quasilinear equations. We want to spend a little time discussing this.

First recall that the problem $*$ above can be solved provided we can find suitable upper and lower boundary barrier functions near each boundary point

x_0, because this gives *a-priori* bounds on $\sup_{\partial\Omega}|Du|$, and then the necessary *a-priori* bounds for $|u|_0 + |Du|_0$ hold by the maximum principle, thus making it possible to apply the Leray-Schauder mathod.

We want to show that the possibility of constructing such barrier functions for arbitrary given $\varphi \in C^2(\overline{\Omega})$ depends on a certain geometric restriction on the boundary, namely that the mean curvature of $\partial\Omega$ (computed relative to the outward pointing unit normal) should be non-negative. We proceed to discuss this, and at the same time explain the terminology for those not already familiar with it.

First, let d be the distance function for $\partial\Omega$; recall that this is defined by

$$d(x) = \begin{cases} \ \ \operatorname{dist}(x, \partial\Omega) \text{ if } x \in \Omega \\ -\operatorname{dist}(x, \partial\Omega) \text{ if } x \in \mathbb{R}^n \setminus \Omega. \end{cases}$$

Recall also that for suitably small $\delta > 0$ we have the following facts on the strip $W \equiv \{x \in \mathbb{R}^n : \operatorname{dist}(x, \partial\Omega) < \delta\}$:

(a) *The nearest point projection of a point $x \in W$ onto $\partial\Omega$ is uniquely defined, and is given by*

$$x \mapsto P(x) \equiv x - d(x)Dd(x), \quad x \in W;$$

(b)
$$Dd(x) \equiv Dd(P(x))$$

(so that Dd is constant on the line segment joining x to $P(x)$, and this line segment meets $\partial\Omega$ orthogonally to the hypersurface $\partial\Omega$ at the point $P(x)$). We also have

(c)
$$|Dd|^2 \equiv 1 \text{ on } W.$$

Notice that by differentiating the relation (c) we get the identity

(d)
$$\textstyle\sum_{j=1}^n (D_j d)(D_i D_j d) \equiv 0 \text{ on } W,$$

and by differentiating the relation (b) and using (d) we get

$$D_i D_j d(x) \equiv \textstyle\sum_{k=1}^n D_i D_k d(P(x))(\delta_{jk} - d(x) D_j D_k d(x)),$$

so that the Hessian matrix $\operatorname{Hess} d$ satisfies

(e)
$$\operatorname{Hess} d(x) = \operatorname{Hess} d(P(x))\,(I + d(x)\operatorname{Hess} d(P(x)))^{-1},$$

provided δ (in the definition of W) is small enough to ensure that the maximum eigenvalue of $\delta \operatorname{Hess} d < 1$. Notice in particular if δ is small enough to ensure that the maximum eigenvalue is in fact $< 1/2$, which we subsequently assume, then (e) gives

(f) $$\|\mathrm{Hess}\, d(x) - \mathrm{Hess}\, d(P(x))\| \le C(n) d(x) \|\mathrm{Hess}\, d(P(x))\|$$

for $x \in W$. Also the relation (d) tells us that $Dd(x)$ is an eigenvector for $\mathrm{Hess}\, d(x)$ and $\mathrm{Hess}\, d(P(x))$ with eigenvalue 0. Thus the other eigenvectors form an orthonormal basis for the tangent space of $\partial\Omega$ at at $P(x)$; if $x \in \partial\Omega$, the eigenvalues corresponding to these tangential eigenvectors are $-\kappa_1, -\kappa_2, \ldots, -\kappa_{n-1}$, where κ_j denote the principal curvatures of $\partial\Omega$ at x. The trace of the Hessian at x, that is the Laplacian of d at x is then $-\sum_{j=1}^{n-1}\kappa_j$, also known as the mean curvature of $\partial\Omega$ at x. Thus non-negative mean curvature of $\partial\Omega$ is equivalent to the condition

