INTRODUCTION TO PDE

ALEX MRAMOR

Abstract. Course notes for NMAK16022U, an introductory graduate course on PDE given at Copenhagen University during block one of the '23-'24 academic year.

Contents

1. Some practical info about the course and the notes

Contact: Alex Mramor, emails: almr@math.ku.dk, amramor-math@outlook.com – I check the second email more often.

Where the course will be: the exercise section is held on Mondays 10–12 at øv - A107, Universitetsparken 5, HCØ by Marco Olivieri. The lectures will be held Tuesdays and Fridays: the Tuesday meeting will be 13–16 at øv - bib 4-0-17, Universitetsparken 1-3, DIKU and the Friday meeting will be 10–12 at øv - A112, Universitetsparken 5, HCØ.

Evaluation: There will be an end of term exam which is what the grade will based off of. There will be weekly ungraded homework, but in the exercise section there will be a quiz based at least loosely off the homework. These quizzes, which won't count towards the grade, will still be checked and should be useful practice for the exam.

About these notes and whats most important: As a rule of thumb, if a statement given below doesn't specifically involve second order PDE or the development of prerequisite theory, it probably isn't material which will be involved in the end of block test. Such statements might only be mentioned in the lectures, depending on time or how much "extra" nonprequisite background they involve (say, from geometry) – these are still worthwhile looking over and I included them because I thought they were still important or had an idea one should at least know about. Concerning the statements out of Evans [5], which is the main source for these notes, or any of the other sources below I tend to follow their arguments pretty closely but I often fill in details to my taste and add commentary. Occasionally I may also indicate other methods of arguing. Having a copy of Evans or access to the other books in the references isnt a bad idea, even just to peruse them, but not essential.

2. What PDE are, what types of complications can there be, and what types of questions we will be interested in

PDEs stand for partial differential equations, where the "partial" here indicates that the equations involve functions and derivatives thereof in several variables. They appear naturally in a vast array of physics and engineering (take Maxwell's equations, for instance, or Schrodinger's equation) but are also intrinisically interesting from a pure perspective and have applications in other fields of pure mathematics.

We start by introducing a few model PDE. I'm a geometric analyst, and being admittedly provincial I'll center the discussion to follow around it but there are many tacks (some perhaps more justified at least from a historical perspective) one can follow. In geometric analysis, a large and relatively new field of mathematics, one theme is studying manifolds with special/"good" geometry and ways to deform manifolds to have good geometry. Such manifolds typically solve in some manner a partial differential equation which resembles at some level the Laplace equation on $\mathbb{R}^n$:

$$\Delta u = \frac{d^2u}{dx_1^2} + \ldots + \frac{d^2u}{dx_n^2} = 0 \tag{2.1}$$

Take, for instance, the minimal surface equation, describing surfaces of locally least area or equivalently surfaces of vanishing mean curvature. Just to be concrete the minimal surface equation for graphs is:

$$(1 + u_y^2)u_{xx} - 2u_x u_y u_{xy} + (1 + u_x^2)u_{yy} = 0 \tag{2.2}$$

Similarly, rules to deform manifolds can often be written in some manner which resembles the heat equation on $\mathbb{R}^n \times \mathbb{R}$:

$$\frac{du}{dt} - \Delta u = 0 \tag{2.3}$$

For example the Ricci flow which was used to solve Poincare's conjecture, and other flows such as the mean curvature flow. There are also cases, for instance in the mathematical study of general relativity (which depending on the flavor belongs or is at least relatively near to geoemetric analysis) that one is interested in equations that resemble the wave equation:

$$\frac{d^2u}{dt^2} - \Delta u = 0 \tag{2.4}$$

As one might expect, the actual equations in geometry/physics/life are more complicated than the "model" equations given above but these equations and their generalizations will form the core of the course. As food for thought let's list off some ways PDE one might be interested in can be different and relatively harder to understand, what extra information might be relevant, and more positively why the list of equations above is satisfactorily large to consider in some sense:

- Notice that in the above examples we didn't really carefully specify any initial conditions/boundary data. For a specifc example with the Laplace equation we will be interested in solving the Dirichlet problem: on a domain U finding a function u satisfying $\Delta u = 0$ in U with $u = f$ for some function f on its boundary. The characteristics of the data prescribed considerably affects the analysis: for an example that might be familiar from topology for a noncontractible domain U in $\mathbb{R}^2$ and a vector field V one cannot generally find a function u whose gradient is V – this is a system of first order PDE. Theres also a strong analogy between PDE (I suppose especially linear PDE) and linear algebra that is good to have in the back of one's head – with this in mind it can be possible to overdetermine/specify too much boundary data akin to trying to solve too many equations in too few variables, and a problem can also be underdetermined which would typically result in nonuniqueness of solution.

- Speaking of systems, the actual PDE one may wish to consider migh not merely be an equation in one function (i.e. scalar equations) as the models are, but instead might be a systems of equations. There are many such PDE that are very important, such as the Navier–Stokes equation or Maxwell's equations. They tend to be harder to study and have different properties from their one dimensional counterparts when such an analogy can be drawn. For instance, in the mean curvature flow two compact flows of hypersurfaces which are initially disjoint stay so under the flow, but its easy to see this is not the case for say curves in $\mathbb{R}^3$ where the mean curvature flow is more strongly given by a system of PDE. In this course we will be mainly interested in scalar PDE: one equation with one scalar valued function to solve for. These can still be applicable in the study of systems of PDE, because some quantities related to the original system may satisfy a scalar PDE.

- An equation one is interested in understanding might involve more than 2 derivatives, or in other words might have order higher than 2. A noncontrived example of such a PDE is the well known Kortweg–deVries eqaution, $u_t + uu_x + u_{xxx} = 0$ which describes waves in shallow water. It seems though that in practice most PDEs people tend to care about are second order (or less) though – take a look at the long list of PDEs in chapter 1 of Evans [5]. If you look, also note that most of them are variants of either the Laplace, heat, or wave equation. As a handwavy justification for why one might expect 2nd order equations to appear often, we remember from physics Newton's law $F = ma$, where F is the force and the acceleration a is the second derivative of position, and in geometry the PDEs involved often dictate the metric or position vector to the curvature, which involves second derivatives of these (in the appropriate context). For some justification for why these models in particular appear often one can see, at least in the constant coefficient case in two dimensions, that one can find a change of variables to write a second order PDE as one of these three equations – not to be taken very seriously of course. Higher order PDE also are often just harder to understand than second order ones – of course one naively expects complexity to increase in order and there is some truth in this. A concrete reason for this is there is generally a lack of the maximum principle (which we'll learn about soon) for higher order PDE.

- They are also oftentimes nonlinear, in that linear combinations of solutions might not give new solutions which complicates things. Take the minimal surface equation written above for instance. However, a sufficiently good understanding of linear PDE can sometimes be used to tackle related nonlinear equations when one has sufficiently good apriori estimates on the solutions (i.e. bounds on solutions only depending on intial data and terms in the PDE). As a very quick sketch of one well known route which should remind you of the big ODE theorem, solutions of a PDE can be written sometimes as the fixed point of an opertor T on an appropriately defined function space, where T is defined as the solution to a related linear PDE. Showing T has a fixed point involves apriori estimates. The introduction of Gilbarg and Trudinger [6] elaborates on this concerning elliptic PDE, which are those related to the Laplace equation above. Nonlinear PDE often have properties mirroring their linear models, but there can be some new features; for instance in the Ricci and mean curvature flows there is a very useful phenomena called pseudolocality, which says roughly that the initial data nearby a point (i.e.

local) more strongly controls the flow than what one would expect from the heat equation.

The takeaway from above is that the study of the three model equations and generalizations thereof will cover a whole lot of phenomena people care about. Now that we know what we care about, what do we care about? Two main types of questions one can ask about a PDE (artificially delineated) are:

- Is a PDE solvable in a certain space of functions? Note this space of functions apriori might not be, say, the space of smooth functions but instead a space of much less regular ones for the tradeoff that they have better topological properties. Then oftentimes with a "weak" solution we'll be able to prove it is more regular. Is the solution unique given fixed initial data? And how continuously (if at all) do solutions to that PDE behave on initial conditions? A PDE with these qualities is said to be well posed.

- What can we say about solutions to a given PDE? For instance, if we can't present it explicitly can we at least say what it looks like qualitatively? Can we prove estimates (i.e. bounds) on solutions of a PDE in some norm without explicitly finding its solution?

Relatedly we ask the following: *How is this course different from an engineering course in PDE?* In the US at least many such courses are heavily focused on finding fairly explicit solutions to PDE by , say, seperation of variables or Fourier transform. These methods can be very useful and are not obsolete at all in the modern study of PDE, especially Fourier analysis as we'll touch on shortly, but the point is that oftentimes one cannot explicitly find a solution to the equations one might wish to study unless the equation is very simple or there is a lot of symmtery at play. In some cases the solution to a PDE we find might not even be "classical" in that it will only be a solution in a certain weak sense. Knowing simply whether a solution to a PDE exists or not can have signifigance on its own though: if a model of a physical situation is valid, it should be solvable sometimes! And even partial information about a solution can be useful and is better than nothing at all.

3. ODE vs. PDE: life is partially harder

Continuing to set the stage for the bulk of the course, we start off with recalling the "big" ODE existence theorem, which says that under very general conditions a solution to an ODE exists – the point of this section is to then give some PDE theorems of this same flavor along with a counterexample to give some further justification for why we will be restricting our attention to just some compartively subclasses of PDE. Specically, consider the problem of solving the ODE given by:

$$\frac{dy}{dt}(t) = F(t, y),\ y(t_0) = y_0 \tag{3.1}$$

Where y is a vector valued function – this represents a system of 1st order ODE but of course any system of ODE can be reduced to such a system. Then the big ODE theorem is the following:

Theorem 3.1. *Let $y_0 \in U$, an open subset of $\mathbb{R}^n$, $I \subset \mathbb{R}$ an interval containing t_0. Suppose F is continuous on $I \times U$ and is Lipschitz in y:*

$$||F(t, y_1) - F(t, y_2)|| \leq L||y_1 - y_2|| \tag{3.2}$$

for $t \in I$, $y_i \in U$. Then the ODE above has a unique solution defined on some subinterval $J \subset I$ containing t_0.

Proof. (Just a sketch to remind ourselves.) Notice by the fundamental theorem of calculus that a solution to the problem 3.1 is equivalent to finding $y(t)$ such that

$$y(t) = y_0 + \int_{t_0}^{t} F(s, y(s))ds \tag{3.3}$$

With this in mind, let define the (nonempty) space of functions X by:

$$X = \{u \in C(J, \mathbb{R}^n) \mid u(t_0) = y_0,\ \sup_{t \in J} ||u(t) - y_0|| \leq K\} \tag{3.4}$$

Where J is a subinterval of I containing t_0. Then if J and K are picked appropriately depending on the lipschitz constant L above, the operator $T : X \to C(\mathbb{R}, \mathbb{R}^n)$ given by

$$T(f) = y_0 + \int_{t_0}^{t} F(s, y(s))ds \tag{3.5}$$

is actually into the space X, and furthermore the bounds can be arranged so that $d(Tf, Tg) < cd(f, g)$ for a constant $c < 1$ (here d is the metric from the sup norm). Then the contraction mapping principle implies the existence of a fixed point of T, which gives a solution to 3.3 for t sufficiently near t_0. □

Fixed point arguments similar to this one are certainly used in finding solutions to PDE but as we've hinted at already the theory of PDE doesn't have quite as strong a result as the one above – its fun to think about where there are difficulties in applying the method above for a general PDE (for instance, how do you decide what T should be?). The next result is a theorem for PDE with constant coefficients that comes pretty close in spirit to the above though:

Theorem 3.2. *(Malgrange–Ehrenpreis) Every non-zero linear differential operator with constant coefficients has a Green's function.*

Proof: This will just be a sketch – going into this rigorously will take us to far afield (although it not terribly hard) but we'll give an outline since it has some nice ideas and foreshadows some of what we do in the sequel. First we have to unpack the terminology. A *linear differential operator of order* m is a map P from, say, $C^k(\mathbb{R}^n) \to C^{k-m}(\mathbb{R}^n)$ which can be written as

$$P = \sum_{|\alpha|\leq m} a_\alpha(x) D^\alpha \tag{3.6}$$

where $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ is a multindex of nonneagtive integers, $|\alpha| = \alpha_1 + \cdots + \alpha_n$, and $D^\alpha = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \cdots x_n^{\alpha_n}}$. P applied to a function f is then given in the obvious way. Constant coefficients of course means the functions $a_\alpha(x)$ are just constants.

Now, onto what we mean by Green's function. For a linear differential operator P, we say that $G(x, x')$ is a Green's function for it if $PG(x, x') = \delta(x - x')$, where δ is the Dirac delta. Actually what would be more accurate to say is that G is a distribution, which is a continuous linear functional functional on the space $C_c^\infty(\mathbb{R}^n)$ i.e. the compactly supported smooth functions (the topology on this space is actually a little hard to describe). Note its easy to use functions to create distributions via integration but not every distribution arises this way. Then a solution to the PDE $Pu = f$ is given by the covolution $G * f$: if f and g are two integrable functions on $\mathbb{R}^n$, then their convolution $f * g$ is given by:

$$f * g(y) = \int_{\mathbb{R}^n} f(x) g(x - y) dx \tag{3.7}$$

Similarly one can define convolution of a smooth function f with a distribution G by setting $G * f(y) = G(x, y)(f(x))$ and this happens to be a smooth function as well. Convolution has lots of useful properties, one of which is that $P(f * g) = (Pf) * g = f * (Pg)$. So, $P(G * f) = (PG) * f = f$ and we've found a solution to the PDE $Pu = f$ which one may call the Poisson equation for P – its sometimes fruitful to think of a linear PDO as a matrix on the infinite dimensional space of smooth functions, and finding G is akin to finding a matrix inverse.

The question then is how can one find a Green's function. Note that for a general partial differential operator finding a Green's function might not be possible, for instance for $f\Delta$ where f is compactly supported – its impossible to solve the Poisson equation $f\Delta = g$ if g is a function with support different from f. An interesting idea to deal with this, which you may have seen in an earlier PDE course, is to apply the Fourier transform: the Fourier transform $\mathcal{F}g$ (also denoted $\widehat{g}$) of a function g is given by:

$$\mathcal{F}g(\xi) = \frac{1}{\sqrt{2\pi n}} \int_{\mathbb{R}^n} f(x) e^{-ix\cdot\xi} dx \tag{3.8}$$

Here we will suppose we are considering compactly supported smooth functions, although in the context of the Fourier transform its better really to talk about elements of Schwartz space which are functions that rapidly decay. A key property of Fourier transform for our purposes is that using integration by parts (compact support used here) differentiation is transformed into multiplication by $-i\xi$, and in particular $\widehat{Pg} = P(-i\xi)\widehat{g}(\xi)$, where $P(D)$ is our constant coefficient linear partial differential operator and $P(-i\xi)$ is a (complex) polynomial in ξ. $P(-i\xi)\widehat{g}(\xi)$ as a function if any only if $\widehat{g}(\xi) = 0$, which is zero if and only if g is zero. So the map $g \to Pg$ is injective on the space of compactly supported smooth functions and so we stand a chance of inverting it to get a Green's function; a natural guess we see to define G applied to a function f might be the inverse Fourier transform of $\frac{1}{P(-i\xi)} f$, from which one would hope to be able to read off G.

This has issues though because the complex polynomial $P(-i\xi)$ may have zeroes for instance, which affects whether the inverse Fourier transform is applicable. One way to deal with this is a clever partition of unity argument by Hormander in his PhD thesis using the so–called Hormander staircase (which is fairly concrete, and there are other concrete arguments in the modern literature). Another way, in fact the original way, one can proceed is show that the inverse of P on the image of $P(D)$ in $C_c^\infty(\mathbb{R}^n)$ exists and is continuous by some Cauchy–like estimates, by which we roughly mean estimates on the value of an entire function multiplied by some polynomial at a point in terms of an integral. The Banach–Hahn theorem can then be used to extend the defintion of this inverse to prove the existence of a Green's function(/distribution).

□

We've seen when the coefficients are allowed to be nonconstant functions a Green's function does not necessarily exist so the method above stands no chance to be extended to the most general case, and there are a great deal of PDE people care about which don't fit the framework of the result above. Still, inspired by the Fourier transform above, given a partial differential operator (now with possibly nonconstant coefficients) people often consider the principal symbol $\sigma(P)(\xi) = \sum_{|\alpha|=m} a_\alpha(x)(i\xi)$ of a differential operator P. One then says P is elliptic if $\sigma(P)(\xi)$ is nonzero for any nonzero choice of ξ; these have very good properties owing in large part to the fact that one can show there exists a so-called parametrix for them, which is "almost" a Green's function.

When the partial differential operator has analytic coefficients there is another quite general theorem due to Cauchy and Kovalevskaya. Its a bit long to write down here in full generality but its good to know about its existence if even vaguely as a "known unknown" – the method of proof in a nutshell is to match terms in taylor expansions. See chapter 4 of Evans – that chapter has lots of neat tricks, by the way. Something that was really shocking, at least apparently at the time, was the following example due to Hans Lewy in 1956 (take a look at the original paper [14] – its not that long!):

Theorem 3.3. *There exists a smooth complex valued function F on $\mathbb{R} \times \mathbb{C}$ so that the differential equation*

$$\frac{\partial u}{\partial \overline{z}} - iz\frac{\partial u}{\partial t} = F(t, z) \tag{3.9}$$

has no solution on any open set.

If F were smooth, then Cauchy–Kovalevskaya would in fact apply to give a solution so that F being merely smooth matters. The idea is that solutions to the PDE 3.9 must be analytic no matter what the RHS is, and that this implies in turn the RHS must be analytic. Notice that here there is no solution on any open set – the topology (a global sort of input) of the domain isn't being used like in the example we gave above. In fact, the set of smooth F which can be used happens to be dense in a natural sense. The LHS on the other hand is linear, with very nonthreatening coefficients, so one would naively expect would be solvable (and hence the suprise). This example shows that a general result, one of the same sort of strength as theorem 3.1, is unreasonable to expect and so we must focus on more specialized classes of PDE.

4. A detour into first order PDE

The course will focus on second order PDE, but because one is less than two and there are some important applications lets say something about first order PDE before we do that. Intuitively, first order PDE should be relatively simple because they involve the least number of derivatives (not to say these are words to live by) and indeed there are some pretty good theorems involving them. One that is particularly important in differential geometry is the following:

Theorem 4.1. *(Frobenius) Let $X_1, \dots X_k$ be k smooth vector fields in $\mathbb{R}^n$. Then if they are linearly independent at every point and the collection is involutive i.e. $[X_i, X_j] = X_iX_j - X_jX_i$ is in the span of $X_1, \dots X_k$ then at every point p there is a integral submanifold Σ passing through it, or a manifold for which $X_1, \dots X_k$ form a basis for the tangent space of Σ.*

Proof: Here we are thinking of vector fields mainly as derivations, corresponding to directional derivatives where the direction is the vector geometrically speaking. One can check the Lie bracket of two vector fields this way is another vector field and that it satisfies a number of good properties, like bilinearity and the Jacobi rule; see [13] for more details about Lie brackets and Lie flows. This proof is borrowed from chapter 1 of Taylor's book [21] (many books on differential geometry will also have a proof). Breaking things down into more simple language and recalling some defintions, an (embedded) submanifold Σ is locally parameterized by/is the image of a smooth function $F : U \subset \mathbb{R}^k \to V \subset \mathbb{R}^n$ with $J = DF$ nonsingular and so that F^{-1} exists and is continuous with respect to the subspace topology. Then what we want to find is such a function that the span of the columns of DF is the same as the span of X_i – a system of first order PDE. If we parameterize $\mathbb{R}^k$ by $(t_1, \cdots t_k)$ then $(F, t_1, \dots t_k)$ are a local coordinate system of Σ in V with corresponding coordinate vector fields $\frac{\partial}{\partial t_i} = DFe_i$ spanning the tangent space of Σ.

Now, the ODE theorem says that for each of the X_i and a point q we can find an integral curve $\mathcal{F}^t_{X_i}(q)$ of X_i going through q; that is, a curve $\gamma(t) = \mathcal{F}^t_{X_i}(q)$ such that $\gamma'(t) = X_i(\gamma(t))$. A natural thing to try to do define such a map $\mathbb{R}^k \to \mathbb{R}^n$ then is to apply the ODE map iteratively: if in our coordinates the origin is mapped to p then we can try to define a map F by:

$$F(t_1, \dots t_k) = \mathcal{F}^{t_1}_{X_1} \circ \cdots \circ \mathcal{F}^{t_k}_{X_k}(p) \tag{4.1}$$

If the Lie brackets $[X_i, X_j]$ aren't all zero/the vector fields don't all commute, which they do in the case if they are already tangent vectors corresponding to coordinate vector fields, it turns out it isn't clear that the vectors DFe_i are in the span of $X_1, \dots X_k$ and actually this shouldn't always be the case. To see this, consider the vectors $X = \frac{\partial}{\partial x} + y\frac{\partial}{\partial z}$, $Y = \frac{\partial}{\partial y}$ defined on $\mathbb{R}^3$. If they were the tangent vectors to a surface in a neighborhood of the origin which passes through it, then by flowing along X from the origin it contains the x axis, and flowing along Y from these points along the x axis we see that the surface would be a portion of the xy–plane. Starting at a point with $y \neq 0$ though and flowing along X from there gives a contradiction though, because of the $y\frac{\partial}{\partial z}$ term. The idea will be to reduce to the case where the vector fields all commute using the involutive assumption crucially.

With the X_i as in the statement we proceed by induction: the $k = 1$ case follows by the ODE theorem. Suppsoing the statement is true for $k-1$ such vector fields, $k \geq 2$, choose a local coordinate system $(v_1, \dots, v_n)$ for $\mathbb{R}^n$ so that $X_k = \frac{\partial}{\partial v_1}$. This follows from the general theorem of existence of slice charts for an embedded submanifold. Now let

$$Y_j = X_j - (X_j u_1)\frac{\partial}{\partial u_1} \text{ for } j < k \text{ and } Y_k = X_k \tag{4.2}$$

Then in the v_i coordinates none of the $Y_1, \dots Y_{k-1}$ involve $\frac{\partial}{\partial v_1}$ so that tbey are an involutive set and we may apply the induction hypothesis to find functions $y_1, \dots y_{k-1} : \mathbb{R}^{k-1} \to \mathbb{R}^n$ such that the span of $\frac{\partial}{\partial y_i}$ is the same as the span of the Y_i. Using slice charts again, these can be extended to coordinates $y_1, \cdots y_n$. Now, define the vector field Z as:

$$Z = Y_k - \sum_{l=1}^{k-1} (Y_k y_\ell)\frac{\partial}{\partial y_\ell} = \sum_{l > k-1} (Y_k y_\ell)\frac{\partial}{\partial y_\ell} \tag{4.3}$$

We wish to show that $[Z, \frac{\partial}{\partial y_j}] = 0$ for $j < k$. First we check it is in the span of $Y_1, \dots Y_{k-1}$. By linearity and that $[-\sum_{l=1}^{k-1}(Y_k y_\ell)\frac{\partial}{\partial y_\ell}, \frac{\partial}{\partial y_j}]$ is clearly in this span we see it suffices to show this for $[Y_k, Y_j]$. Using again that in the v_i coordinates none of the $Y_1, \dots Y_{k-1}$ involve $\frac{\partial}{\partial v_1}$ and that $Y_k = \frac{\partial}{\partial v_1}$ this bracket is in the span claimed. On the other hand from the second equality of 4.3 $[Z, \frac{\partial}{\partial y_j}]$ is in the span of $\frac{\partial}{\partial y_k}, \dots, \frac{\partial}{\partial y_n}$. Since the intersection of these two spaces is the zero vector, $[Z, \frac{\partial}{\partial y_j}] = 0$.

This implies that $\frac{\partial}{\partial y_1}, \dots, \frac{\partial}{\partial y_{k-1}}, Z$ are a commuting set of vector fields, so as indicated above we can find an integral submanifold for them. Because these are combinations of the original X_k, we get the statement. $\square$

This is used, amongst other interesting applications, in proving the correspondence between Lie sublagebras and Lie subgroups of a Lie group so is good to know about even for the more algebraically minded. There is another powerful method for solving for first order ODE called the **method of characteristics** – see chapter 3 of Evans. We'll develop some of it (far) below when we get to the wave equation in deriving d'Alembert's formula. The idea in a nutshell is that a solution can often be given by the union of solutions (the characteristics) to related ODEs which can be solved often explicitly, giving a staisfactory representation of a solution to the PDE we were original interested in. Second order PDE are the next PDE after first order ones, at least ordering by order, and as mentioned seem to be the most relevent in applications.

5. The fundamental solution/Green's function for Laplace equation on $\mathbb{R}^n$

We now turn to the Laplace equation; of the model equations above it is the one that will be focused on probably most in the course. Now, there are a handful of obvious solutions to the Laplacian, such as the constant and linear functions, but these are pretty cheap because they work by having all their second derivatives equal to zero. In this section we will produce a less trivial solution which will turn out to actually be the Green's function for the Laplacian (aka Δ) on $\mathbb{R}^n$; as an aside about terminology a Green's function for a differential operator on $\mathbb{R}^n$ is also often called a fundamental solution.

The main point of the solution we will find for later developments is that it is a Green's function, and actually it can be found using Fourier transform methods as indicated in theorem 3.2 above, with no extra complications. This is clearly a more principled approach to find the Green's function, but instead following Evans we'll find a solution using a good guess that happliy turns out to give it. Our starting observation is that the Laplace equation is very symmetric, and so its sensible to try to find a rotationally symmetric solution – that is a solution $u(x) = v(r)$, where $r = \sqrt{x_1^2 + \cdots x_n^2}$ is the distance to the origin. Using such an ansatz (educated guess) is helpful because we reduce the number of variables involved, in this case hopefully reducing a PDE to an ODE if all goes well. First we do some calculations when $r \neq 0$.

$$\frac{\partial r}{\partial x_i} = \frac{2x_i}{2\sqrt{x_1^2 + \cdots x_n^2}} = \frac{x_i}{r} \tag{5.1}$$

This gives by the chain rule that:

$$u_{x_i} = v'(r) \cdot \frac{x_i}{r} \tag{5.2}$$

$$u_{x_i x_i} = v''(r) \cdot \frac{x_i^2}{r^2} + v'(r)\frac{r - x_i \frac{x_i}{r}}{r^2} = v''(r) \cdot \frac{x_i^2}{r^2} + \frac{v'(r)}{r}(1 - \frac{x_i^2}{r^2}) \tag{5.3}$$

Now, using that $\Delta u = \sum_{i=1}^{n} u_{x_i x_i}$ and that $\sum_{i=1}^{n} x_i^2 = r^2$ we have

$$\Delta u(x) = \sum_{i=1}^{n} u_{x_i x_i} = v''(r) + v'(r) \cdot \frac{n-1}{r} \tag{5.4}$$

So we get an ODE! Note that if instead of the laplace equation some of the coefficients on the $u_{x_i x_i}$ terms were different from others we wouldn't have gotten such a clean formula only involving r (with none of the x_i appearing explicitly). This ODE can be solved (exercise!) to find that

$$\Phi(x) = \begin{cases} -\frac{1}{2\pi}\log(|x|), n = 2 \\ \frac{1}{n(n-2)\alpha(n)}|x|^{2-n}, n \geq 3 \end{cases} \tag{5.5}$$

is a smooth solution to the Laplace equation on $\mathbb{R}^n \setminus \{0\}$, where $\alpha(n)$ is the volume of the unit n-ball; there are more solutions of the same form but these are picked so integrals involving them work out nicely below. Now, we want to claim soon that its the Green's function of the Laplacian on $\mathbb{R}^n$ – this is perhaps reasonable to hope for because Φ is harmonic away the origin and blows up (so looks sort of like the Dirac delta at least for $n \geq 3$) as one approaches it. To justify commuting some limits and integrals first we'll want to know a bit more about Φ, namely the following:

Lemma 5.1. $\Phi \in L^1_{loc}(\mathbb{R}^n)$

Proof: What this claim is saying is that for any point $p \in \mathbb{R}^n$, there is some neighborhood U of p for which $\int_U |\Phi| dx$ exists and is bounded. Φ is clearly a measurable function, and smooth away from 0, so we really only need to show that the integral of it over a ball is bounded. We estimate

$$\int_{B(0,a)} \Phi dx = \int_0^a \int_{S(0,r)} \Phi dS dr \leq \begin{cases} -C\int_0^a \log(r) r dr \\ C\int_0^a r^{2-n} r^{n-1} dr \end{cases} \tag{5.6}$$

For a dimensional constant C. We see that the integrands in both are uniformly bounded (and tend to zero as a does) giving the claim. $\square$

However, by the same sort of argument, note that Φ is not in $L^1(\mathbb{R}^n)$ because it decays too slow at inifinity. Anyway our next claim is that Φ gives a Greens function (specifically by setting $G(x,y) = \Phi(x-y)$.

Theorem 5.2. *Let $f \in C_c^2(\mathbb{R}^n)$. Then $u(x) = \int_{\mathbb{R}^n} \Phi(x-y)f(y)dy = \int_{\mathbb{R}^n} \Phi(y)f(x-y)dy$ is in $C^2(\mathbb{R}^n)$ and satisfies $-\Delta u = f$.*

Proof: Note that the first equality above follows just by change of variables and is helpful because f is smooth with compact support – we naturally want to calculate Δu to see what we get. Now by the lemma above $|\Phi(y)f_{x_i}(x-y)|$, $|\Phi(y)f_{x_ix_i}(x-y)|$ are both uniformly bounded in L^1 so, by the dominated convergence theorem, we have $\Delta u(x) = \int_{\mathbb{R}^n} \Phi(y)\Delta_x f(x-y)dy$. Because of Φ's bad behavior at the origin we further split this up as

$$\Delta u(x) = \int_{B(0,\epsilon)} \Phi(y)\Delta_x f(x-y)dy + \int_{\mathbb{R}^n \setminus B(0,\epsilon)} \Phi(y)\Delta_x f(x-y)dy = I_1(\epsilon) + I_2(\epsilon) \quad (5.7)$$

Now the first term, $I_1(\epsilon)$, tends to zero as ϵ does as we saw in the proof of the lemma above. This is advantagous for us because on the set $\mathbb{R}^n \setminus B(0,\epsilon)$ Φ is smooth and in fact harmonic – in the following we will essentially use integration by parts twice (Green's formula) to move the Laplacian back onto Φ; note that since $(-1)^2 = 1$ that $\Delta_x f(x-y) = \Delta_y f(x-y)$. With this in mind for the second term we use Green's formula, that $\int_\Omega u\Delta v = -\int_\Omega Du \cdot Dv + \int_{\partial\Omega} u\frac{\partial v}{\partial \nu}$ to write:

$$I_2(\epsilon) = -\int_{\mathbb{R}^n \setminus B(0,\epsilon)} D_y\Phi(y) \cdot D_y f(x-y)dy + \int_{S(0,\epsilon)} \Phi(y)\frac{df}{d\nu}(x-y)dS(y) = I_3(\epsilon) + I_4(\epsilon) \quad (5.8)$$

Using the same reasoning in the lemma again, $I_4(\epsilon)$ tends to zero as ϵ does so we are left with considering $I_3(\epsilon)$. Integrating by parts/using Green's formula again gives:

$$I_3(\epsilon) = \int_{\mathbb{R}^n \setminus B(0,\epsilon)} \Delta\Phi(y)f(x-y)dy - \int_{S(0,\epsilon)} \frac{d\Phi}{d\nu}(y)f(x-y)dS(y) = -\int_{S(0,\epsilon)} \frac{d\Phi}{d\nu}(y)f(x-y)dS(y) \quad (5.9)$$

Where the second equality is because Φ is harmonic away from the origin. Now we need to calculate $\frac{d\Phi}{d\nu}(y) = \nu \cdot D\Phi(y)$ where ν to be clear is the unit inward normal. Noting that on the sphere $S(0\epsilon)$ the unit normal is given by $\nu = -y/|y| = -y/\epsilon$ and also on it $y \cdot y = \epsilon^2$ we have the following (valid on the sphere)

$$\frac{d\Phi}{d\nu}(z) = \begin{cases} -y/\epsilon \cdot -\frac{1}{2\pi}(\frac{1}{\epsilon} \cdot \frac{y}{\epsilon}) = \frac{1}{2\pi\epsilon} \\ -y/\epsilon \cdot \frac{1}{n(n-2)\alpha(n)}(\frac{2-n}{\epsilon^{n-1}} \cdot \frac{y}{\epsilon}) = \frac{1}{n\alpha(n)\epsilon^{n-1}} \end{cases} \quad (5.10)$$

These are exactly the areas of the sphere of radius ϵ for the 2 and $n \geq 3$–sphere of radius ϵ respectively. This implies that

$$-\int_{S(0,\epsilon)} \frac{d\Phi}{d\nu}(y) f(x-y) dS(y) = -\fint_{S(0,\epsilon)} f(x-y) dS(y) \tag{5.11}$$

where the right hand side is the average of $-f$ over the sphere of radius ϵ centered at the point x. Because f is continuous, the value of this as $\epsilon \to 0$ is $-f(x)$, giving the claim. □

It can be useful to that the Green's function encodes a lot of useful information about solutions to a PDE and the geometry of the underlying space it is set on (if one considers PDE on a curved manifold), and so the study of Green's functions for various operators is an important topic in its own right.

6. Mean and maximum principles for harmonic functions

Above we just solved the Poisson equation for the Laplacian, at least when the RHS is in C_c^∞ and the domain is $\mathbb{R}^n$. There's more one could ask for of course, for instance what about on a smooth domain with data prescribed along the boundary? We'll leave it be for the immediate future and just be content with the fact for now that there are solutions to the Laplace and Poisson equations out there to prove things about, and instead go ahead to showing the most primordial of all PDE properties, the mean value principle for harmonic functions:

Theorem 6.1. *Let $U \subset \mathbb{R}^n$ be an open domain and suppose $u \in C^2(U)$ is harmonic i.e. $\nabla u = 0$. Supposing the ball $B(x,r) \subset U$ we have:*

$$u(x) = \fint_{S(x,r)} u dS = \fint_{B(x,r)} u dy \tag{6.1}$$

Proof: We start with showing the first equality. Let $\phi(r) = \fint_{S(x,r)} u(y) dS(y)$. We want to show that $\phi(r)$ is constant/$\phi' = 0$, which will give that $\phi(r) = \lim_{r \to 0} \phi(r) = u(x)$ like in the proof above. To calculate the derivative of ϕ first we perform a change of variables to get rid of the r dependence in the domain of integration, simplifying matters. We consider a new variable $z = \frac{y-x}{r}$, so that $S(x,r)$ is sent to $S(0,1)$. Recalling the change of variables formula the jacobian of this transformation is simply $\frac{1}{r^n}$, which is absorbed by the scaling constant when we consider the averaged integral. This gives that $\fint_{S(x,r)} u(y) dS(y) = \fint_{S(0,1)} u(x+rz) dS(z)$. Then we calculate (clearly passing the derivative through is no problem):

$$\phi'(r) = \fint_{S(0,1)} Du(x + rz) \cdot z dS(z) \tag{6.2}$$

With this calculation done we just change variables back, and then use Green's formula (with the other function being 1, so that its derivative vanishes) to write:

$$\phi'(r) = \fint_{S(0,r)} Du(y) \cdot \frac{y - x}{r} dS(y) = \fint_{S(0,r)} \frac{\partial u}{\partial \nu} dS(y) = \frac{r}{n} \fint_{B(0,r)} \Delta u dy = 0 \tag{6.3}$$

As explained this gives us our first equality; the factor on the average integral over the ball is there because the previous ones were over spheres. For the second equality we use that the integral over the ball can be written as an interated integral over spheres along with what we just showed:

$$\int_{B(x,r)} u dy = \int_0^r (\int_{S(x,s)} u dS) ds = u(x) \int_0^r n\alpha(n) s^{n-1} ds = \alpha(n) r^n u(x) \tag{6.4}$$

Dividing through by $\alpha(n)r^n$ gives us the second equality. □

If a $C^2(U)$ function u isn't harmonic then $\Delta u \neq 0$ at some point $p \in U$ and so, for a very small ball B about p, has a sign. Inspecting the above proof then we see:

Theorem 6.2. *If $u \in C^2(U)$ satisfies*

$$u(x) = \fint_{S(x,r)} u dS = \fint_{B(x,r)} u dy \tag{6.5}$$

then it is harmonic.

With the mean value theorem we can prove the (strong) maximum principle. Maximum principles in various guises are truly some of the most important tools used in geometric analysis so its good to remember this one. I'll give one proof, using the mean value property, and then I'll give another which works more generally, is easy, and is a bit more how I think about things (see Gilbarg and Trudinger). The first statement below is usually just called the "maximum principle" (I tend to call it the regular maximum principle, while others call it the weak max principle – but its pretty mighty!). The second one, claiming rigidity, is the strong version.

Theorem 6.3. *Suppose $u \in C^2(U) \cap C(\overline{U})$ is harmonic within U.*

(1) Then

$$\max_{\overline{U}} u = \max_{\partial U} u$$

(2) Furthermore, if U is connected and there exists a point $x_0 \in U$ such that

$$u(x_0) = \max_{\overline{U}} u$$

then u is constant within U.

Proof: First the proof given in Evans, using the mean value principle: suppose there is a point $x_0 \in U$ where the maximum M of u is achieved. Since U is open the distance $d(x_0, \partial U)$ between x_0 and ∂U is positive; from the mean value property we have for $0 < r < d(x_0, \partial U)$ that

$$M = u(x_0) = \fint_{B(x_0,r)} u dy \leq M \tag{6.6}$$

From the defintion of M equality holds only if $u = M$ identically within $B(x_0, r)$. This gives that the set $\{x \in U \mid u(x) = M\}$ is open. On the other hand by the continuity of u it is closed. Hence, if u achieves its maximum M in U and U is connected it is equal to M everywhere in U, giving item (2). Item (2) implies item (1) because the max of u on $\overline{U}$ is achieved apriori in U or ∂U, and in the former case from (2) it will also be achieved on the boundary since its just a constant function (on that connected component).

Now let's give a second proof of (1) which broadly works for general linear elliptic operators as well (we should get around to introducing these fairly shortly but if you are curious take a peak at chapter 3 of [6]). First note that if v is a function so that $\Delta v > 0$ (such functions are examples of so–called subharmonic ones), then the maximum of v on $\overline{U}$ must be attained on ∂U because, by the second derivative test, $\Delta v \leq 0$ at points in U where its maximum is achieved. With this in mind consider $v = u + \epsilon e^{cx_1}$, for some $c > 0$. Then:

$$\Delta v = \Delta u + \Delta \epsilon e^{cx_1} = \epsilon c^2 e^{cx_1} > 0 \tag{6.7}$$

By what we said, the maximum of v must be attained on ∂U. Taking $\epsilon \to 0$ shows that it is true for u too, giving the claim. □

To see why the second proof is more general, note that it also works if we add, say, $b(x)\frac{\partial}{\partial x_i}$ to the laplacian where b is some bounded function because $\nabla v = 0$ at critical points.