(g) $$\Delta d(x) \le 0 \text{ on } \partial\Omega.$$

We henceforth assume this. What we then want to prove is

Lemma. *If $\partial\Omega$ is a bounded C^2 domain and if (g) holds, then for any $\varphi \in C^2\overline{\Omega}$, and any constant $M > 0$ there are upper and lower barrier functions $w^\pm \in C^2(\overline{W}^+)$, where $W^+ \equiv \{x \in \Omega : \mathrm{dist}\,(x, \partial\Omega) < \delta\}$, satisfying*

(**) $$\begin{cases} \mathcal{M}(w^+) \le 0, \quad \mathcal{M}(w^-) \ge 0 \text{ on } \overline{W}^+ \\ w^\pm \equiv \varphi \text{ on } \partial\Omega \\ w^+ \ge M, \quad w^- \le -M \text{ on } S \equiv \{x \in \Omega : \mathrm{dist}\,(x, \partial\Omega) = \delta\}. \end{cases}$$

Proof: Without loss of generality we can take the constant δ in the above discussion of the distance function to be $\le 1/2$. We look for $w = w^+$ in the form

$$w(x) = \varphi(P(x)) + \gamma(d(x)), \quad x \in W,$$

where $W = \{x \in \mathbb{R}^n : \mathrm{dist}\,(x, \partial\Omega) < \delta\}$ as above, and where P is the nearest point projection described above.

By direct computation we have, writing $\tilde\varphi(x) = \varphi(P(x))$,

(1) $$w_i = \tilde\varphi_i + \gamma'(d)d_i, \quad w_{ij} = \tilde\varphi_{ij} + \gamma''(d)d_i d_j + \gamma'(d)d_{ij},$$

where the subscripts denote partial derivatives. Notice that by virtue of (c) above and the fact that $\tilde\varphi$ is constant along line segments in the direction Dd, we have the relations

(2) $$\textstyle\sum_{j=1}^n d_j\tilde\varphi_j = 0 \ \sum_{j=1}^n d_j d_{ij} \equiv 0, \ \sum_{j=1}^n d_j d_j \equiv 1.$$

Then by direct computation using these relations we see that

$$\begin{aligned}(1+|Dw|^2)^{3/2}\mathcal{M}(w) &\equiv (1+|D\tilde\varphi|^2 + (\gamma'(d))^2)(\Delta\tilde\varphi + \gamma'(d)\Delta d + \gamma''(d)) \\ &\quad - \textstyle\sum_{i,j=1}^n(\tilde\varphi_i\tilde\varphi_j + (\gamma')^2 d_i d_j + \gamma'\tilde\varphi_i d_j + \gamma'\tilde\varphi_j d_i)(\tilde\varphi_{ij} + \gamma'' d_i d_j + \gamma' d_{ij}) \\ &\le (1 + (\gamma')^2 + |D\tilde\varphi|^2)(C + \gamma'\Delta d + \gamma'') + C - \gamma'\tilde\varphi_i\tilde\varphi_j d_{ij} + \\ &\qquad C(\gamma')^2 - \gamma''(\gamma')^2 - 2\tilde\varphi_i\tilde\varphi_{ij}d_j\gamma' \\ &\le (1+(\gamma')^2 + |D\varphi|^2)\gamma'\Delta d + C(1+(\gamma')^2) - C\gamma'',\end{aligned}$$

provided we take $\gamma' \ge 1$ and $\gamma'' \le 0$.