7. Some first consequences of the mean and maximum principles

As a first consequence, we get uniqueness for boundary value problems to Poisson's equation (we'll return to the topic of solving this – the precription of boundary data is new – later):

Theorem 7.1. *Let $g \in C(\partial U)$, $f \in C(U)$. Then there exists at most one solution $u \in C^2(U) \cap C(\overline{U})$ of the boundary value problem*

$$\begin{cases} -\Delta u = f \ \text{in } U \\ u = g \ \text{on } \partial U \end{cases} \tag{7.1}$$

Proof: Suppose u_1 and u_2 are two such solutions. Then their difference $u_1 - u_2$ is a solution to the Dirichlet problem with boundary data equal to zero. The maximum prinicple then says that $u_1 - u_2$ is nonpositive. Repeating the same argument with $u_2 - u_1 = -(u_1 - u_2)$ we see it must also be nonnegative, giving that it is zero.

□

Uniqueness is pretty common in general for elliptic boundary value problems, but not a hard and fast rule. For a geometric example if one considers two round circles laying in parallel planes in $\mathbb{R}^3$, then if these planes are close enough there are (at least) two minimal surfaces spanning them: one which looks like a catenoid bridging the two loops and another simply given by two parallel flat discs with the circles as boundary. Next we discuss some theorems concerning bounds on and the regularity of solutions to the Laplace equation, more or less in increasing strength (depending on perspective). We start with the famous Harnack's inequality:

Theorem 7.2. *For each connected open set $V \subset\subset U$ there exists a positive constant C, depending only on V, such that*

$$\sup_V u \leq C \inf_V u \tag{7.2}$$

for all nonnegative harmonic function u in U.

Proof: Here the double inclusion means that even the closure of V is contained in U, and that the closure is compact. Obviously the inequality can't be true if u switches signs, of course. Fixing V throughout if we consider two points $x, y \in V$ we have $u(y) \leq \sup_V u$ and $u(x) \geq \inf_V u$ so that $\frac{1}{C}u(y) \leq u(x)$; similarly $u(x) \leq Cu(y)$ so a nicely phrased consequence is that $\frac{1}{C}u(y) \leq u(x) \leq Cu(y)$ – this inequality holding for all x and y implies the statement above of course.

Now, let $r = \frac{1}{4}d(V, \partial U)$ and choose $x, y \in V$ with $|x - y| < r$. Then by the mean value property and the use of the nonnegativity of u in the second inequality we have:

$$u(x) = \fint_{B(x,2r)} u dz \geq \frac{1}{\alpha(n) 2^n r^n} \int_{B(y,r)} u dz \ = \frac{1}{2^n} \fint_{B(y,r)} u dz = \frac{1}{2^n} u(y) \tag{7.3}$$

Thus $\frac{1}{2^n}u(y) \leq u(x) \leq 2^n u(y)$ for $x, y \in V$ when $|x-y| < r$. Now if we cover V with balls B_i of radius less than r we can extract a finite subcover, say N of them, using that $\overline{V}$ is compact. Since V is connected for any $x, y \in V$ there is a (Harnack) chain of these balls $B_1, \ldots B_k$ where $x \in B_1$, $y \in B_k$, and $B_i \cap B_{i-1} \neq \emptyset$. We can then apply the inequality in balls successively (compaing using points in the intersections of the balls) and that k must be less than N so that

$$u(x) \leq 2^{n(N+1)}u(y) \tag{7.4}$$

for all $x, y \in V$. The constant depends on V where the bound on the number of balls needed in the cover is used. □

Harnack's inequality holds pretty generally and even for heat like equations – in the Ricci and mean curvature flows there is an important Harnack inequality call Hamilton's harnack inequality which is useful in the singularity analysis of these flows. We'll discuss a harnack inequality for the heat equation later. Next we prove the following, which is also a pretty common property:

Theorem 7.3. *If $u \in C(U)$ satisfies the mean value property for each ball $B(x,r) \subset U$, then $u \in C^\infty(U)$.*

Proof: Remember from above that if u satisfies the mean value property and is twice differentiable, then it must be harmonic; here we are only assuming apriori that it is continuous however. Also as pointed out in Evans note that no claim about the continuity of a possible extension of u to ∂U is made – such questions of regularity up to the boundary often have to be dealt with separately.

Anyway denote by η a standard mollifier, which roughly speaking is a smooth function that looks like a bump concentrated at the origin and is radial i.e. is only a function of $r = |x|$. See appendix C of Evans for more precision. Also denote by η_ϵ to be $\frac{1}{\epsilon^n}\eta(x/\epsilon)$ – note because the support of η lays in a ball of radius 1 the support of η_ϵ lays in a ball of radius ϵ. The idea below is that if we mollify u with η_ϵ in the set $U_\epsilon = \{x \in U \mid d(x, \partial U) > \epsilon\}$ we get a smooth function u^ϵ; it will be smooth because we can pass the derivative through the integral sign onto η. On the other hand we will see by the mean value property that its equal to u, giving us the claim as we let $\epsilon \to 0$. Getting to it:

$$\begin{aligned} u^\epsilon(x) = u * \eta_\epsilon(x) &= \int_{U_\epsilon} \eta_\epsilon(x-y)u(y)dy \\ &= \frac{1}{\epsilon^n}\int_{B(x,\epsilon)} \eta(\frac{|x-y|}{\epsilon})u(y)dy \end{aligned} \tag{7.5}$$

Now we break the integral over the ball up into integrals over spherical shells; since η is radial its constant on each sphere we can pull it out and use the mean value property:

$$\begin{aligned} &= \frac{1}{\epsilon^n}\int_0^\epsilon \eta(\frac{r}{\epsilon})(\int_{S(x,r)} udS)dr \\ &= \frac{1}{\epsilon^n}u(x)\int_0^\epsilon \eta(\frac{r}{\epsilon})n\alpha(n)r^{n-1}dr \\ &= u(x)\int_{B(0,\epsilon)} \eta_\epsilon dy = u(x) \end{aligned} \tag{7.6}$$

In the last equality we are using that the integral of η is normalized to be 1. As explained already this gives the claim. □

One simple property that is used over and over again is taking convergent subsequences of functions – either to solve a PDE by solving it on a set of simpler domains and taking a sequence, or by considering a "contradictory" sequence of solutions to some problem, extracting a subsequece due to some sort of compactness. and using known rigidity results to argue by contradiction. In other words, having results that say a space of solutions to some sort of problem is compact can be very helpful. The next result is called Harnack's convergence theorem and is useful particularly in the former situation just described:

Theorem 7.4. *Let $\{u_n\}$ be a monotone increasing sequence of harmonic functions in a domain U and suppose that for some point $y \in U$ that the sequence $\{u_n(y)\}$ is bounded. Then the sequence converges uniformly on any bounded subdomain $V \subset\subset U$ to a harmonic function.*

Proof: Since the sequence $\{u_n(y)\}$ is bounded and the sequence of functions is monotone it converges, so in particular for any $\epsilon > 0$ there exists a number N so that $0 \le u_m(y) - u_n(y) < \epsilon$ for all $m \ge n > N$. By linearity the difference $u_m - u_n$ is harmonic and by the monotonicity its nonnegative so for a fixed choice of $V \subset\subset U$ Harnack's inequality 7.2 gives a constant C so that

$$\sup_V |u_m(x) - u_n(x)| < C\epsilon \tag{7.7}$$

Note the constant C here depends just on V and so implies that the sequence of functions $\{u_n\}$ is a Cauchy seuqence in the sup norm converges uniformly to some continuous function u in $\overline{V}$. Since each of the u_i are harmonic they satisfy the mean value property so, since the convergence is uniform, one can see u does as well. By the theorem above u must be smooth, so by the converse of the mean value property its harmonic. □

The next statement can be thought of as a sharpening of theorem 7.3, giving explicit bounds/estimates on the derivatives of u at a point in terms of its L_1 norm in a neighborhood of it:

Theorem 7.5. *Assume that u is harmonic in U. Then*

$$|D^{\alpha}u(x_0)| \leq \frac{C_k}{r^{n+k}}||u||_{L^1(B(x_0,r))} \tag{7.8}$$

for each ball $B(x_0,r) \subset U$ and each multiindex α of order $|\alpha| = k$. Here the constants are:

$$C_0 = \frac{1}{\alpha(n)}, C_k = \frac{(2^{n+1}nk)^k}{\alpha(n)} \tag{7.9}$$

Proof: The proof is by induction on k, with the $k = 0$ case following directly from the mean value theorem and that $\int f \leq \int |f|$. For the induction step there are two important observations: if u solves the laplace equation then so does derivatives of u, since derivatives commute, and we can use the divergence theorem to "strip off" derivatives to let us use the inductive hypothesis. Let's see the argument for the $k = 1$ case first (strictly speaking, this isn't necessary). Using the first obeservation and the mean value property:

$$\begin{aligned} |u_{x_i}(x_0)| &= |\fint_{B(x_0,r/2)} u_{x_i}dx| \\ &= |\frac{2^n}{\alpha(n)r^n}\int_{S(x_0,r/2)} u\nu_i dS| \quad \leq \frac{2n}{r}||u||_{L^{\infty}(S(x_0,r/2))} \end{aligned} \tag{7.10}$$

The last line is just a crude estimate of the integral in terms of the max of u on the sphere and isn't using the inductive hypothesis. Now for $x \in S(x_0, r/2)$ we note by the triangle inequality that $B(x, r/2) \subset B(x_0, r)$, so *now* we can use the inductive hypothesis to bound $||u||_{L^{\infty}(S(x_0,r/2))}$ by $\frac{1}{\alpha(n)}||u||_{L^1(B(x_0,r))}$. Combining this with the chain of (in)equalities above gives

$$|D^{\alpha}u(x_0)| \leq \frac{2^{n+1}n}{\alpha(n)}\frac{1}{r^{n+1}}||u||_{L^1(B(x_0,r))} \tag{7.11}$$

When $\alpha = 1$. Note that the first fraction on the RHS does agree with what we called C_1. Now consider a multiindex α with $|\alpha| = k$ and that the estimates are known for all multiindices of length $\leq k-1$. Of course $D^\alpha u$ is harmonic, and we can write $D^\alpha u = (D^\beta u)_{x_i}$ for some i and some multiindex β of length $k-1$. Denoting by $v = D^\beta u$ we have from the $k=1$ work in the $k=1$ case:

$$|D^\alpha u(x_0)| = |v_{x_i}(x_0)| \leq \frac{nk}{r}||v||_{L^\infty(S(x_0,r/k))} \tag{7.12}$$

Using that for $x \in S(x_0, r/k)$, $B(x, \frac{k-1}{k}r) \subset B(x_0, r)$ by the inductive hypothesis we have

$$||v||_{L^\infty(S(x_0,r/k))} \leq \frac{C_{k-1}}{(\frac{k-1}{k}r)^{n+k-1}}||u||_{L^1(B(x_0,r))} \tag{7.13}$$

Combining this with the above, we have

$$|D^\alpha u(x_0)| \leq \frac{nk}{r}\frac{C_{k-1}}{(\frac{k-1}{k}r)^{n+k-1}}||u||_{L^1(B(x_0,r))} = \frac{nk^{n+k}C_{k-1}}{(k-1)^{n+k-1}}\frac{1}{r^{n+k}}||u||_{L^1(B(x_0,r))} \tag{7.14}$$

Writing out C_{k-1}, the factor in front of $1/r^{n+k}$ above is

$$\frac{nk^{n+k}2^{nk+k-n-1}n^{k-1}(k-1)^{k-1}}{\alpha(n)(k-1)^{n+k-1}} = \frac{n^k k^{n+k}2^{nk+k-n-1}}{\alpha(n)(k-1)^n} \tag{7.15}$$

Now, borrowing 2^{-n-1} from the $2^{nk+k-n-1}$ term, we see $\frac{k^{n+k}}{2^{n+1}(k-1)^n} < k^k$ using that $2(2k-2)^n \geq k^n$ when $k \geq 2$. Hence the LHS above is bounded by $\frac{n^k k^k 2^{nk+k}}{\alpha(n)} = C_k$ giving the claim. □

Derivative bounds for solutions to the more general Poisson's equation, $-\Delta u = f$ are also possible at least for $\alpha = 1$ by the maxmimum principle and even more general statements will be discussed (much) later. A nice consequence of these estimates is Liouville's theorem for Harmonic functions; the name and statement should remind you of a similar statement in complex analysis – this isn't a coincidence! We'll discuss their relationship more a short time later:

Theorem 7.6. *Suppose that $u : \mathbb{R}^n \to \mathbb{R}$ is harmonic and bounded. Then u is constant.*

Proof: Fixing $x_0 \in \mathbb{R}^n$ we see that on the ball $B(x_0, r)$, $||u||_{L^1(B(x_0,r))} \leq Cr^n$ where C is some constant in terms of the assumed bound on $|u|$ and dimensional constants. By the $k=1$ derivative bounds from above we have $|Du(x_0)| \leq \frac{C_1 C}{r}$, which tends to zero as $r \to \infty$. Since x_0 was arbitrary we get that $Du = 0$ so that it is constant. □

Using the derivative estimates its easy to see that any bounded sequence of harmonic functions has a subsequence which converges uniformly on compact subdomains to a limit function in the C^k topology for any k using the Arzela–Ascoli theorem, and the limit function is itself harmonic. Comparing to theorem 7.4 this claim is in some ways stronger and weaker than it but the refinement in the topology of convergence we may suppose is cerrtainly a strengthening and the topology under consideration in applying these sort of results does matter often – indeed the derivative estimates can be used to strengthen theorem 7.4. We can also refine theorem 7.3 to see that solutions to the Laplace equation are actually analytic; recall that theorem says that functions which are continuous and have the mean value property are smooth, and so since they are C^2 they are harmonic by the converse to the mean value property so the conclusion below is true for the functions in that statement as well.

Theorem 7.7. *Assume u is harmonic in U. Then u is analytic in U.*

Proof: Recalling the definition of analytic function, we recall that we need to show that the taylor series of u centered at any point x_0 of U, $\sum_{\alpha,|\alpha|=0}^{\infty} \frac{D^\alpha u(x_0)}{\alpha!}(x - x_0)^\alpha$ (recall what these mean for multiindices), converges in some neighborhood of x_0 and actually agrees with u near x_0 as well. Recall that there are smooth but not analytic functions even on $\mathbb{R}$: the canonical example is defined by:

$$f(x) = \begin{cases} 0 \text{ when } x \leq 0 \\ e^{-1/x} \text{ when } x > 0 \end{cases} \tag{7.16}$$

This happens to be a smooth function but its Taylor series at $x = 0$ has all coefficients equal to zero – of course this doesn't agree in any neighborhood of f about zero because $f > 0$ for all $x > 0$. Now, letting $r = \frac{1}{4}d(x_0, \partial U)$ we see that $M = \frac{1}{\alpha(n)r^n}||u||_{L^1(B(x_0,2r))}$ is well defined and finite. Since $B(x, r) \subset B(x_0, 2r) \subset U$ for each $x \in B(x_0, r)$, the derivative estimate above can be used to see

$$||D^\alpha u||_{L^\infty(B(x_0,r))} \leq M(\frac{2^{n+1}n}{r})^{|\alpha|}|\alpha|^{|\alpha|} \tag{7.17}$$

We basically want to plug these estimates into the definition of Taylor series to estimate it, but since there are some terms which could concievably grow fast compared to $\alpha!$ (e.g. $|\alpha|^{|\alpha|}$) we need to do some estimating. Recalling the Taylor expansion (at $x = 0$) of e^x evaluated at k we see that $\frac{k^k}{k!} \leq e^k$ for all positive integers and hence $|\alpha|^{|\alpha|} \leq e^{|\alpha|}|\alpha|!$. As a consequence of the multinomial theorem $n^k = \sum_{|\alpha|=k} \frac{|\alpha|!}{\alpha!}$ so that for a given multiindex of length k $|\alpha|! \leq n^{|\alpha|}\alpha!$. Combining these gives $|\alpha|^{|\alpha|} \leq e^{|\alpha|}|\alpha|! \leq e^{|\alpha|}n^{|\alpha|}\alpha!$; plugging this into the derivative estimates gives:

$$||D^\alpha u||_{L^\infty(B(x_0,r))} \leq CM(\frac{2^{n+1}n^2 e}{r})^{|\alpha|}\alpha! \tag{7.18}$$

One can check that if $|x - x_0|$ is sufficiently small then, using these estimates to majorize the corresponding terms in the Taylor series centered at x_0, the Taylor series converges by standard series comparison/convergence theorems but as discussed that isn't quite good enough; we need to check that if $x - x_0$ is sufficiently small then the error between $u(x)$ and the Taylor series expansion up to multiindices of length $N-1$ of u tends to zero as $N \to \infty$. In particular we claim this is true for $|x-x_0| < \frac{r}{2^{n+2}n^3 e}$. To check this we apply Taylor's theorem with remainder, a consequence of the mean value theorem in calculus, to the 1-d function $g(s) = u(x_0 + s(x - x_0))$ at $s = 1$. We get from the theorem (at the second equality):

$$R_N(x) = u(x) - \sum_{k=0}^{N-1} \sum_{|\alpha|=k} \frac{D^\alpha u(x_0)}{\alpha!}(x - x_0)^\alpha = \sum_{|\alpha|=N} \frac{D^\alpha u(x_0 + t(x - x_0))}{\alpha!}(x - x_0)^\alpha \tag{7.19}$$

Where $0 \leq t \leq 1$. Plugging in our derivative estimates and that bound on $|x - x_0|$ we assumed/claimed worked we have:

$$|R_N(x)| \leq CM \sum_{|\alpha|=N} (\frac{2^{n+1}n^2 e}{r})^N (\frac{r}{2^{n+2}n^3 e})^N \quad \leq CMn^N \frac{1}{(2n)^N} = \frac{CM}{2^N} \to 0 \text{ as } N \to \infty \tag{7.20}$$

□

As a consequence of u being analytic its not so hard to see it satisfies the *unique continuation property*, that if two harmonic functions agree on an open subset of a connected domain they agree everywhere. This holds for more general elliptic PDE even when the result above doesn't hold. A nice survey on this topic using the frequency function, which is relatively elementary, is [8].

8. A diversion about complex analysis

We pause here to compare the situation for harmonic functions to holomorphic ones, which depending on your complex analysis background you might notice (and know why already, but its good to be reminded) have a lot of the same great properties. In complex analysis one considers (complex valued) functions defined on $\mathbb{C} \sim \mathbb{R}^2$. Writing $z = x + iy$ and $\overline{z} = x - iy$ so that $x = \frac{z+\overline{z}}{2}$ and $y = \frac{z-\overline{z}}{2i}$ a function f on $\mathbb{C}$ can be considered as a function of z and $\overline{z}$. Then a function f on an open set $U \subset \mathbb{C}$ is one where

$$\frac{\partial f}{\partial \overline{z}} = 0 \tag{8.1}$$

at every point of U, or so that it depends only on z. Here $\frac{\partial}{\partial \overline{z}} = \frac{1}{2}(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y})$, as you'd basically expect except for the sign change. Writing a function f as $u(x, y) + iv(x, y)$ we find collecting real and imaginary parts (using $i^2 = -1$)

$$\frac{\partial f}{\partial \overline{z}} = (\frac{\partial}{\partial x} + i\frac{\partial}{\partial y})(u + iv) = \frac{1}{2}(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y}) + \frac{i}{2}(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}) \tag{8.2}$$

So, if f is holomorphic we must have that

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \text{ and } \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x} \tag{8.3}$$

These equations are called the Cauchy–Riemann equations, which are a system of coupled first order PDE for u and v. If we suppose that u and v are in $C^2(U)$. then differentiating the first equation with respect to x and the second with respect to y and using that mixed partials commute to go between them we have:

$$\frac{\partial^2 u}{\partial x^2} = \frac{\partial^2 v}{\partial x \partial y} = \frac{\partial^2 v}{\partial y \partial x} = -\frac{\partial^2 u}{\partial y^2} \tag{8.4}$$

The same reasoning applies to v as well. They immediately give:

Theorem 8.1. *Suppose that $f = u + iv$, where u and v are real valued C^2 functions on an open subset U of $\mathbb{C}$, and f is holomorphic on U. Then u and v are harmonic on U.*

Conversely if we have a harmonic function g, then one can see that letting $u = \frac{\partial g}{\partial x}$ and $v = -\frac{\partial g}{\partial y}$ then u and v satisfy the Cauchy–Riemann equations giving that the function $f = u + iv = g_x - ig_y$ is holomorphic. In the case that f has a primitive, which is to say there exists a holomorphic function F for which $\frac{dF}{dz} = g$, then one can see that $F = g + ih$, where h is the so–called harmonic conjugate of g. Such a primitive can always be found on simply connected domains of $\mathbb{C}$, which says that holomorphic functions are in direct correspondence with harmonic ones on them.

A nice consequence of this relationship for us is that complex polynomials (i.e. polynomials in z) happen to be holomorphic, so by taking their real and imaginary parts we find a way to produce many different harmonic functions. Now as mentioned above holomorphic functions $f = u+iv$ are very special, for instance like the harmonic ones they are analytic. This can be interpreted as a byproduct of u and v being harmonic, although this isn't how its usually done in a complex analysis course. Instead the main tool is Cauchy's integral formula, which says that for a holomorphic function f on a domain U then for a point $z \in U$:

$$f(z) = \oint_\gamma \frac{f(\zeta)}{\zeta - z} d\zeta \tag{8.5}$$

Where γ is a closed curve on U that we can just assume here is a small circle around z; the RHS can be thought of as convolution with $\frac{1}{w}$ and it should remind you a little bit of the Green's function. The smoothness of f, for instance, follows by justifying passing derivatives under the integral sign and using basically that $1/w$ is smooth away from the origin. Indeed, the correspondence above (theorem 8.1 and the discussion after) and the Cauchy integral formula give an alternate route to show some of the results in the previous section for harmonic functions on $\mathbb{R}^2$. To proceed along lines closer to this for harmonic functions without reference to complex analysis persay one can show, using Green's formulas (see eq. 2.18 in [6]; we'll also discuss this shortly) that for a harmonic function u on a smooth domain U that:

$$u(x) = \int_{\partial U} (u(y) \frac{\partial \Phi}{\partial \nu}(y - x) - \Phi(y - x) \frac{\partial u}{\partial \nu}(y)) dS(y) \tag{8.6}$$

where here y is in U and in particular not on the boundary. Because $\Phi(x - y)$ for $x \neq y$ is smooth and even analytic, one can see that u is from the representation above; one should also be able to derive derivative estimates.

9. Green's function for the Laplacian on general domains

We've just assembled a nice collection of facts about Harmonic functions and now we turn back to their existence, particularly to solve Poisson's problem eventually in a general (eventually C^2) bounded domain of $\mathbb{R}^n$. Or, in other words we want to eventually say something about solving the problem

$$\begin{cases} -\Delta u = f \text{ in } U \\ u = g \text{ on } \partial U \end{cases} \tag{9.1}$$

Where above f and g are sufficiently regular – from the results in this section and the next we can handle $f \in C^2(U)$ and $g \in C^0(\partial U)$ although this probably isn't sharp. Now sure, if $f \in C_c^\infty(U)$ one could try to proceed by solving Posson's equation in $\mathbb{R}^n$

using theorem 5.2 and then restrict it to a domain U; the problem is that we aren't really controlling what the value of the solution is on ∂U where we want it to equal g.

We calculate inspired by theorem 5.2 anyway with a C^2 function u, to see what boundary terms we get which will hopefully point us in the right direction. Considering a point $x \in U$ and $\epsilon > 0$ small enough so that $B(x,\epsilon) \subset U$ we define $V_\epsilon = U \setminus B(x,\epsilon)$. Using that Φ is harmonic away from the origin and Green's formula we have:

$$\begin{aligned}\int_{V_\epsilon} \Phi(y-x)\Delta_y u(y)dy &= \int_{V_\epsilon} u(y)\Delta_y \Phi(y-x) + \Phi(y-x)\Delta_y u(y)dy \\ &= \int_{\partial V_\epsilon} u(y)\frac{\partial \Phi}{\partial \nu}(y-x) - \Phi(y-x)\frac{\partial u}{\partial \nu}(y)dS(y) \\ &= J_1(S(x,\epsilon)) + J_1(\partial U) + J_2(S(x,\epsilon)) + J_2(\partial U)\end{aligned} \tag{9.2}$$

Where in the third line we use that $\partial V_\epsilon = S(x,\epsilon) \cup \partial U$ and denote by these terms are the corresponding boundary integrals. Now, because $\frac{\partial u}{\partial \nu}(y)$ is bounded the $J_2(S(x,\epsilon))$ term tends to 0 as ϵ does because Φ is in L^1_{loc}. Arguing as in theorem 5.2 the $J_1(S(x,\epsilon))$ term tends to $-u(x)$ as $\epsilon \to 0$. This gives that $u(x) = J_1(\partial U) + J_2(\partial U) - \int_{V_\epsilon} \Phi(y-x)\Delta_y u(y)dy$. If u is harmonic note we get the equation at the end of the last section.

For solving Poisson's equation what would be reasonable to try from this representation formula is to plug in $-f$ for $\Delta_y u(y)$, and then something into the boundary terms for g somehow. The problem is that these terms involve normal derivatives of u and not simply u which the prescribed boundary data involves. To deal with this what one can do is to add on a corrector term, a term ϕ^x such that

$$\begin{cases} -\Delta \phi^x = 0 \text{ in } U \\ \phi^x = \Phi(y-x) \text{ on } \partial U \end{cases} \tag{9.3}$$

The choice of this is so that the term $J_2(\partial U)$ is cancelled out. Assuming we have such a ϕ^x we calculate:

$$\begin{aligned}-\int_U \phi^x(y)\Delta u(y)dy &= \int_{\partial U} u(y)\frac{\partial \phi^x}{\partial \nu}(y) - \phi^x(y)\frac{\partial u}{\partial \nu}(y)dS(y) \\ &= \int_{\partial U} u(y)\frac{\partial \phi^x}{\partial \nu}(y) - \Phi(y-x)\frac{\partial u}{\partial \nu}(y)dS(y)\end{aligned} \tag{9.4}$$

The second term above is exactly what we need to cancel out $J_2(\partial U)$, of course. So, letting $G(x, y) = \Phi(y - x) - \phi^x(y)$ we get

$$u(x) = -\int_{\partial U} u(y) \frac{\partial G}{\partial \nu}(x, y) dS(y) - \int_U G(x, y) \Delta u(y) dy \tag{9.5}$$

Now, to solve Poisson's problem in U we should just need to plug in g for u in the first term and $-f$ for Δu in the second (at least at a formal level). This is really great, except that we now have to solve 9.3 – here f is zero and the boundary data is known which is an improvement compared to 9.1 but it still not necessarily easy without some symmetry assumptions which allow for some natural guesses. For instance, considering 9.3 and that Φ is radial and harmonic away from the origin a natural thing to try to do is to build ϕ^x out of Φ by setting $\phi^x = \Phi(y - \tilde{x})$ where $\tilde{x}$ is the same distance from every point on the boundary that x is, while also laying outside U. For the case that U is a halfspace one can do this by letting $\tilde{x}$ be reflection across the boundary plane. When U is the unit ball one has to do something a little bit more complicated, by taking $\tilde{x}$ to be the image of x under inversion across the unit sphere but then scaling by the norm of x. See Evans for more details. As an upshot we have the following, which we record for the sequel:

Theorem 9.1. *The Poisson problem 9.1 call be solved when U is a ball for $f \in C^2(U)$ and $g \in C^0(\partial U)$*

To be precise, Evans considers the Dirichlet problem on the ball, or when $f = 0$. Arguing as immediately below one can see easily that $f \in C^2(U)$ can then be covered.

10. Perron's method of subharmonic functions

Generally finding the corrector function above is hard, but just having theorem 9.1 in hand is enough to argue for more general domains: next we describe solving Poisson's problem 9.1 in a general (bounded) domain U using Perron's method, following section 2.8 of [6] more or less. Eventually we will know a few different approaches one can take to solve this problem but this one is nice because its relatively elementary and will formally introduce us to the notions of weak solution and barriers. What we will actually spend our time on is solving the Dirichlet problem in U:

$$\begin{cases} -\Delta u = 0 \text{ in } U \\ u = g \text{ on } \partial U \end{cases} \tag{10.1}$$

Or in other words the Poisson problem when $f = 0$ with the same assumptions in the previous section: U a C^2 domain and $g \in C^0(\partial U)$. Then to solve the Poisson problem in U for nonzero f we can solve the Poisson problem $-\Delta u = f$ on $\mathbb{R}^n$ using the Green's function on $\mathbb{R}^n$, finding a solution u_1: this argument used two derivatives of f so we take $f \in C^2(\overline{U})$ (this can then be extended to a C^2 function on all of $\mathbb{R}^n$ by extension theorems). Restricting u_1 to $\overline{U}$ we can then let $\overline{g} = g - u_1 \mid_{\partial U}$ and solve the Dirichlet problem above with it (i.e. as g above) to find a function u_2. Letting $u = u_1 + u_2$ then we see on ∂U $u \mid_{\partial U} = u_1 \mid_{\partial U} + g - u_1 \mid_{\partial U} = g$, and on U we see, using the linearity of the Laplacian, that $\Delta u = \Delta u_1 + \Delta u_2 = 0 + f = f$ so that u solves the Poisson problem on U.

Now, we say a $C^2(U)$ function is subharmonic (superharmonic) if $\Delta u \geq 0$ (≤ 0). The important fact about sub/superharmonic functions for the Perron method is the following comparison result, where $S = \partial B$:

Lemma 10.1. *Let u be a subharmonic (superharmonic) function on a ball $B \subset\subset U$ and h a harmonic function on B. Then if $u \leq h$ ($\geq$) h on S we also have $u \leq h$ ($\geq h$) in B.*

Proof: We'll consider just the subharmonic case because the superharmonic case because, if u is (sub/super)harmonic, $-u$ is (super/sub)harmonic. Denoting by $v = h - u$, v is a superharmonic function on B which is nonnegative along S. Mimicking (one of) the proofs of the maximum value principle for harmonic functions, if we let $w = v - \epsilon e^{cx_1}$ for some $c > 0$ and any $\epsilon > 0$ we see w is strictly superharmonic or in other words so that $\Delta w < 0$. By the second derivative test if the minimum of w is achieved within the interior of B it must be a point where $\Delta w \geq 0$, a contradiction showing the minimum of w is along S. Taking $\epsilon \to 0$ as before implies the same for v. Since v is nonnegative on the boundary, we have the claim. □

Or, in more plain terms, a subharmonic function will lay below a harmonic one with the same boundary data. This is easy to visualize in the one dimensional case, because then the harmonic functions are just the linear ones and subharmonic functions are convex so have graphs the roughly look like upwards facing paraboloids. The idea of the Perron method then is to realize a harmonic function as the supremum of subharmonic ones, which again we see is reasonable from the 1-d case.

In order to carry out Perron's method we will need to consider operations on subharmonic functions which might not result in something twice differentiable though, and so we will use the property of sub/superharmonic functions given in the lemma above to generalize the defintion of these functions to the space of merely continuous ones. The following can be interpreted as an instance of a weak solution for a PDE/PDI (inequality):

Definition 10.1. *A $C^0(U)$ function is subharmonic (superharmonic) if, when $B \subset\subset U$ and h is a harmonic function on B, $u \le (\ge) h$ on S implies the same on all of the ball, or in other words that the conclusion of the lemma above holds.*

Note that if u is superharmonic, $-u$ is subharmonic so by and large it suffices to just show properties for subharmonic functions. Considering that we already have apriori knowledge of many such harmonic "competitors" from theorem 9.1, this is a promising defintion. The way its defined is a very common theme in PDE and adjacent fields: often the function space on which solutions are defined are too small to perform some operation one would like to do, such taking limits, so one widens the class of functions by taking a property "classical" solutions enjoy and crafting a defintion of weak solution around that. In hindsight, note we could have similarly defined, for instance, a weak notion of harmonic function via the mean value property for continuous functions although these turned out to all be smooth anyway.

The next few lemmas collect some essential facts on sub/superharmonic functions which will be needed in the Perron method. These are stated only for subharmonic functions, but one can check that if u is superharmonic then $-u$ is subharmonic so immediately give analogous statements superharmonic ones. The first shows that the maximum principle holds for C^0 subharmonic functions, just as it does for classical (C^2 with $\Delta u \ge 0$) subharmonic functions.

Lemma 10.2. *Supposing that $u \in C^0(\overline{U})$ is subharmonic, then u satisfies the strong maximum principle: if U is connected and there exists a point $x_0 \in U$ such that*

$$u(x_0) = \max_{\overline{U}} u$$

then u is constant within $\overline{U}$. As a consequence u satisfies the "regular" maximum principle as in the harmonic case.

Proof: Suppose that the maximum of u in $\overline{U}$ is attained at $x_0 \in U$ as in the statement, and let $B \subset\subset U$ be a small ball about x_0. From theorem 9.1 we can find a harmonic function h on B with the same values as u along S, and by the definition of subharmonic $h - u$ is nonnegative in B. Hence $u(x_0) \le h(x_0) \le \max_{\partial B} h = \max_{\partial B} u \le u(x_0)$, using the maximum principle for h in the second inequality, so that equality must hold throughout. By the strong maximum principle for h then h is constant on B, in particular S, so that u is constant on S as well an equal to $u(x_0)$ there. Now, we can restart this argument with any of the points along S, and iterate it further again and again. So using this by considering $x, y \in U$ and a chain of appropriately picked balls/spheres all compactly contained in U so that x and y are connected by a chain of subarcs of these to the original ball S, we get that $u(x) = u(y)$ so that u is constant in U giving the claim. □

With the rough idea of the Perron method given above in mind, one can imagine the following simple observation below will be useful because it says a "value increasing" operation on a collection of subharmonic functions will stay within in the class of subharmonic functions:

Lemma 10.3. *Let $u_1, u_2, \ldots, u_N$ be subharmonic in U. Then the function $u(x) = \max\{u_1(x), \ldots, u_N(x)\}$ is also subharmonic in U.*

Proof: Take a ball $B \subset\subset U$ and $u \leq h$ on S as in the previous lemma. Then since each of the $u_i \leq u$ on S they are less than h on S as well so by subharmonicity less than h on B. Hence $u \leq h$ in B too. □

Considering a continuous function u on U and a ball $B \subset\subset U$, we can consider the harmonic function $h = u$ along S and so define a new function given by:

$$\tilde{u}(x) = \begin{cases} h(x) \ x \in B \\ u(x) \ x \in U \setminus B \end{cases} \tag{10.2}$$

This is called the harmonic lifting of u in B. When u is subharmonic, its immediate from the defintion that any harmonic lifting of it will be larger than it. So, if the harmonic lifting of a subharmonic function is subharmonic one can imagine that it can be used to "improve" a sequence of subharmonic ones to see the limit is harmonic. The following lemma says this wish comes true:

Lemma 10.4. *The harmonic lifting $\tilde{u}$ of a subharmonic function u is subharmonic.*

Proof: To check this we consider an arbitary ball $B' \subset\subset U$ (of course, not just the ball we lifted on), a harmonic function h on that for which $\tilde{u} \leq h$ on S', and we must show that $\tilde{u} \leq h$ on all of B'. Now, notice that since u is subharmonic we have $u \leq \tilde{u}$, and also that they are equal outside the ball B where the lifting was done. The first observation gives that in particular $u \leq h$ on S, and the second observation combined with the subharmonicity of u again gives that in $B' \setminus B$ $\tilde{u} \leq h$. As a consequence of this $\tilde{u} \leq h$ on $\partial(B' \cap B)$, so since $\tilde{u}$ is harmonic in $B' \cap B$ we have that it is bounded by h in this set as well.

□

Now we are ready to describe the Perron method in more detail. Given a bounded domain U and a bounded function g on ∂U (we'll just take it to be continuous, which implies its bounded), a $C^0(\overline{U})$ subharmonic function is called a subfunction relative to g if it satisfies $u \leq g$ on ∂U. Although we won't need these immediately a superharmonic function is called a superfunction relative to g if its not less than g along ∂U. If we denote by S_g the set of subfunctions relative to g on U, then as alluded to the following holds:

Theorem 10.5. *The function $u(x) = \sup_{v \in S_g} v(x)$ is harmonic in U.*

Proof: Because constant functions are harmonic and hence subharmonic in the classical sense, we see the set S_g is nonempty considering the function $v = \min_{\partial U} g$ (using g is continuous and U is bounded this is $> -\infty$). By the definition of subfunction at any point $x \in U$ and any subfunction $v(x) \leq \max_{\partial U} g$, this supremum is a well defined finite number everywhere. Of course this space of functions S_g could be huge, perhaps even uncountable and its not really clear we could realize u as the limit of a single sequence of functions simultaneously at every point, but fixing a point x we may consider a sequence $v_k \in S_g$ with $v_k(x) \to u(x)$ since the reals are second countable. By replacing v_k in the sequence with $\max\{v_1, \dots v_k, \min_{\partial U} g\}$ we may suppose it is bounded below (at every point, not just x), increasing, and still a sequence of subfunctions using lemma 10.3. Now we pull out our harmonic lifting trick: fixing $B(x,r) \subset\subset U$ we then replace the v_k in B with their harmonic lifts to get a seuquence of functions $\tilde{v_k} \geq v_k$. Since the $v_k(x) \to u(x)$ and the $\tilde{v_k}$ are subharmonic from the lemma above we have $\tilde{v_k}(x) \to u(x)$, which towards our goal of showing u is harmonic is promising. On the flip side this convergence of harmonic functions to u in B we only have holding at a single point right now.

Since the v_k are an increasing sequence their harmonic lifts are and since the sequence $\tilde{v_k}$ is an increasing sequence of, in fact, harmonic functions in $B(x,r)$ we may employ the Harnack convergence theorem, theorem 7.4 (or the other convergence theorem above that used the derivative estimates) to see that in $B(x, r/2)$ the sequence $\tilde{v_k}$ converges to a harmonic function v. By the defintion of u we have

$v \leq u$ and to show the theorem it suffices to see they are equal in the ball. Arguing by contradiction if they aren't equal, there is a subharmonic function v' and a point $y \in B(x, r/2)$ such that $v(y) < v'(y) < u(y)$. Defining yet another sequence $w_k = \max\{v_k, v'\}$ (v_k from before), and doing the harmonic lifting in $B(x, r)$ along this sequence, we then get a harmonic function w on $B(x, r/2)$ such that $v \leq w \leq u$. Importantly, because $v < v'$ strictly at y and the harmonic lifting is an increasing operation we have $v \neq w$. On the other hand by the mean value property $v(x) = \fint_{B(x,r/2)} v \leq \fint_{B(x,r/2)} w = w(x) \leq u(x) = v(x)$, so we must have equality throughout implying since $v \leq w$ that they are equal in the ball. So, we've reached a contradiction giving the claim. □

So, we have a method then which produces a harmonic function and if for instance $g < 0$ at a point we can be sure that from the defintion of subfunction it will be nonzero, so we've produced something that will sometimes be nonzero so different from what using the Green's function on $\mathbb{R}^n$ would produce (although one supposes it could still be pretty boring, like a constant). What we really want to know of course is if the solution produced will actually agree with g on ∂U. This brings us to the concept of barrier argument, which for a PDE can be generally thought of as the idea that solutions to a (related) PDE/PDI, oftentime which we understand well, can be used to control the behavior of the solution we are actually interested in via the maximum principle. For instance, in many PDE only solutions which are extremely symmetric are very well understood, but these can be used as barriers to still tell us a lot about more general solutions.