Now by the bounds (f), (g) we have

$$(1 + (\gamma')^2 + |D\varphi|^2)\gamma'\Delta d \le C(\gamma'(d))^3 d \le C(\gamma'(d))^2,$$

provided we agree to select γ such that

(3) $$d\gamma'(d) \le 1, \quad \gamma'(d) \ge 1.$$

Thus, subject to the resrictions (3) and $\gamma'' \le 0$, we obtain

$$(1+|Dw|^2)^{3/2}\mathcal{M}(w) \le C(\gamma')^2 - C\gamma''.$$

We note that by selecting $\gamma(d) = \log(1 + kd)$, with $k \ge 2$ and $\log(1 + k\delta) \ge 2M$, we then have the required restrictions $\gamma' \ge 1$, $d\gamma'(d) \le 1$, and $\gamma'' \le 0$. Therefore $\mathcal{M}(w) \le 0$ in W and (since $\gamma(\delta) \ge 2M$) $w \ge M$ on $\{x \in \Omega : \mathrm{dist}\,(x, \partial\Omega) = \delta\}$.

Next we want to derive some interior and global estimates for continuity and the gradient of solutions of the minimal surface equation. The results here are make it possible to prove very strong compactness theorems (see the later discussion below).

The key result here is the following; we emphasize that there is no restriction on the domain Ω, except for boundedness, in this.

Theorem 1. *Suppose Ω is a bounded subset of $\mathbb{R}^n$, $\varphi \in C^2(\mathbb{R}^n)$, and $u \in C^2(\Omega) \cap C^0(\overline{\Omega})$ is a solution of the minimal surface equation on Ω with $u = \varphi$ on $\partial\Omega$. Then, with $\gamma \ge 1$ any constant such that $|\varphi|_{C^2} \le \gamma$, we have*

$$|D(u-\varphi)^2| \le 3e^{10\gamma^2}.$$

Proof: Recall that the minimal surface equation is

$$\textstyle\sum_{i=1}^n \frac{D_i u}{\sqrt{1+|Du|^2}} = 0,$$

or, equivalently,

$$\textstyle\sum_{i,j=1}^n g^{ij} D_i D_j u = 0,$$

where $g^{ij} = \delta_{ij} - (1+|Du|^2)^{-1}(D_i u)(D_j u)$.

Next we recall that by differentiating the divergence-form of the equation we get the identity

$$\textstyle\sum_{i,j=1}^n D_i\big(v^{-1} g^{ij} D_j u_\ell\big) = 0, \quad \ell = 1,\dots,n,$$

where $u_\ell = D_\ell u$ and $v = \sqrt{1+|Du|^2}$. By multiplying by u_ℓ and suming on ℓ, this yields the identity

$$\textstyle\sum_{i,j,\ell}^n v^{-1} g^{ij} u_{i\ell} u_{j\ell} - \sum_{i,j=1}^n D_i(g^{ij} D_j v) = 0.$$

Equivalently, since $D_j v = -v^{-2} D_j(v^{-1})$, this identity can be written

$$v^{-1}\Big(Q + \frac{1}{v}\textstyle\sum_{i,j=1}^n D_i\big(v g^{ij} D_j\big(\frac{1}{v}\big)\big)\Big) = 0, \tag{1}$$

where

$$Q = \frac{1}{v^2}\textstyle\sum_{i,j,k,\ell} g^{ij} g^{k\ell} u_{ik} u_{j\ell} \ge 0.$$

(See Problem 31.1.) Thus in particular we obtain

$$L\big(\frac{1}{v}\big) \le 0, \tag{2}$$

where L is the second order operator given by

$$L(\psi) = \frac{1}{v}\textstyle\sum_{i,j=1}^n D_i\big(v g^{ij} D_j \psi\big).$$

Geometrically the quantity Q is the sum of squares of principal curvatures of graph u and L is the Laplacian on graph u in the coordinate system $x^1,\dots,x^n$, but we make no use of these facts here. The reader familiar with the appropriate background in geometry will also notice that the equation (1) is the Jacobi-field equation for the quantity $e_{n+1}\cdot\nu$, where $\nu = v^{-1}(-Du, 1)$ is the upward unit normal of graph u; $e_{n+1}\cdot\nu$ is the normal component of the initial velocity of graph u under the vertical translation $x^{n+1} \to x^{n+1} + t$. Again we will not explicitly use these geometric facts.