In our context, given a point $\xi \in \partial U$ we say that a $C^0(\overline{U})$ function w is a barrier at ξ relative to U if:

(1) w is superharmonic in U, and
(2) $w = 0$ at ξ but is > 0 at all other points in $\overline{U}$

Barrier functions are not guaranteed to exist; a boundary point will be called regular if there exists a barrier at that point. Our defintion of barrier is the right notion by the following lemma:

Lemma 10.6. *Let u be the harmonic function produced by the Perron method above. Then if $\xi \in \partial U$ is a regular boundary point, then $u(x) \to g(\xi)$ as $x \to \xi$.*

Proof: (NB: if you are also reading [6] they include the assumption " and g is continuous at ξ" but here we are just assuming g is continuous from the start.) Since ξ

is a regular boundary point, there is a barrier function w at ξ. Fix an $\epsilon > 0$, we set the following notation/values:

(1) let $M = \sup |g|$ which is finite since U is bounded and g is continuous,
(2) also using continuity of g pick δ so that $|g(x) - g(\xi)| < \epsilon$ if $|x - \xi| < \delta$,
(3) Using continuity of w and boundedness of U, pick k so that $kw(x) \geq 2M$ if $|x - \xi| \geq \delta$ for $x \in \partial U$.

With these constants in mind, we define the functions $v_1(x) = g(\xi) - \epsilon - kw(x)$ and $v_2(x) = g(\xi) + \epsilon + kw(x)$. From the choice of constants one can see, checking the cases $x \in B(\xi, \delta)$ and $x \in B(\xi, \delta)^c$ separately, that $v_1 < g$ and $v_2 > g$ on ∂U. Since w is superharmonic v_1 is a subfunction then with respect to g and similarly v_2 is a superfunction. One can check sums of subharmonic functions are subharmonic, so since $-v_2$ is a subharmonic function and that the boundary condtion satisfied by $v_1 - v_2$ is nonpositive that $v_1 - v_2 \leq 0$ in U by the maximum principle shown above so that v_2 is greater than any subfunction. Thus by the definition of u we see that $v_1 < u < v_2$ on ∂U, so in particular $|u(\xi) - g(\xi)| < \epsilon$. Letting $\epsilon \to 0$ gives the claim. □

So, to know that $u = g$ on ∂U we just need to exhibit a barrier at every point. A natural place to start of course are the known explicit solutions like our old friend the Green's function, using that harmonic functions are simultaneously subharmonic and superharmonic. Using them we can construct barriers if U satisfies the so–called exterior sphere condition, which is that for every point $\xi \in \partial U$ there exists a ball $B = B_R(y) \subset U^c$ such that $\overline{B} \cap \overline{U} = \xi$. Specifically if this condition is fullfilled then the function(s) w given by

$$w(x) = \begin{cases} R^{2-n} - |x-y|^{2-n} \text{ for } n \geq 3 \\ \log \frac{|x-y|}{R} \text{ for } n = 2 \end{cases} \tag{10.3}$$

will be a barrier at ξ. Here, the exterior sphere condition is used to "isolate" the singularity of the Green's function, where its not defined, away from U; elsewhere we recall its harmonic. The exterior sphere condition is also used in getting positivity of w away from ξ: these functions are radial about the point y and are increasing functions in r on the spheres $S(y, r)$ because they come (up to scalling) from Green's functions. Recalling that the boundary of a domain U is C^k if in a neighborhood of every point in the boundary ∂U can be written as the graph of a C^k function, and that the second derivative (if defined) of this graph is roughly the curvature of

∂U, one can see for bounded U that if ∂U is C^2 it will satisfy the bounded exterior sphere condition. Putting this all together, we have:

Theorem 10.7. *The problem 9.1 is solvable when U is a bounded domain with C^2 boundary, $f \in C^2(\overline{U})$, and $g \in C^0(\partial U)$.*

Note that this isn't a sharp result, particularly when $n = 2$. See [6] for more discussion on this matter. I'll end by mentioning in passing Schwartz's alternating method, which bears some similiarties to Perron's method in that it uses one can solve the Dirichlet problem on a special class of domains and which can be "bootstrapped" to more complicated ones. As an idea of how it works, take a domain given by the union of two overlaping discs D_1, D_2. Then the boundary data pescribes boundary data on at least some of the boundaries of the two discs considered sepearately, except where the boundaries $\partial D_1, \partial D_2$ lay in the intersection $D_1 \cap D_2$. In the method one plugs in by hand initial data along "missing part" of the boundary of D_1, solve the Dirichlet problem on that disc, and then use that to define the missing boundary data for D_2. Then one solves the Dirichlet problem on D_2, and uses that to update what the missing boundary data should D_1 (hence, the name). Going back and forth one can see the functions settle out and converge to a solution to the Dirichlet problem on $D_1 \cup D_2$. See [11] for a more detailed account.

11. The energy of a function and the heat equation

One might recall from physics that associated to many physical models there is an associated energy/action, a functional on the space of relevent space of functions, for which solutions to the model correspond to critical points. This point of view can be extended to many PDE without reference persay to any physical interpretation and is often quite fruitful: such problems are called variational problems. The Poisson problem is an example of a variational problem and the correct energy to consider is the following:

$$I[w] = \int_U \frac{1}{2}|Dw|^2 - wf dx \tag{11.1}$$

where w belongs to the adminissible set:

$$\mathcal{A} = \{w \in C^2(\overline{U}) \mid w = g \text{ on } \partial U\} \tag{11.2}$$

The following, which can be taken as justification for this energy, is called Dirichlet's principle:

Theorem 11.1. *Assume $u \in C^2(\overline{U})$ solves Poisson's equation. Then $I[u] = \min w \in \mathcal{A}$. Conversely if $u \in A$ is a minimum for the energy I it solves Poisson's equation.*

Proof: First we suppose u satisfies Poisson's equation, and consider some other function $w \in \mathcal{A}$. Since $-\Delta u = -f$ we have:

$$0 = \int_U (-\Delta u - f)(u - w)dx \tag{11.3}$$

Using that $u - w = 0$ on ∂U since both are in $\mathcal{A}$ (w by assumption, and u again since it solves Poisson's equation) we have by integration by parts then:

$$0 = \int_U Du \cdot D(u-w) - f(u-w)dx = \int_U Du \cdot Du - Du \cdot Dw - fu + fwdx \tag{11.4}$$

Rearranging terms then, we see we have:

$$\int_U |Du|^2 - ufdx = \int_U Du \cdot Dw - wfdx \tag{11.5}$$

Considering our goal, to compare the energy of u to that of w, we see we must be getting close. Using that $|Du \cdot Dw| \leq |Du||Dw| \leq \frac{1}{2}|Du|^2 + \frac{1}{2}|Dw|^2$ we see:

$$\int_U |Du|^2 - ufdx = \int_U Du \cdot Dw - wfdx \leq \int_U \frac{1}{2}|Du|^2 + \frac{1}{2}|Dw|^2 - wfdx \tag{11.6}$$

Subtracting over $\frac{1}{2}|Du|^2$ from the far left gives that $I[u] \leq I[w]$, as claimed. Now we consider the other direction, that if $u \in \mathcal{A}$ is a minimum point of the energy I then u solves the Poisson equation. Towards this end consider an arbitrary $v \in C_c^\infty(U)$ and write $i(\tau) = I[u + \tau v]$, where $\tau \in \mathbb{R}$. Since $u + \tau v \in \mathcal{A}$ for all τ the i should have a minimum at $\tau = 0$; by the dominated convergence theorem one can see i is differentiable so in particular we should have $i'(0) = 0$. To see what this implies, we write out $i(\tau)$ in a way which separates out τ a bit more:

$$i(\tau) = \int_U \frac{1}{2}|Du + \tau Dv|^2 - (u + \tau v)fdx \quad = \int_U \frac{1}{2}|Du|^2 + \tau Du \cdot Dv + \frac{\tau^2}{2}|Dv|^2 - uf - \tau vfdx \tag{11.7}$$

We can easily compute the τ derivative of this term by term (moving the derivative through the integral) to see the following, where the second equality is by integration by parts:

$$0 = i'(0) = \int_U Du \cdot Dv - vfdx = \int_U (-\Delta u - f)vdx \tag{11.8}$$

The integrand in the last term, with v separated out, is called the first variaton of I and is exactly the first equation in the Poisson problem. It turns out the second variation is also often useful say in geometric problems but we don't need to consider it right now. Now, v was an arbitrary function in $C_c^\infty(U)$. Suppose that $-\Delta u - f$ wasn't equal to zero on U. Then by continuity there is a point x and ball

$B(x,r) \subset U$ where it is nonzero and doesn't change sign. Picking v to be a bump function supported on it and equal to one in $B(x, r/2)$ then gives a contradiction, so that indeed $-\Delta u - f = 0$ in U. That $u = g$ on ∂U is built directly into the defintion of the space $\mathcal{A}$, so u solves the Poisson equation on U. □

The question of existence for the Poisson problem then is equivalent to the existence of minimizers for the functional I. Of couse apriori its not even obvious that $\min_{w\in\mathcal{A}} I[w] > -\infty$ but this rephrasing of the problem opens up new potential avenues of attack which can be quite useful. Supposing the minimum is bounded and denoting this number by m one way to proceed, called the direct method, is to consider a sequence of functions $u_i \in \mathcal{A}$ such that $I[u_i] \to m$. Then our hope would be that the limit $\{u_i\}$ actually converges to a function u and that $I[u] = m$ (requiring that I is lower semicontinuous). The convergence it turns out is tricky and will require us to consider broader spaces of functions which are complete under the notion of convergence best suited to the problem – we'll return to it later (hopefully).

Another idea is to take an initial function and try to deform it in a way which consistently decreases the associated energy – that is given a function use it as initial data for a PDE that looks something life $u_t = L(u)$ for some partial differential operator L. If a solution to such a PDE exists classically for all time, which of course itself could potentially be a big request and mean a lot of work or simply not be true, we can hopefully avoid any technical issues involving the "broader spaces of functions" that we alluded to above. One can see from 11.8 that when $f = 0$ $L = \Delta$ works, or in other words as $t \to \infty$ a solution to the problem

$$\frac{\partial u}{\partial t}(x,t) = \Delta_x u, \ u(x,0) = f(x) \tag{11.9}$$

Should converge to a harmonic equation if all works out well. Of course, the PDE above is exactly the heat equation, which we discuss next.

12. The fundamental solution of the heat equation on $\mathbb{R}^n$

Mirroring the development above for Laplace's equation, we start with the fundamental solution (Green's function) for the heat equation on $\mathbb{R}^n$ – this is often called the heat kernel. One way to find the heat kernel is by Fourier transform. The idea, along similar lines as indicated in theorem 3.2 above is that under Fourier transform (in only spatial coordinates) the heat equation will be transformed to an ODE which can be easily explictly solved, and when one does so and applies the inverse fourier transform the Green's function can be read off – this might be discussed more in the exercise section and is written down in section 4.3 of [5]. Another method for arriving at the Green's function is to proceed as we did for the Laplace equation, which is by starting from an ansatz based off a symmetry of the problem and getting lucky. In chapter 2 of [5] they argue by looking for solutions of the form

$$u(x,t) = \frac{1}{t^{\alpha}} v(\frac{x}{t^{\beta}}) \tag{12.1}$$

For α, β, v to be determined. To give a little bit of motivation for this ansatz, by the chain rule and using that a derivative is taken just once in time but twice spatially one can see that if $u(x,t)$ is a solution to the heat equation, then so does $u(\lambda x, \lambda^2 t)$ for $\lambda \in \mathbb{R}$. One sometimes says this is how the PDE scales, and sending $(x,t) \to (\lambda x, \lambda^2 t)$ is refered to as parabolic rescaling. As a result, the ratio $\frac{|x|}{\sqrt{t}}$, where here t is positive, is preserved under the scaling. So, considering the aesthetically pleasing philosophy that for a given symmetry of a PDE there should be a solution to that PDE which respects it, one hopes that there might be a solution $u(x,t)$ of the form $v(\frac{r}{\sqrt{t}})$ of the heat equation on $\mathbb{R}^n \times (0,\infty)$ for an appropriate function v. One would say here we would be looking for a solution invariant under parabolic rescaling.

Starting with this ansatz can be made to work but the ansatz given above is closer to the form of the Green's function and so leads to the answer quicker and there are some heuristics based off, say, mass invariance for the factor of t in front. By assuming that v is radial and proceeding much as in the case for the Laplacian one finds that for u to solve the heat equation there is a related ODE for v to solve. The constants α, β are decided in the course of things to be $n/2, 1/2$ respectively to make the equation one finds for v simpler. Instead of belaboring this here/in lecture for the sake of time working this out will be left as an exercise. If we make the additional assumption that v and its derivative decays to zero, natural in our hunt for fundamental solution, we find solutions of the form $\frac{b}{t^{\frac{n}{2}}} e^{\frac{-|x|^2}{4t}}$. If we impose the condition that its mass/total integral for fixed $t > 0$ is one, then we see $b = \frac{1}{(4\pi)^{n/2}}$ – this uses Fubini's theorem and the well known fact that $\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}$. So, we've sketched out a path to a solution of the form

$$\Phi(x,t) = \frac{1}{(4\pi t)^{n/2}} e^{\frac{-|x|^2}{4t}} \tag{12.2}$$

to the heat equation for $(x,t) \in \mathbb{R}^n \times (0,\infty)$. One may extend the defintion of this function to be zero for $t < 0$. The main claim of this section is that it is indeed the fundamental solution to the heat equation/heat kernel/Green's function for the heat equation on $\mathbb{R}^n$:

Theorem 12.1. *Assume g is a continuous and bounded function. Then if we set u to be*

$$u(x,t) = \int_{R^n} \Phi(x-y,t)g(y)dy \tag{12.3}$$

then u has the following properties:

(1) $u \in C^\infty(\mathbb{R}^n \times (0,\infty))$,
(2) $u_t(x,t) - \Delta u = 0$ *for* $(x,t) \in \mathbb{R}^n \times (0,\infty)$,
(3) $\lim_{(x,t)\to(x^0,0)} u(x,t) = g(x_0)$ *for each point* $x^0 \in \mathbb{R}^n$

Proof: Item (1) follows from the heat kernel being smooth with uniformly bounded derivatives of all orders on $\mathbb{R}^n \times [\delta,\infty)$ for each $\delta > 0$, justifying pulling derivatives through the integral sign. Concerning item (2) by the same reasoning with regards to limits we see that

$$u_t - \Delta u = \int_{R^n} [(\Phi_t - \Delta_x \Phi)(x-y,t)]g(y)dy = 0 \tag{12.4}$$

Using that Φ solves the heat equation on $\mathbb{R}^n \times (0,\infty)$ and a trivial use of the chain rule. For item (3), by the continuity of g we have for each $\epsilon > 0$ ther exists $\delta > 0$ such that $|g(y) - g(x^0)| < \epsilon$ if $|y - x^0| < \delta$. So, if $|x - x^0| < \delta/2$ we have, using that Φ integrates to 1 on each timeslice:

$$\begin{aligned} |u(x,t) - g(x^0)| &= |\int_{\mathbb{R}^n} \Phi(x-y,t)(g(y) - g(x^0))dy| \\ &\leq \int_{B(x^0,\delta)} \Phi(x-y,t)|g(y) - g(x^0)|dy \\ &+ \int_{\mathbb{R}^n \setminus B(x^0,\delta)} \Phi(x-y,t)|g(y) - g(x^0)|dy = I + J \end{aligned} \tag{12.5}$$

Becase the integral of Φ is one and in $B(x^0,\delta)$ we have $|g(y) - g(x^0)| < \epsilon$, $I < \epsilon$. For the second term by the triangle inequality we have:

$$|y - x^0| \leq |y - x| + |x - x^0| \leq |y - x| + \frac{\delta}{2} \leq |y - x| + \frac{1}{2}|y - x^0| \tag{12.6}$$

Where in the first inequality we used that $|x - x^0| < \frac{\delta}{2}$ and for the second inequality we used that we integrate over the set $\mathbb{R}^n \setminus B(x^0,\delta)$ in the second term. This inequality implies that $|y - x| \geq \frac{1}{2}|y - x^0|$, which will be used in the third inequality below:

$$\begin{aligned} J &\leq 2||g||_{L^\infty} \int \int_{\mathbb{R}^n \setminus B(x^0,\delta)} \Phi(x-y,t)dy \\ &\leq \frac{C}{t^{n/2}} \int_{\mathbb{R}^n \setminus B(x^0,\delta)} e^{\frac{-|x-y|^2}{4t}} dy \\ &\leq \frac{C}{t^{n/2}} \int_{\mathbb{R}^n \setminus B(x^0,\delta)} e^{\frac{-|y-x^0|^2}{16t}} dy \end{aligned} \tag{12.7}$$

Here C is a dimensional constant we can write C explicitly, of course, but it doesn't matter for our purposes. The point of this inequality for J is that we've now bounded it by an integral where the center point of the ball involved in the domain and the point that y is offset by in the integrand agree, and we can use the change of coordinates $z = \frac{y-x^0}{\sqrt{t}}$ to see the final term above is equal to:

$$C \int_{\mathbb{R}^n \setminus B(0,\delta/\sqrt{t})} e^{-\frac{z^2}{16}} dz \tag{12.8}$$

Now, the integrand is positive and in $L^1(\mathbb{R}^n)$ and so its easy to see that for any $\epsilon > 0$ there exists R large enough so that its integral over $\mathbb{R}^n \setminus B(0,R)$ is less than epsilon. Since $\delta/\sqrt{t} \to \infty$ we see then that the integral above tends to zero as t does, completing the proof. □

Parameterized in t, one can see that the heat kernel $\frac{1}{(4\pi t)^{n/2}} e^{\frac{-|x|^2}{4t}}$ looks like a family of Gaussians (bell curves) which as $t \to 0^+$ become concentrated more and more tightly about the origin and as t increases become more and more diffuse. So, considering a solution u as defined by the integral above we see that as $t \to \infty$ we should have u converge to a constant which is harmonic, so in other words it should be doing what we expect in this case. Considering the heat equation corresponds to the temperature distribution in a medium, this is sensible considering everyday experience (e.g. a pie left out on a table will eventually cool down to room temperature).

Something that is a little bit more subtle and arguably nonsense from a physics perspective is that the heat equation has *infinite propogation speed*, notice that if g is bounded, continuous, nonnegative but not identically zero then $u(x,t)$ defined by convolution of g with Φ as above:

$$u(x,t) = \frac{1}{(4\pi t)^{n/2}} \int_{\mathbb{R}^n} e^{\frac{-|x-y|^2}{4t}} g(y)dy \tag{12.9}$$

is strictly positive at every point x for $t > 0$. And in fact, given a point $p \in \mathbb{R}^n$ one can see that, even if g is just supported in the ball $B(0,1)$, $u(p,t)$ can be arranged to be as large as we want for any positive $t > 0$ by taking g to be suitably large. In contrast many curvature flows have an important and useful quality called "pseudolocality." Roughly said, pseudolocality says that if a manifold is very flat in a very large ball initially then it is relatively close to being flat in a (much) smaller ball for a short period of time later, independent of how curved the manifold is elsewhere.

13. Duhamel's principle

An interesting phenomena about the heat equation and other linear evolution equations – roughly speaking linear PDE where there is a time variable involved, is that solutions to the inhomogenous problem, i.e. those of the form:

$$\begin{cases} (u_t - \Delta u)(x,t) = f(x,t) \text{ in } \mathbb{R}^n \times (0,\infty) \\ u = 0 \text{ on } \mathbb{R}^n \times \{0\} \end{cases} \tag{13.1}$$

where f is possibly nonzero, can be fruitfully as the sum of solutions to the homogenous problem(s):

$$\begin{cases} (u_t - \Delta u)(x,t) = 0 \text{ in } \mathbb{R}^n \times (s,\infty) \\ u(x,s) = f(x,s) \text{ on } \mathbb{R}^n \times \{s\} \end{cases} \tag{13.2}$$

Of course, in the nonhomogenous problem if we want to include a nonzero boundry term we can by solving the related homogenous problem and adding it to a solution of 16.12. In other words, there is in some sense a way to trade (in terms of problem to solve) the forcing function and the boundary data for the heat equation. One might suppose this principle is reasonable using the linearity of heat equation because it seems basically plausible (obvious disclaimer: I am not a physicist) that a solution to the nonhomogenous problem up to a given time could be approximated by a number solutions to the homogenous problem with initial heat "pulses" given by the driving function, say if one imagines diping their hands in and out of a cold water bath quickly versus just leaving it in. One imagines the approximation would get better as the pulses are considered spaced more finely as well. More precisely for the heat equation the claim is the function

$$u(x,t) = \int_0^t \int_{\mathbb{R}^n} \Phi(x-y,t-s) f(y,s)dyds = \int_0^t \int_{\mathbb{R}^n} \frac{1}{(4\pi t)^{n/2}} e^{\frac{-|x-y|^2}{4t}} f(y,s)dydt \tag{13.3}$$

is a solution to the nonhomogenous problem, where here $f \in C_1^2(\mathbb{R}^n, [0,\infty))$ has compact support. What the notation $C_k^\ell(\mathbb{R}^n, [0,T]$ (apparently nonstandard, but what Evans uses) notation means is that for a function u in the space its spatial derivatives are ℓ times differentiable and that the temporal derivatives are k times differentiable with no claims on the mixed derivatives in x and t; for instance, a function in $C_1^2(\mathbb{R}^n, [0,\infty))$ has $u, D_x u, D_x^2 u, u_t \in C(\mathbb{R}^n \times [0,\infty))$.

The proof that u as defined above satisfies the nonhomogenous equation is obfuscated somewhat by the singular nature of ϕ as $t \to 0$, so we give a simple ODE example first highlighting as well the generality of the argument. Suppose we wish to solve the ODE

$$\begin{cases} \frac{du}{dt} - cu(t) = f(t), t > 0 \\ u(0) = 0 \end{cases} \tag{13.4}$$

where f is some suitably good function – this is written to look like the heat equation of course. You've probably seen this how to approach this sort of problem early on in an ODE class under the name "variation of parameters," called such because in the homogenous case there is a constant of integration that is replaced with a function for an ansatz for the nonhomogenous case (which isn't strictly necessary, but that's getting off topic). At any rate we see that the problem

$$\begin{cases} \frac{dx}{dt} - cx(t) = 0, t > s \\ x(s) = f(s) \end{cases} \tag{13.5}$$

is trivial to solve, with solution given by $x(t) = f(s)e^{c(t-s)}$. Then our claim translated over in this toy case is that the function $u(t) = \int_0^t f(s)e^{c(t-s)}ds$ solves the nonhomogenous problem. Calculating we have:

$$\frac{d}{dt}u(t) = \frac{d}{dt}\int_0^t f(s)e^{c(t-s)}ds = c\int_0^t f(s)e^{c(t-s)} + f(t) = cu(t) + f(t) \tag{13.6}$$

So that u is indeed a solution. Roughly speaking, in the second equality the first term is what one would get if the upper limit in the integral was fixed and we differentiate the integrand and the extra term is from the fundamental theorem of calculus – this extra term is what snuck in the nonhomogenous term and is the mathematical reason why Duhamel's principle works – and one should be able to work backwards from this observation since the integrand can be general to find Duhamel's formula. With this in mind, our specific claim for the heat equation is the following:

Theorem 13.1. *Defining u as above in 13.3, then*

(1) $u \in C_1^2(\mathbb{R}^n \times (0,\infty))$,
(2) $u_t - \Delta u = f$ *for* $(x,t) \in \mathbb{R}^n \times (0,\infty)$,
(3) $\lim_{(x,t)\to(x^0,0)} u(x,t) = 0$ *for each point* $x^0 \in \mathbb{R}^n$

Proof: Because Φ has a singularity at $(0,0)$ (note that it limits to zero for other points $(x,0)$, because the exponential term "beats" $t^{-n/2}$) we first change variables to see that

$$u(x,t) = \int_0^t \int_{\mathbb{R}^n} \Phi(y,s) f(x-y,t-s)dyds \tag{13.7}$$

Because $f \in C_1^2(\mathbb{R}^n \times (0,\infty))$, since $(0,\infty) \subset [0,\infty)$, and Φ is locally integrable, one can justify passing derivatives through the integrals to see all for any fixed t u is twice differentiable in x and given by replacing f in the formula above with its spatial derivatives. The temporal derivative of u also exists but calculating it is a little bit more subtle and is where the magic happens as discussed in the toy example above, because t is also in the bounds for the integral:

$$u_t(x,t) = \int_0^t \int_{\mathbb{R}^n} \Phi(y,s) f_t(x-y,t-s)dyds + \int_{\mathbb{R}^n} \Phi(y,t) f(x-y,0)dy \tag{13.8}$$

Putting this together gives that

$$u_t - \Delta u(x,t) = \int_0^t \int_{\mathbb{R}^n} \Phi(y,s)(\frac{\partial}{\partial t} - \Delta_x) f(x-y,t-s)dyds + \int_{\mathbb{R}^n} \Phi(y,t) f(x-y,0)dy \tag{13.9}$$

Using that Φ isn't as well behaved approaching $t = 0$, we wish to do our usual trick of breaking $[0,t]$ up into $[0,\epsilon] \cup (\epsilon,t]$ for some $0 < \epsilon < t$. What comes often comes after in these arguments is an integration by parts, but the derivatives are on x and t and not y and s so we fix that first. We see that $\frac{\partial}{\partial t} f(x-y,t-s) = -\frac{\partial}{\partial s} f(x-y,t-s)$, and using that $(-1)^2 = 1$ that $\Delta_x f(x-y,t-s) = \Delta_y f(x-y,t-s)$ so we have:

$$\begin{aligned} u_t - \Delta u(x,t) &= \int_0^t \int_{\mathbb{R}^n} \Phi(y,s)(-\frac{\partial}{\partial s} - \Delta_y) f(x-y,t-s)dyds + \int_{\mathbb{R}^n} \Phi(y,t) f(x-y,0)dy \\ &= \int_\epsilon^t \cdots + \int_0^\epsilon \cdots + \int_{\mathbb{R}^n} \Phi(y,t) f(x-y,0)dy \\ &= I_\epsilon + J_\epsilon + K \end{aligned} \tag{13.10}$$

Starting with the middle term J_ϵ first, because $f \in C_1^2(\mathbb{R}^n, [0,\infty))$ and has compact support we have

$$|J_\epsilon| \leq (||f_t||_{L^\infty} + ||D^2 f||_{L^\infty}) \int_0^\epsilon \int_{\mathbb{R}^n} \Phi(y,s)dyds \leq \epsilon C \tag{13.11}$$

Where in the last line we are using the normalization of the mass of Φ; so, J_ϵ tends to zero which will be used later. For the I_ϵ integral we have by integration by parts that:

$$\begin{aligned} I_\epsilon = &\int_\epsilon^t \int_{\mathbb{R}^n} [(\frac{\partial}{\partial s} - \Delta_y)\Phi(y,s)] f(x-y,t-s)dyds \\ &+ \int_{\mathbb{R}^n} \Phi(y,\epsilon) f(x-y,t-\epsilon)dy \\ &- \int_{\mathbb{R}^n} \Phi(y,t) f(x-y,0)dy \end{aligned} \tag{13.12}$$

The second two terms are the temporal boundary terms one gets from integration by parts. There are no spatial boundary integrals because we are integrating in spatial coordinates over $\mathbb{R}^n$ and f is compactly supported. Because Φ solves the heat equation for $t > 0$ the very first term is zero and the third term is precisely $-K$, we thus have:

$$u_t - \Delta u(x,t) = I_\epsilon + J_\epsilon + K = \int_{\mathbb{R}^n} \Phi(y,\epsilon) f(x-y,t-\epsilon)dy + J_\epsilon \tag{13.13}$$

The left hand side doesn't depend on ϵ, so taking $\epsilon \to 0$ we see since $J_\epsilon \to 0$ and theorem 12.1 that $u_t - \Delta u(x,t) = f(x,t)$ giving item (2). For item (3), note from its definiton and the normalization of Φ that:

$$||u(x,t)||_{L^\infty} \leq t||f||_{L^\infty} \tag{13.14}$$

which tends to zero as u does. □

14. The maximum principle for the heat equation

At this point, we could further develop the solvability of the heat equation on general domains but mirroring the development for the Laplace equation (and to mix it up a bit) let's move on to discuss some properties of solutions to the heat equation. It turns out that just like the laplace equation, the heat equation satisfies mean and maximum principles, although they are naturally more complicated to state. As a somewhat amazing historical tidbit the mean value formula for the heat equation apparently didn't appear until the 1970s! (see [25]). To discuss it first we define the so–called heat ball; for fixed $x \in \mathbb{R}^n, t \in \mathbb{R}$, and $r > 0$ we denote:

$$E(x, t; r) = \{(y, s) \in \mathbb{R}^{n+1} \mid s \leq t, \Phi(x - y, t - s) \geq \frac{1}{r^n}\} \tag{14.1}$$

This will be the set that is integrated over below. As a bit of motivation for why this is natural, the regular (Euclidean) balls we averaged over for harmonic equations could have been written similarly with respect to the fundamental solution to the Laplacian using that they were radial. Then the mean value property for the heat equation is the following assertion:

Theorem 14.1. *Let $u \in C_1^2(U_T)$ solve the heat equation. Then*

$$u(x, t) = \frac{1}{4r^n} \int\int_{E(x,t;r)} u(y, s) \frac{|x - y|^2}{(t - s)^2} dyds \tag{14.2}$$

Note that in the defintion of heat ball that the point (x, t) is at the "top" of the ball, which makes sense because the value of $u(x, t)$ shouldn't depend on the value of u at future times at least going off of the physical interpretation of the heat equation as "diffusing heat" (or what if, like, the future affects the past, man???). The proof of this is similar to the proof of the mean value property for harmonic functions where one considers the RHS above as a function in r and shows that it is constant. Although its not terribly long, it is trickier than for the harmonic case and since it won't be needed in the sequel we leave the reader to look it up themselves in, for instance, [5] – its handy to know about of course. It also turns out that as in the case for harmonic functions the converse statement is also true.

The reason in particular we won't need it is because the rest of the results we wish to show: uniqueness, smoothness, and derivative estimates all follow from the maximum principle which, as in the elliptic case, can be proved without it by similar means. To state the maximum principle for caloric functions (that is, solutions to the heat equation) though first we need to set some terminology, where the last item will be used when discussing some of its consequences:

Definition 14.1. *(special sets in spacetime)*

1. *The parabolic cylinder of a (open, bounded) domain $U \subset \mathbb{R}^n$ is the set $U \times (0, T] \subset \mathbb{R}^n \times \mathbb{R}$.*
2. *The parabolic boundary of U_T is the set $\Gamma_T = \overline{U}_T \setminus U_T$*
3. *Reminisicent of the defintion of heat ball, we denote $C(x, t; r) = \{(y, s) \mid |x - y| \leq r, t - r^2 \leq s \leq t\}$*

Of course in the above, $T > 0$ and we'll also take it to be finite unless stated otherwise. Notice that from the definitions that the "top" of the box given by $\partial\overline{U}_T$ is not included in Γ_T since in defining U_T we crossed with the clopen interval $(0, T]$. Now we are ready to state the maximum principle for the heat equation:

Theorem 14.2. *Assume that* $u \in C^2_1(U_T) \cap C(\overline{U}_T)$ *solves the heat equation in* U_T. *Then*

$$\max_{\overline{U}_T} u = \max_{\Gamma_T} u \tag{14.3}$$

Proof: Since U is bounded and T is finite by convention $\overline{U}_T$ is compact, and so by compactness that the supremum of u is obtained at some point in $\overline{U}_T$. Suppose for the sake of contradiction that $\max_{\overline{U}_T} u$ is obtained in the set $\overline{U}_T \setminus \Gamma_T$, say at the spacetime point (x_0, s), and is strictly greater than the values of u at the boundary. First we suppose that $s < T$. Now at this point, considering the space and time directions separately, we see by the derivative tests we must have $\Delta u \leq 0$ and $u_t = 0$ which is nearly good enough for a contradiction unless of course $\Delta u = 0$ there: if $\Delta u < 0$ then $(\partial_t - \Delta)u > 0$ at (x_0, s) giving a contradiction. Supposing then we are indeed in the edge case, we consider the function $v = u - \epsilon t$. If ϵ is sufficiently small v will still have an interior maximum, so that at this point $v_t = 0$. On the other hand since u satisfies the heat equation v satisfies $(\partial_t - \Delta)v = -\epsilon$ so that $\Delta v > \epsilon > 0$, giving a contradiction to the second derivative test.

Now, if $s = T$ then we see we must at least have $u_t \geq 0$, with the $u_t = 0$ case being the same as above so suppose $u_t > 0$ at (x_0, s). In this case the heat equation says that $\Delta u > 0$ as well, giving a contradiction again by the second derivative test. □

This can also be shown using the mean value property and in fact using it one can also see the strong maximum principle for the heat equation holds, but this fact will not be needed below (again, see [5]). In fact its not needed for the proof of the strong maximum principle either, which is something that holds pretty generally for parabolic PDE. A proof of the strong maximum principle using the parabolic harnack inequality will be given below for solutions on $\mathbb{R}^n$ that don't grow too rapidly, using a proof that generalizes easily. Some comments about its generalizations: as with the elliptic case this proof generalizes easily to more general parabolic scalar PDE (roughly speaking, adding more terms onto the heat equation). Along similar lines to the proof above one can often bound solutions to parabolic PDE in terms of related ODE (basically, justifying throwing out terms like the Laplacian) which can be readily solved or at least understood better. Important in the study of curvature flows like the mean curvature flow $\frac{dX}{dt} = \vec{H}$ or the Ricci flow $\frac{dg_{ij}}{dt} = -\frac{1}{2}\operatorname{Ric}_{ij}$ there are also maximum prinicples for matrices/tensors that satisfy parabolic PDE. These are important because outside of some notable exceptions flows aren't so useful unless some positivity of the curvature tensor or second fundamental form is assumed from the start, which will typically be shown to be preserved by a maximum principle.

15. A bedtime story about the mean curvature flow

Relatedly for the mean curvature flow the maximum principle implies that if two mean curvature flows of hypersurfaces are initially disjoint, they stay disjoint under the flow at least when one of them is compact – this is called the avoidance principle for the mean curvature flow. This can be used for instance to show rigorously "neckpinch" singularities can occur under the flow by considering as initial data a dumbbell shaped sphere M formed by attaching two large spheres (the "bells") with a segment of a skinny cylinder (the "neck"). It actually wasn't rigorously known for a period if these actually occured for comapct initial data! We wish to see the flow M_t of M will form a singularity along the neck, which is to say the neck region of M stays roughly cylinderical and the radius of the approximating cylinder tends to zero in finite time. Nestled inside M on either side of the attaching cylinder we may place large round spheres, which by the avoidance principle keep the bells of the flow of the dumbbell from shrinking quickly because we know the flow of these explicitly for symmetry reasons: essentially this keeps the cylinderical part of M looking cylinderical for at least some period along the flow; the problem is now that it might take too long (if ever) for the neck to squash down as indicated above. To see that this occurs, one can use as a barrier Angenent's shrinking donut, which is a torus which shrinks homethetically under the flow. If the neck of the initial data is thin enough it can be threaded through the hole of the donut and, since we know how fast the donut crushes down to a point, we get an upper bound on the time it takes for the radius of the neck to have to tend to zero again by the avoidance prinicple. See [2] for the construction of the donut.

It turns out if the initial data is convex (so looks like the sphere, or an ellipsoid) neckpinches *can't* occur, as shown in [9], and that the flow will always shrink to round circles in the curve shortening flow case [7], which is the analogue of the mean curvature flow for curves on a surface. That the curve shortening flow is relatively well behaved gives a number of interesting geometric and topological corollaries, such as the isoperimetric inequality, the 3 geodesics theorem, or a statement about the diffeomorphism group of the 2–sphere. That neckpinches, or even more complicated singularites (see for instance the construction by Copenhagen's very own N.M. Møller [10]) can occur messes up some of the potential applications of the mean curvature flow to related problems in geometry and topology though. One way to rule out the presence of more complicated singularities, besides for instance assumptions concerning the curvature of the initial data, is through the notion of entropy introduced by Colding and Minicozzi in [4] – Toby Colding is a very prominent Danish mathematician and this concept has been heavily employed in recent research in the area. If you like textbooks, a nice comprehensive book on the mean curavture flow and other extrinsic geometric flows is [1].

16. Consequences of the maximum principle for the heat equation

Getting back to actual course material, we next discuss two uniqueness theorems one for the heat equation in bounded domains and one when the initial domain is all of $\mathbb{R}^n$. The proof of the first statement is immediate:

Theorem 16.1. *Let $g \in C(\Gamma_T), f \in C(U_T)$ where U is bounded. Then there exists at most one solution $u \in C^2_1(U_T) \cap C(\overline{U}_T)$ of the initial/boundary–value problem*

$$\begin{cases} u_t - \Delta u = f \text{ in } U_T \\ u = g \text{ on } \Gamma_T \end{cases} \tag{16.1}$$

Proof: Suppose there were two solutions u_1, u_2. Then $u_1 - u_2$ solves the heat equation and is zero along Γ_T so that it is nonpositive by the maximum principle. Similarly by considering $-(u_1 - u_2)$ it is nonnegative so must be zero giving they are equal. □

Now in the proof of the maximum principle (theorem 14.2) above we used the domain U was bounded in knowing there was the point (x_0, s) where the supremum of u was actually achieved; if we were to follow the scheme of the proof of uniqueness above when $U = \mathbb{R}^n$ we would need a noncompact maximum principle however. Obviously these tend to be more subtle than the compact case – they can be useful though because there are many cases where one wishes to consider a PDE on a noncompact domain even when the problem one was first interested in was on a compact one! (in fact we'll see such an instance shortly) Below we give a noncompact maximum principle under a growth bound; interjecting some more jargon when the domain U is all of $\mathbb{R}^n$ the initial value problem is called the Cauchy problem:

Theorem 16.2. *Suppose* $u \in C_1^2(\mathbb{R}^n \times (0,T]) \cap C(\mathbb{R}^n \times [0,T])$ *solves*

$$\begin{cases} u_t - \Delta u = 0 \ \textit{in } \mathbb{R}^n \times (0,T) \\ u = g \ \textit{on } \mathbb{R}^n \times \{t=0\} \end{cases} \tag{16.2}$$

and satisfies the growth estimate

$$|u(x,t)| \leq Ae^{a|x|^2} \tag{16.3}$$

for constants $A, a > 0$*. Then*

$$\sup_{\mathbb{R}^n \times (0,T)} u = \sup_{\mathbb{R}^n} g \tag{16.4}$$

Proof: The growth rate assumption we'll soon see is related to the growth of the heat kernel. First we suppose that $4aT < 1$, so that there is some $\epsilon > 0$ for which $4a(T+\epsilon) < 1$. Fixing $\mu > 0$ we define the function

$$v(x,t) = u(x,t) - \mu \frac{1}{(T+\epsilon-t)^{n/2}} e^{\frac{|x|^2}{4(T+\epsilon-t)}} \tag{16.5}$$

Which we see is u shifted down by some multiple of the heat kernel time shifted by $T+\epsilon$. Because both u and the heat kernel solve the heat equation we have $v_t - \Delta v = 0$ in $\mathbb{R}^n \times (0,T]$. Then for some $r > 0$ considering the bounded domain $U = B(0,r)$ we have by the maximum principle for bounded domains that

$$\max_{\overline{U}_T} v = \max_{\Gamma_T} v \tag{16.6}$$

Now, we note since the heat kernel is positive and we are subtracting it off u that $v(x,0) < g(x)$. If the maximum of v in $\overline{U}_T$ is achieved on the $t=0$ slice for arbitrarily small μ and sufficiently large r it will give us what we want then by taking these quantities to 0 and ∞ respectively, but the danger is that the maximum might instead occur along the sides of the parabolic cylinder. Now for $x \in \Gamma_T$ (the upshot being that we can replace $|x|^2$ with r^2) the growth bound gives

$$\begin{aligned} v(x,t) &= u(x,t) - \mu \frac{1}{(T+\epsilon-t)^{n/2}} e^{\frac{|x|^2}{4(T+\epsilon-t)}} \\ &\leq Ae^{a|x|^2} - \mu \frac{1}{(T+\epsilon-t)^{n/2}} e^{\frac{r^2}{4(T+\epsilon-t)}} \\ &\leq Ae^{ar^2} - \mu \frac{1}{(T+\epsilon)^{n/2}} e^{\frac{r^2}{4(T+\epsilon)}} \end{aligned} \tag{16.7}$$

The last line using that $t \in [0,T]$ so in particular is positive. By the assumption we made in the start, that $4a(T+\epsilon) < 1$, we have $\frac{1}{4(T+\epsilon)} = a + \gamma$ for some $\gamma > 0$. Thus

$$v(x,t) \le Ae^{ar^2} - \mu(4a+4\gamma)^{n/2}e^{(a+\gamma)r^2} = e^{ar^2}(A - \mu(4a+4\gamma)^{n/2}e^{\gamma r^2}) \tag{16.8}$$

The point of this is that the second term in parentheses beats the first as $r \to \infty$ because of the extra exponential term in r and is why it was helpful to subtract off the heat kernel term. In particular for r large enough $(A - \mu(4a+4\gamma)^{n/2}e^{\gamma r^2} < -2A$; on the other hand the growth bound applied for $t = 0$ gives $g(x) > -2Ae^{ar^2}$ for $x \in \overline{U}$ (the $t = 0$ slice of Γ_T) so that indeed the maximum of v occurs not just on Γ_T but on $\overline{U}$, where v is less than $g(x)$. As metioned already taking $\mu \to 0$ gives the claim for $4aT < 1$. For bigger T the result can be applied on $[0,T]$ cut up into suitably small subintervals. □

There are other noncompact maximum principles out there, and just some names that are handy to remember (at least to have buried in the subconcious, in case you are in desperate need someday) are the Omori–Yau maximum principle and the Ecker–Huisken maximum principle. As a corollary of this we have uniqueness in the noncompact case as well:

Theorem 16.3. *Let $g \in C(\mathbb{R}^n), f \in C(\mathbb{R}^n \times [0,T])$. Then there exists at most one solution $u \in C_1^2(U_T) \cap C(\overline{U}_T)$ of the initial/boundary–value problem*

$$\begin{cases} u_t - \Delta u = f \text{ in } \mathbb{R}^n \times (0,T) \\ u = g \text{ on } \mathbb{R}^n \times \{t = 0\} \end{cases} \tag{16.9}$$

satisfying the growth estimate $|u(x,t)| \le Ae^{a|x|^2}$ for constants $A, a > 0$.