For the computations below it is convenient to note that the operator L can also be written in the non-divergence form

$$L(\psi) = \textstyle\sum_{i,j} g^{ij} D_i D_j \psi. \tag{3}$$

This is readily checked by direct computation; it involves checking the identity

$$\textstyle\sum_i D_i\big(v g^{ij}\big) = 0, \quad j = 1,\dots,n.$$

We shall also need the fact that for any point $\xi \in \mathbb{R}^n$ we have

$$\begin{aligned}\textstyle\sum_{i,j} g^{ij}(u_i - \xi^i)(u_j - \xi^j) &= \textstyle\sum_{i,j}\big(g^{ij} u_i u_j - 2g^{ij} u_i \xi^j + g^{ij}\xi^i\xi^j\big)\\ &\ge \frac{|Du|^2}{1+|Du|^2} - \frac{\sum_j \xi^j D_j u}{1+|Du|^2},\end{aligned} \tag{4}$$

keeping in mind that the matrix (g^{ij}) has eigenvalues $1, (1+|Du|^2)^{-1}$, and that Du is the eigenvector corresponding to eigenvalue $(1+|Du|^2)^{-1}$.

Now let $\eta = e^{-2K\gamma}(e^{Kt} - 1)$, where $K > 0$ is to be chosen. Notice that then $0 \le \eta(t) \le 1$ for $t \ge 0$. We consider the function

$$w = \frac{\eta((u-\varphi)_+)}{\varepsilon + v^{-1}},$$

where $(u-\varphi)_+ = \max\{u-\varphi, 0\}$ and where $\varepsilon > 0$.

Since $\eta((u-\varphi)_+) \le 1$ and $\varepsilon > 0$, it is easy to check that, unless $u \le \varphi$ everywhere in Ω, the function w has a positive maximum value M attained at some $x_0 \in \Omega$. Notice that then $\eta(u-\varphi) - M(\varepsilon + v^{-1})$ has maximum value zero at the point x_0. Then at the point x_0, we have $L(\eta(u-\varphi) - M(\varepsilon + v^{-1})) \le 0$. By virtue of (2), (3), this gives the inequality

$$\eta''(u-\varphi)\big(\textstyle\sum_{i,j} g^{ij}(u_i - \varphi_i)(u_j - \varphi_j)\big) + \eta'(u-\varphi)\sum_{i,j} g^{ij} D_i D_j \varphi \le 0, \tag{5}$$

which evidently gives

$$K\textstyle\sum_{i,j} g^{ij}(u_i - \varphi_i)(u_j - \varphi_j) \le \gamma,$$

and by (4) this gives

$$\Big(\frac{|Du|^2}{1+|Du|^2} - \frac{\gamma|Du|}{1+|Du|^2}\Big) \le \frac{\gamma}{K}. \tag{6}$$

Now if $M \ge 4\gamma + 2$, then since $\eta(u-\varphi) \le 1$ we would have $v \ge 4\gamma + 2$, hence $|Du| \le 4\gamma + 1$, which evidently implies from (6) that $1/4 \le \gamma/K$, which is impossible if we select $K = 5\gamma$.

Thus we deduce that $M \le 4\gamma + 2$, hence (after letting $\varepsilon \downarrow 0$)

$$\big(e^{5\gamma(u-\varphi)_+} - 1\big)v \le (4\gamma + 2)e^{10\gamma^2}.$$

Since $e^y - 1 \ge y$ for $y \ge 0$, this gives

$$(u-\varphi)_+ v \le 2e^{10\gamma^2}.$$

Similarly $(\varphi - u)_+ \le 2e^{10\gamma^2}$, and hence $|u-\varphi||Du| \le 2e^{10\gamma^2}$, so (since $|u - \varphi|\,|D\varphi| \le 2\gamma^2$)

$$|u-\varphi||D(u-\varphi)| \le 2e^{10\gamma^2} + 2\gamma^2,$$

which is the required inequality.

We note now a couple of important corollaries of Theorem 1.

Corollary 1. (Interior gradient estimate.) *If u is a C^2 solution of the MSE on the open set $\Omega \subset \mathbb{R}^n$, and if $B_\rho(x_0) \subset \Omega$ and if $u \le u_{x_0} + M$ on $B_\rho(x_0)$, then*

$$|Du(x_0)| \le C_1 e^{C_2 M/\rho},$$

where C_1, C_2 are constants depending only on n.

Proof: Let $\varphi(x) = (M+\rho)\rho^{-2}|x - x_0|^2 - \rho + u(x_0)$, and let $\tilde\Omega = \{x \in B_\rho(x_0) : u(x) > \varphi(x)\}$. Notice that then $\overline{\tilde\Omega} \subset B_\rho(x_0)$ ($B_\rho(x_0)$ is the <u>open</u> ball), and $u = \varphi$ on $\partial\tilde\Omega$, and $u(x_0) - \varphi_{x_0} = \rho$. Thus Theorem 1 applies on $\tilde\Omega$; since $\gamma \le 4n^2(M+\rho)$ in this case, the corollary is then proved.

Corollary 2. *If $u \in C^2(\Omega) \cap C^0(\overline{\Omega})$, and if $u = \varphi$ on $\partial\Omega$, where $\varphi \in C^2(R^n)$, then*

$$|u(x) - u(y)| \le 2|\varphi(x) - \varphi(y)| + Ce^{5\gamma^2}|x-y|^{1/2} \quad \forall\, x, y \in \overline{\Omega},$$

where $\gamma = 1 + |\varphi|_{C^2}$ and C depends only on n.

Note: This says in particular that u is Hölder continuous with exponent 1/2 on $\overline{\Omega}$. In view of the fact that the domain Ω is an arbitrary domain, this is rather surprising; such a result is certainly <u>not</u> true for harmonic functions.

Proof: By Theorem 1 we have

$$|Dw| \le Ce^{10\gamma^2},$$

where w is the Lip $(\mathbb{R}^n)$ function obtained by extending $(u-\varphi)^2$ to zero outside Ω. (See Problem 31.2.) By integration along line segments, we now easily obtain the required inequality.

Finally we want to prove the following consequence of Corollary 1.

Theorem 2. *Suppose $u \in C^2(\mathbb{R}^n)$ satisfies the MSE on all of $\mathbb{R}^n$, and suppose also that there is a fixed constant C such that*

$$u(x) \le C|x| \quad \forall\, x \in \mathbb{R}^n \text{ with } |x| \ge 1.$$

Then u is affine (i.e., linear + constant).

Proof: By Corollary 1, we have

$$|Du(x_0)| \le C_1 \exp(C_2(R+\rho)/\rho)$$

for any $x_0 \in \mathbb{R}^n$, with $R = |x_0|$ and $\rho > 0$ arbitrary. Then letting $\rho \uparrow \infty$, we deduce that $|Du|$ is bounded on $\mathbb{R}^n$. But then $w = D_\ell u$ is a bounded solution on $\mathbb{R}^n$ of the uniformly elliptic equation

$$\textstyle\sum_{i,j=1}^n D_i\left(a_{ij} D_j w\right) = 0,$$

where $a_{ij} = (1+|Du|^2)^{-1/2}(\delta_{ij} - (1+|Du|^2)^{-1} D_i u D_j u)$.

Then by the De Giorgi Nash theory (see Lecture 17) we have

$$|D_\ell u(x) - D_\ell u(y)| \le C(|x-y|/\rho)^\alpha, \quad \forall\, x, y \in B_{\rho/2}(0)$$

for any $\rho > 0$, where $C > 0$ and $\alpha \in (0,1)$ are fixed constants. Letting $\rho \uparrow \infty$, we deduce that $D_\ell u \equiv$ constant.