Now, it turns out that by power series methods that without the growth assumption the theorem above is actually false: Tychonoff showed there exists infinitely many solutions to the Cauchy problem with $g = 0$ (of course, the $u = 0$ solution is the obvious one). These are sometimes refered to as "nonphysical solutions" which is reasonable because, since $g = 0$ it would be as though a long (well, infintely long) metal bar at uniform temperature suddenly becomes hotter and colder at some places – one way to look at this perhaps is that, combined with infinite speed of propogation, that the heat equation isn't a perfect model for heat! Not a big surprise necessarily, its a very simple equation. Next we proceed to discuss the regularity of solutions to the heat equation:

Theorem 16.4. *Suppose $u \in C_1^2(U_T)$ is caloric/solves the heat equation in U_T. Then $u \in C^\infty(U_T)$.*

Proof: Note that as for the corresponding theorem on elliptic functions no claim on regularity along the boundary is made. Now fix $(x_0, t_0) \in U_T$ and, recalling the defintions in 14.1, choose r small enough so that $C = C(x_0, t_0, r) \subset U_T$. We also define for later use the cylinders $C' = C(x_0, t_0, \frac{3}{4}r)$ and $C'' = C(x_0, t_0, \frac{1}{2}r)$. The idea to proceed bares some similarity to the proof in the harmonic case although instead of using the mean value theorem (which we didn't prove for the heat equation anyway) we will use uniqueness of the heat equation along with the heat kernel.

Because we wish to employ the representation formula, we start reminiscent of the harmonic case by considering the mollification $u_\epsilon = \eta_\epsilon * u$ of u. Because here $u \in C_1^2(U_T)$ one can see from the defintion of convolution and pulling derivatives past integrals that actually u_ϵ solves the heat equation too. Then, we consider a smooth bump function ζ which is equal to one on C' and is zero near the boundary of C, which we then extend to be zero on all of $(\mathbb{R}^n \times [0, t_0]) \setminus C$. We multiply these two functions to get a smooth function $v = \zeta(x,t)u(x,t)$ defined on $\mathbb{R}^n \times [0, t_0]$. Now, indeed v does not solve the heat equation but the point roughly is that the discprepancy is smooth, compactly supported, and so can be plugged into the representation formula with the heat kernel. An easy calculation gives:

$$v_t - \Delta v = \zeta_t u - 2D\zeta \cdot Du - u\Delta\zeta \tag{16.10}$$

denoting the RHS above by $\tilde{f}$, consider the function

$$\tilde{v}(x,t) = \int_0^t \int_{\mathbb{R}^n} \Phi(x-y, t-s)\tilde{f}(y,s)dyds \tag{16.11}$$

Now by Duhamel's principle this solves

$$\begin{cases} (u_t - \Delta u)(x,t) = f(x,t) \text{ in } \mathbb{R}^n \times (0,\infty) \\ u = 0 \text{ on } \mathbb{R}^n \times \{0\} \end{cases} \tag{16.12}$$

Of course, v also solves this equation so because v and $\tilde{v}$ both are bounded (and hence exponentially bounded) the noncompact uniqueness theorem above says they are equal, which is good because we've got v in terms of convolution with the heat kernel (apriori this wasn't necessarily the case). Using that $\zeta = 0$ away from the cylinder C we may integrate by parts, using the formula above for $\tilde{f}$, to get that

$$v(x,t) = \int\int_C [\Phi(x-y, t-s)(\zeta_s(y,s) + \Delta\zeta(y,s)) + 2D_y\Phi(x-y, t-s)\cdot D\zeta(y,s)]u_\epsilon(y,s)dyds \tag{16.13}$$

Now, if $(x,t) \in C''$, where ζ is identically equal to one, the LHS is just $u_\epsilon(x,t)$. Denoting by $K(x,t,y,s) = \Phi(x-y,t-s)(\zeta_s(y,s) + \Delta\zeta(y,s)) + 2D_y\Phi(x-y,t-s) \cdot D\zeta(y,s)$ we have then for $(x,t) \in C''$ that

$$u_\epsilon(x,t) = \int\int_C K(x,t,y,s)u_\epsilon(y,s)dyds \tag{16.14}$$

Because $u_\epsilon \to u$ uniformly as $\epsilon \to 0$, we get the same representation formula for u. Because the support of ζ is contained in the set where Φ is smooth one can see K is smooth and also compactly supported. As a consequence we can pass the spatial derivatives on the LHS through the integral which fall onto K so that u is smooth as claimed. □

Note that we used mollification of u above as in the harmonic case, but instead of showing $u_\epsilon = u$ via the mean value property we mollified to justify a representation formula which one could then see was true for the original function. Later we recall we derived a represenation formula for harmonic functions using Green's formulas in section 9. This representation formula similar to as indicated (although not carried out in detail) for harmonic functions can be used to give more explict gradient estimates:

Theorem 16.5. *There exists for each pair of integers $k, \ell = 0, 1, \ldots$ a constant $C_{k,\ell}$ such that*

$$\max_{C(x,t;r/2)} |D_x^k D_t^\ell u| \leq \frac{C_{k\ell}}{r^{k+2\ell+n+2}} ||u||_{L^1(C(x,t;r))} \tag{16.15}$$

for all cylinders $C(x,t;r/2) \subset C(x,t;r) \subset U_T$ and all caloric functions u in U_T.

Proof: From the representation above and that K is a fixed function (for a given r at least) its clear there should be some constants which fullfill the bound above. An important point is that we also know how the statement scales in r and this is a consequence of how the domain(s) C scale whose definition in turn is related to the scaling properties of the heat equation. By translation, we may suppose that $(x,t) = (0,0)$ (this doesn't affect the derivatives of u), and we start by supposing that $C(1) = C(0,0;1) \subset U_T$. Defining analogously $C(1/2)$ then as in the proof above for $(x,t) \in C(1/2)$ we have

$$u(x,t) = \int\int_C K(x,t,y,s)u(y,s)dyds \tag{16.16}$$

Where K is some smooth function with compact support – all we need to know about it is that its fixed throughout. Bounding $|D_x^k D_t^\ell K| \leq C_{k\ell}$ by some number $C_{k\ell}$ we see then we have

$$\max_{C(1/2)} |D_x^k D_t^\ell u| \le C_{k\ell} ||u||_{L^1(C(1))} \tag{16.17}$$

Now we consider a general cylinder $C(r) \subset U_T$ and proceed by a scaling argument (I admitedly belabor this, but hopefully in a clarifying way). We recall that if $u(x,t)$ satisfies the heat equation then so does $v(x,t) = u(rx, r^2 t)$. On the other hand, this rescaling (sending $(x,t) \to (rx, r^2 t)$) takes $C(1), C(1/2)$ to $C(r), C(r/2)$ respectively. So, if we want to prove an inequality for u on the sets $C(r), C(r/2)$ respectively we should consider v on the sets $C(1), C(1/2)$. Because v satisfies the heat equation, we have from above:

$$\max_{(x,t)\in C(1/2)} |D_x^k D_t^\ell v(x,t)| \le C_{k\ell} ||v||_{L^1(C(1))} \tag{16.18}$$

On the other hand by the chain rule $D_x^k D_t^\ell v(x,t) = D_x^k D_t^\ell u(rx, r^2 t) = r^{2\ell+k} D_y^k D_s^\ell u(y,s)$ evaluated at the point $(y,s) = (rx, r^2 t)$. Note that from above that because $(x,t) \in C(1/2)$ that $(rx, r^2 t) \in C(r/2)$, and this map is a bijection between these sets, so that

$$\max_{(x,t)\in C(1/2)} |D_x^k D_t^\ell v(x,t)| = \max_{(y,s)\in C(r/2)} |r^{2\ell+k} D_y^k D_s^\ell u(y,s)| = r^{2\ell+k} \max_{(x,t)\in C(r/2)} |D_x^k D_t^\ell u(x,t)| \tag{16.19}$$

where in the last line we just changed the notation back to what we were using and pulled the r factor out. Because these are equalities the last line is still bounded by $C_{k\ell}||v||_{L^1(C(1))}$. But $||v||_{L^1(C(1))} = \frac{1}{r^{n+2}}||u||_{L^1(C(r))}$ by the regular change of variables rule. Hence

$$\max_{C(r/2)} |D_x^k D_t^\ell u| \le \frac{C_{k\ell}}{r^{k+2\ell+n+2}} ||u||_{L^1(C(r))} \tag{16.20}$$

which, since we translated (x,t) to $(0,0)$, gives us what we want. □

Now, we notice that in contrast to the corresponding estimates for harmonic functions that we don't have explicit estimates for the constants $C_{k\ell}$ above, although if we were more explicit with our choice of bump function we could readily do so. It turns out that there is a constant C for which $C_{k0} < C^k k!$ while $C_{0\ell} < C^k (k!)^2$ suggesting it might not be the case (and indeed there are examples) that caloric functions will always be analytic in time. For a fixed t they will be analytic in space coordinates however. Using these estimates one can, as in the harmonic case, also prove a Liousville type theorem for caloric functions but we see here that the domains $C(r)$ stretch far backwards in time besides just being large spatially for r large, so the statement is only for ancient solutions (i.e. those defined on $(-\infty, T]$ for some

$T > -\infty$). Incidentally ancient solutions are important in the singularity analysis of the mean curvature flow, and indeed there are no closed ancient solutions to the mean curvature flow contained for all times in a fixed bounded set by the comparison principle.

Another method to gain gradient estimates, bounded above in terms of the L^∞ norm of u is possible using just the maximum principle – it's more robust from the perspective that it doesn't use the representation formula for the heat equation. Since I got the itch to discuss this let's state and sketch it for the first derivatives:

Theorem 16.6. *Suppose that u is a caloric function in U_T and $C(x,t;r) \subset U_T$. Then*

$$\max_{C(x,t;r/2)} |Du| \leq \frac{C(n)}{r} ||u||_{L^\infty(C(x,t;r))} \tag{16.21}$$

Proof: (sketch) Let ζ be a smooth bump function as before which is equal to 1 on $C(x,t;r/2)$ and zero near the boundary of $C(x,t;r)$. Then we consider the function $v = \zeta^2|Du|^2 + Au^2$ where A is a constant to be determined. By and some easy estimating one can see that $v_t - \Delta v \leq 0$ if A is large enough depending on ζ so that the maximum principle, after checking it applies for such v where only the inequality in the heat equation holds, implies the maximum of v is attained along the parabolic boundary of $C(x,t;r)$. Since ζ vanishes along the parabolic boundary this implies that v is bounded above by the L^∞ norm of u on $C(x,t;r)$. Since $\zeta = 1$ on $C(x,t;r/2)$ the claim follows. □

Higher derivative estimates also follow. Of course, although its nothing really new it follows from this that the analogous statement is true for harmonic functions as well. Along similar lines, of plugging in a clever function into the heat equation and using the maximum principle, to wrap up this section we prove the parabolic harnack inequality. We start with a variant of the so–called differential Harnack inequality of the famous geometers Li and Yau [15] (now we are in the 1980s!), proceeding basically as in chapter 2 of [1]:

Theorem 16.7. *Let $u \in C^\infty(\mathbb{R}^n \times (0,T)) \cap C(\mathbb{R}^n \times [0,T))$ be a positive caloric function such that $u \leq e^{Ae^{a|x|^2}}$ for positive constants A, a. Then*

$$\Delta \log u + \frac{n}{2t} \geq 0 \tag{16.22}$$

Proof: This is mainly a computation so will just be left as a sketch: on may calculate $(\partial_t - \Delta)$ of $2t\Delta \log u + n$ is nonnegative, so by the growth bound on u the noncompact maxmimum principle with some modifications can be applied. Since this quantity is initially positive then it stays positve for $t > 0$, and dividing through by t gives the claim. □

Of course, in the above the choice of auxillary function was inspired – one suggestion that it is special is that its zero when u is the fundamental solution. Differential Harnack inequalities are already useful in curvature flows because the strong maximum principle can be used with them to prove some rigidity statements – surely this is covered in [1] but a nice source covering a relatively simple case is chapter 4 of [16]. Moving along, as a consequence it gives the following Harnack inequality for the heat equation:

Theorem 16.8. *Let $u \in C^\infty(\mathbb{R}^n \times (0,T]) \cap C(\mathbb{R}^n \times [0,T])$ be a positive caloric function such that $|u| \le e^{Ae^{a|x|^2}}$ for positive constants A, a. Then*

$$\frac{u(x_2,t_2)}{u(x_1,t_1)} \ge (\frac{t_2}{t_1})^{-n/2} e^{-\frac{|x_2-x_1|^2}{4(t_2-t_1)}} \tag{16.23}$$

for all $x_1, x_2 \in \mathbb{R}^n$ and $t_1 < t_2$ in $(0,T)$.

Proof: As the adjective "differential" in the previous statement might suggest, we integrate it to get this result. More precisely, let $\gamma(t) : [t_1, t_2] \to \mathbb{R}^n$ be a path from x_1 to x_2. By the chain rule, we calculate that $\frac{d}{dt} \log(u(\gamma(t),t)) = \partial_t \log u + \frac{D_{\gamma'} u}{u}$. Now, since u solves the heat equation we can see that $\partial_t \log u = \Delta \log u + \frac{|\nabla u|^2}{u^2}$. Because $\Delta \log u + \frac{n}{2t} \ge 0$, $\Delta \log u \ge -\frac{n}{2t}$. Putting this all together, using the Cauchy–Schwarz inequality, and throwing out the positve term gives:

$$\begin{aligned}\frac{d}{dt}\log(u(\gamma(t),t)) &= \partial_t \log u + \frac{D_{\gamma'} u}{u} \\ &= \Delta \log u + \frac{|\nabla u|^2}{u^2} - \frac{|Du|}{u}|\gamma'| \\ &\ge -\frac{n}{2t} + \frac{|\nabla u|^2}{u^2} - \frac{|Du|}{u}|\gamma'| \\ &\ge -\frac{n}{2t} - \frac{|Du|}{u}|\gamma'|\end{aligned} \tag{16.24}$$

Now we integrate from t_1 to t_2 to find

$$\log \frac{u(x_2,t_2)}{u(x_1,t_1)} \ge -\frac{n}{2}\log\frac{t_2}{t_1} - \frac{1}{4}\int_{t_1}^{t_2} |\gamma'(t)|^2 dt \tag{16.25}$$

We can pick γ to be whatever curve between x_1 and x_2 we want defined on the interval $[t_1.t_2]$, so we take it be the straight line parameterized by $\gamma(t) = x_1 + \frac{t-t_1}{t_2-t_1}(x_2 - x_1)$

so that $|\gamma'(t)|^2 = (\frac{|x_2-x_1|}{t_2-t_1})^2$, so that $\frac{1}{4}\int_{t_1}^{t_2} |\gamma'(t)|^2 dt = \frac{|x_2-x_1|^2}{4(t_2-t_1)}$. Exponentiating then gives the claim. □

Along more or less similar lines one can prove a parabolic Harnack inequality for general parabolic PDE (see chapter 7 of [5]), and using the Harnack inequalty one can proceed to prove the strong maximum principle – note that the argument above only used the regular maximum principle. Of course, it holds on more general domains than $\mathbb{R}^n$ but its phrased this way with the Harnack inequality above in mind:

Theorem 16.9. *Assume that $u \in C_1^2(\mathbb{R}^n \times (0,T)) \cap C(\mathbb{R}^n \times [0,T))$ solves the heat equation in U_T. Then if $\sup\limits_{\mathbb{R}^n \times [0,T)} u$ is attained in $(\mathbb{R}^n \times (0,T))$ and $|u| \leq e^{Ae^{a|x|^2}}$ for positive constants A, a then u is constant.*

Proof: Suppose the supremum is attained at the point $(x_2, t_2) \in \mathbb{R}^n \times (0,T)$ and denote it by M (since its actually attained at a point, its finite). Then $M + \epsilon - u$, where $\epsilon > 0$, is a positive solution to the heat equation for which the Harnack inequality applies. Letting $\epsilon \to 0$ gives that for $t < t_2$ that $u(x,t) = M$ identically. The uniqueness statement above then implies its true for $t \geq t_2$ as well. □

17. Energy methods, backwards uniqueness, and illposedness

In section 11 we introduced the concept of energy functional. In this section we return to energy methods to show some uniqueness results, including the well–known backwards uniqueness property for the heat equation, and discuss an example of an illposed PDE. First we give a new proof of uniqueness for solutions to the heat equation over bounded domains U with C^1 boundary (for integration by parts):

Theorem 17.1. *There exists at most one solution $u \in C_1^2(\overline{U}_T)$ of the problem:*

$$\begin{cases} u_t - \Delta u = f \ \textit{in} \ U_T \\ u = g \ \textit{on} \ \Gamma_T \end{cases} \tag{17.1}$$

Proof: As in the proof by the maximum principle we suppose there are two solutions u_1, u_2. Then their difference solves the heat equation and is zero along Γ_T. Denoting their difference $w = u_1 - u_2$, we consider the following energy quantity (of course, different from the Dirichlet energy):

$$e(t) = \int_U w^2(x,t)dx \tag{17.2}$$

Then the time derivative of $e'(t)$ is equal to the following, using next that w solves the heat equation and then integration by parts:

$$\begin{aligned} e'(t) &= 2\int_U ww_t dx \\ &= 2\int_U w\Delta w dx \\ &= -2\int_U |Dw|^2 dx \leq 0 \end{aligned} \tag{17.3}$$

Because $e(t)$ is intially zero it must be zero for all $t \geq 0$, implying $u_1 = u_2$. □
The boundedness of U I suppose is used when integrating by parts: if a solution is assumed to decay suitably as one considers points further and further from the origin I suppose it could work with some minor variations in the case U is unbounded (but not equal to $\mathbb{R}^n$). Now we discuss backwards uniqueness to the heat equation. The remarkable thing about the statement is that we are not supposing that u_1 and u_2 agree everywhere on Γ_T (only on the sides, not the bottom) so the maximum principle does not apply:

Theorem 17.2. *Suppose $u_1, u_2 \in C^2(\overline{U}_T)$ are two solutions to the heat equation that agree along $\partial U \times [0, T]$ and agree at time $t = T$. Then $u_1 = u_2$ everywhere within U_T.*

Proof: As in the previous proof we consider $w = u_1 - u_2$ and the energy defined by $e(t) = \int_U w^2(x,t)dx$. As before, $e'(t) = -2\int_U |Dw|^2 dx$. Now, we calculate the second derivative of e, $(e')'$:

$$\begin{aligned} e''(t) &= -4\int_U Dw \cdot Dw_t dx \\ &= 4\int_U \Delta w w_t dx \\ &= 4\int_U (\Delta w)^2 dx \end{aligned} \tag{17.4}$$

Where in the second and third equalities we used integration by parts and that w is caloric respectively. Considering that in the proof above we found $e'(t) = 2\int_U w\Delta w dx$ one may notice that e, e', and e'' can be related by Cauchy–Schwarz (where here the inner product is $\langle f, g\rangle = \int fg$) to get that

$$(e'(t))^2 = 4(\int_U w\Delta w dx)^2 \leq (\int_U w^2 dx)(4\int_U (\Delta w)^2 dx) = e(t)e''(t) \tag{17.5}$$

This is an ODE inequality which we hope to exploit, by showing it implies $e(t) = 0$ if $e(T) = 0$ for $T > t$ – this is should inspire confidence because its a fairly simple ODE, and we understand ODE pretty well. Rearranging it formally we see we get $\frac{e'(t)}{e(t)} \leq \frac{e''(t)}{e'(t)}$. Considering for a positive function $g(t)$ that the derivative $\log g(t)$ is $\frac{g'(t)}{g(t)}$ we see both sides are derivatives of log of e and its derivative respectively, suggesting (although this form isn't what we'll use at the end of the day) a good quantity to start looking at is $\log e(t)$ – this at least points us in the right direction. Being a bit more careful, $e(t)$ is clearly nonnegative and if its zero for all $0 \leq t \leq T$ there is nothing to show so otherwise, by continuity, there exists a subinterval $I = [t_1, t_2) \subset [0, T]$ for which $e(t) > 0$ on I and $e(t_2) = 0$. Writing $f(t) = \log e(t)$ defined on I then we calculate:

$$f'(t) = \frac{e'(t)}{e(t)}, \; f''(t) = \frac{e''(t)}{e(t)} - \frac{e'(t)^2}{e(t)^2} \tag{17.6}$$

Dividing the inequality $(e'(t))^2 \leq e(t)e''(t)$ through by $e(t)^2$, we see then that $f''(t) \geq 0$ so is convex. In particular for $0 < \tau < 1, t \in I$ we have:

$$f((1-\tau)t_1 + \tau t) \leq (1-\tau)f(t_1) + \tau f(t) \tag{17.7}$$

Since $f = \log e(t)$, exponentiating this gives:

$$0 \leq e((1-\tau)t_1 + \tau t) \leq e(t_1)^{1-\tau} e(t)^{\tau} \tag{17.8}$$

Now, this is true for $t \in I = [t_1, t_2)$ but by continuity of $e(t)$ we can take $t = t_2$ as well, implying since τ could run between 0 and 1 that $e(t) = 0$ for $t \in I$ giving a contradiction.

□

The backwards uniqueness of the heat equation raises the question: can we reverse the heat equation? That is, if we have prescribed data $f(x)$ at time $T > 0$ where f is a function on some domain U, can we find a solution $u(x, t)$ to the heat equation on $[0, T]$, with prescribed boundary values along $\partial U \times [0, T]$, for which $u(x, T) = f(x)$? The backwards uniqueness theorem says such a solution if it exists is unique. By the chain rule, one can see that equivalently the question is if one can always solve the following, naturally called the backwards heat equation:

$$\begin{cases} v_t + \Delta v = 0 \text{ in } U_T \\ v = g \text{ on } \Gamma_T \end{cases} \tag{17.9}$$

A solution v to the problem above would give a solution u to the original problem by setting $u(x, t) = v(x, T - t)$. Now, so far we've always been able to solve the

PDE we've been interested in but here is an example where there must be some time T_1 such that, for $T > T_1$, there is no solution (we've touched at similar issues with geometric flows). The problem, or at least a problem, is that the heat equation has a smoothing effect. With this in mind consider then a solution $u(x,t)$ to the heat equation over, say, the time interval $[0,1]$ so that $u(x,0) = f(x)$, where f is merely continuous: we have existence theorems for such inital data on $\mathbb{R}^n$ via the fundamental solution, and will discuss the existence of solutions for more general boundary data below. Let's just suppose to be specific that u is zero along $\partial U \times [0,1]$, which it turns out we may solve for.

Now, suppose then we plug in the function $u(x,1)$ for the intial value in the backwards heat equation, and extend by zero along the sides to get a function g. Suppose we were able to solve the backwards heat equation for, say, $T = 2$. Now backwards uniqueness for the heat equation implies that a solution v to this problem should be unique, and in particular that $v(x,1) = f(x)$ so is only continuous. But on the other hand, $v(x,1)$ should be smooth (even analytic) by the relationship between the backwards heat equation and regular one, giving a contradiction. Really, the situation is a bit more subtle than I'm letting on – for instance, it turns out one can see the backwards heat equation is well posed for some relatively small set of functions using the Fourier transform and the Fourier transform can be used to see more precisely where the issue lays. The point is that it isn't well posed within some natural function spaces of initial data.

18. Existence of the heat flow on general domains, and its long term fate

Having quenched our thirst for properties about solutions to the heat equation, we turn back to solving the heat equation on more general domains. We'll show that with appropriate boundary data the solution will converge in the long term to a solution to Dirichlet problem. The reason one should expect this to be true is because its the gradient flow of the Dirichlet energy; this is the justification I gave for investigating the heat equation in the first place in section 11 but of course its intrinsically interesting/useful for other reasons, for instance from the physics perspective. The main source for this section is chapter 4 of [11]; most of this will just be sketched to various degrees of completeness for the sake of time and knowledge of the full proofs won't be required for the exam but I think its still worthwhile. To be precise we will start by considering solving the general problem:

$$\begin{cases} u_t - \Delta u = \phi(x,t) \text{ in } U \times (0,\infty) \\ u(x,t) = g(x,t) \text{ on } \partial U \times (0,\infty) \\ u(x,0) = f(x) \text{ on } x \in U \end{cases} \tag{18.1}$$

(Here in keeping with [11] we broke the parabolic boundary of U_T into the "sides" and "bottom" respectively.) The method we will present, which by no means is the only method but is pretty elementary and ties in nicely with ideas we've already seen, is to show existence of solutions via an integral representation formula under some mild conditions, with our starting point being the fundamental solution. Actually matching the assumptions we had in the Perron method we will suppose that U is bounded with C^2 boundary. The problem with just using the heat kernel roughly speaking as for the Laplace equation are boundary terms and a corrector term will need to be added if one wants to find a heat kernel for a general domain. The first lemma will be used essentially to construct one, where $\Phi(x,y,t) = \Phi(x-y,t)$:

Lemma 18.1. *Where U is as above, let $\gamma \in C^0(\partial U \times [0,T])$. If we set v as*

$$v(x,t) = -\int_0^t \int_{\partial U} \frac{\partial \Phi}{\partial \nu_y}(x,y,s)\gamma(y,t-s)dyds \tag{18.2}$$

Then $v \in C^\infty(U \times [0,T])$, $v(x,0) = 0$ for $x \in U$, and for all $x_0 \in \partial U$, $t \in (0,T]$:

$$\lim_{x \to x_0} v(x,t) = \frac{\gamma(x_0,t)}{2} - \int_0^t \int_{\partial U} \frac{\partial \Phi}{\partial \nu_y}(x_0,y,s)\gamma(y,t-s)dyds \tag{18.3}$$

Proof: The smoothness of v follows by the smoothness of the heat kernel. Because there is a time integral in the definition of v, its clear that $v(x,0) = 0$ as well. The hardest part is the last identity in the statement, which requires showing $-\int_0^{t_0} \int_{\partial U \cap B(x_0,\delta)} \frac{\partial \Phi}{\partial \nu_y}$ limits to $1/2$ in the limit of $\delta, t_0 \to 0$. Since the boudary of U is C^2 for δ small enough it can be approximated suitably well by a straight line, for which the normal to it can be explicitly written. Then the problem comes down to calculating an integral which boils down eventually to the Gamma function; again the details can be found in [11]. □

With this in hand, now suppose we wanted to solve the problem

$$\begin{cases} u_t - \Delta u = 0 \text{ in } U \times (0, \infty) \\ u(x,t) = g(x,t) \text{ on } \partial U \times (0, \infty) \\ u(x,0) = 0 \text{ on } x \in U \end{cases} \tag{18.4}$$

A thing one can try is to set u to be

$$u(x,t) = -\int_0^t \int_{\partial U} \frac{\partial \Phi}{\partial \nu_y}(x,y,s)\gamma(y,t-s)dyds \tag{18.5}$$

For some appropriate function γ. We have u is at least smooth in $U \times (0,T)$ and it vanishes on $U \times \{0\}$ by the lemma. By the lemma for it to at least agree with g on the sides of the cylinder we need to have:

$$g(x_0,t) = \frac{\gamma(x_0,t)}{2} - \int_0^t \int_{\partial U} \frac{\partial \Phi}{\partial \nu_y}(x_0,y,s)\gamma(y,t-s)dyds \tag{18.6}$$

Which, rearranging slightly, is equivalent to finding γ so that

$$\gamma(x_0,t) = 2g(x_0,t) + 2\int_0^t \int_{\partial U} \frac{\partial \Phi}{\partial \nu_y}(x_0,y,s)\gamma(y,t-s)dyds \tag{18.7}$$

Or in other words we want to find a fixed point of the function $\gamma(x_0,t) \to 2g(x_0,t) + \int_0^t \int_{\partial U} \frac{\partial \Phi}{\partial \nu_y}(x_0,y,s)\gamma(y,t-s)dyds$ (valid at once for all $x_0 \in \partial U$). If we set $\gamma_0(x_0,t) = 2g(x_0,t)$ and $\gamma_n = 2g(x_0,t) + 2\int_0^t \int_{\partial U} \frac{\partial \Phi}{\partial \nu_y}(x_0,y,s)\gamma_{n-1}(y,t-s)dyds$, if this sequence converges to a function γ it will be a fixed point by the regular argument. Proceeding formally for a moment, using that $\gamma_0 = 2g(x_0,t)$ the fantastical fixed point can be written as:

$$\gamma(x_0,t) = 2g(x_0,t) + 2\int_0^t \int_{\partial U} \sum_{k=1}^{\infty} S_k(x_0,y,t-s)g(y,s)dyds \tag{18.8}$$

where $S_1 = 2\frac{\partial K}{\partial \nu_y}(x_0,y,t)$, $S_{k+1} = 2g(x_0,t) + 2\int_0^t \int_{\partial U} \frac{\partial \Phi}{\partial \nu_y}(x_0,y,s)S_k(x_0,y,t-s)dyds$ as can be verified by induction (γ_n of course is with just the first n terms in the series above). For such γ to actually exist the series in the representation above must converge, and showing this comes down to some estimation of the heat kernel which is technical as in the lemma with an induction argument. Inserting this γ back into how we define u we get and a time translation gives:

$$u(x,t) = -\int_0^t \int_{\partial U} \Sigma(x,y,t-s)g(y,s)dyds \tag{18.9}$$

Where for the record Σ is given by:

$$\Sigma(x,y,t) = 2\frac{\partial K}{\partial \nu_y}(x_0,y,t) + 2\int_0^t \int_{\partial U} \frac{\partial \Phi}{\partial \nu_z}(x,z,t-s)\sum_{k=1}^{\infty} S_k(z,y,s)dyds \qquad (18.10)$$

Because the convergence of the series above turns out to uniform on emay differentiate terrm by term through the integral to see that u satisfies the heat equation. By how γ was found, $u = g$ along ∂U so that in all 18.2 is solvable. For reference below we record this more officially as a theorem:

Theorem 18.2. *On a bounded C^2 domain U problem is uniquelly solvable for all continuous functions g and the solution u is given by the following, where Σ is as above:*

$$u(x,t) = -\int_0^t \int_{\partial U} \Sigma(x,y,t-s)g(y,s)dyds \qquad (18.11)$$

Now at this point one can see arguing essentially as in the beginning of section 10 that the general problem 18.1 is solvable. In particular, to solve the general problem 18.1 one can extend the functions ϕ, f continuously to all of $\mathbb{R}^n$ (supposing they are continuous up to the boundary) and use the heat kernel representation formula and Duhamel's principle from above to solve 18.1 with a function u_1 when restricting back to $U \times [0,T]$, all except the function u_1 probably won't agree with g along the sides of the cylinder. To deal with this we may subtract the restriction of u_1 off of g, solving 18.2 for it to get a function u_2, and then $u = u_1 + u_2$ will completely solve 18.1 when all the functions are sufficiently regular, although we won't actually need this full result in the following.

As promised, our final goal of this section is to solve the Dirichlet problem on a bounded C^2 domain, giving a proof independent of the Perron method. To continue towards that and to close up an assertion from above we officially introduce the heat kernel for a general domain U, by which we mean the following:

Definition 18.1. *Let $U \subset \mathbb{R}^n$ be a domain. A function $q(x,y,t)$ that is defined for $x, y \in \overline{U}$, $t > 0$ is called the heat kernel of U if:*

(1) $q_t - \Delta_x q = 0$ *for* $x, y \in U, t > 0$
(2) $q(x,y,t) = 0$ *for* $x \in \partial U$
(3) $\lim\limits_{t\to 0} \int_U q(x,y,t)f(x)dx = f(y)$ *for all* $y \in U$

for all $f \in C(U)$.

We argue essentially as above to see that U then has a heat kernel:

Theorem 18.3. *Any bounded domain $U \subset \mathbb{R}^n$ with C^2 boundary has a heat kernel $q(x, y, t)$. For x, t fixed q is C^1 in y on $\overline{U}$. Furthermore the heat kernel is positive in U for all $t > 0$.*

Proof: Where Φ as usual is the fundamental solution of the heat equation, in the sketch above we let $g(x, t) = -\Phi(x, y, t)$ restricted to ∂U to get for each y a solution $\mu(x, y, t)$. Then we put $q(x, y, t) = \Phi(x, y, t) + \mu(x, y, t)$. Clearly q satisfies items (1) and (2) and because Φ satisfies item (3) and $\mu(x, y, t) \to 0$ as t does (by the lemma essentially) q also satisfies item (3). The claim that q is positive follows from the strong maximum principle and that by item (3) q must be positive when t and $|x-y|$ are small: in fact knowing that q is nonnegative suffices for our purposes which is a consequence of the regular maximum principle. □

Of course, μ essentially plays the role of the corrector function in the discussion we gave above for Green's functions for the Laplacian, although instead of before where we found it by solving a related PDE in some special domains by using their symmetry we found μ essentially by a fixed point argument. The point of bringing this up for the long term convergence is that it turns out $\int_0^t \int_{\partial U} \Sigma(x, y, t-s) g(y, s) dy ds = \int_0^t \int_{\partial U} \frac{\partial q}{\partial \nu_y}(x, y, t-s) g(y, s) dy ds$. To see this, with μ as above we have by theorem 18.2:

$$\mu(x, y, t) = \int_0^t \int_{\partial U} \Sigma(x, z, t-s) \Phi(z, y, s) dz ds \tag{18.12}$$

and arguing similar to proof of lemma 18.1 we can see:

$$\frac{\partial \mu}{\partial \nu_y}(x, y, t) = \frac{\Sigma(x, y, t)}{2} + \int_0^t \int_{\partial U} \Sigma(x, z, t-s) \frac{\partial \Phi}{\partial \nu_y}(z, y, s) dz ds \tag{18.13}$$

The second term above in turn is just $-\frac{\partial \Phi}{\partial \nu_y}(x, y, t)$ by another use of theorem 18.2 for $x \in \partial U$. Rearranging then gives the claim. It will be used right below:

Theorem 18.4. *Let $U \subset \mathbb{R}^n$ be a bounded C^2 domain and let $g \in C(\partial U)$. Then there exists a function $u(x, t)$ solving the problem*

$$\begin{cases} u_t - \Delta u = 0 \text{ in } U \times (0, \infty) \\ u(x, t) = g(x) \text{ on } \partial U \times (0, \infty) \\ u(x, 0) = 0 \text{ on } x \in U \end{cases} \tag{18.14}$$

and $\lim_{t \to \infty} u(x, t)$ is a harmonic function in U equal to g along ∂U.

Proof: Such a function u exists as a special case of theorem 18.2. Since q is non-negative and zero along the boundary $\frac{\partial q}{\partial \nu_y}(x,y,t) \leq 0$. Since q vanishes along the boundary of $\partial U \times (0,\infty)$ as $\lim_{t\to\infty} q(x,y,t) = 0$ by a barrier argument (i.e. one can construct a subsolution to the heat equation that tends to zero as $t \to \infty$ which dominates the function we are interested in); this is a general fact for caloric functions which will probably be outlined in a homework problem. With all this in mind for any $t_2 > t_1$ we calcuate:

$$\begin{aligned}
|u(x,t_2) - u(x,t_1)| &= |\int_{t_1}^{t_2}\int_{\partial U} \frac{\partial q}{\partial \nu_y}(x,y,s)g(y)dyds| \\
&\leq \max_{\partial U}|g| \int_{t_1}^{t_2}\int_{\partial U} -\frac{\partial q}{\partial \nu_y}(x,y,s)dyds \\
&= -\max_{\partial U}|g| \int_{t_1}^{t_2}\int_{U} \Delta q(x,y,s)dyds \text{ (divergence theorem)} \\
&= -\max_{\partial U}|g| \int_{t_1}^{t_2}\int_{U} q_t(x,y,s)dyds \text{ (sol'n to heat equation)} \\
&= -\max_{\partial U}|g| \int_{U} q(x,y,t_2) - q(x,y,t_1)dy
\end{aligned} \tag{18.15}$$

And for t_1, t_2 sufficiently large the last line can be made as small as one wishes, implying that $u(x,t)$ converges uniformly to a function $u_\infty(x)$ as $t \to \infty$ which is equal to $g(x)$ along its boundary. Because derivatives commute u_t also solves the heat equation, and since g doesn't depend on time one can justify u_t is zero along ∂U so like q must tend to zero as $t \to \infty$. This implies by the heat equation that $\lim_{t\to\infty} \Delta u(x,t) = 0$. By the interior gradient estimates the convergence of $u(x,t)$ to $u_\infty(x)$ after possibly passing to a subsequence by Arzela–Ascoli can be taken to be in say C^3 norm, so that $\Delta u_\infty = 0$ showing that it solves the Dirichlet problem. □

As mentioned in [11], starting with nonzero initial data along U will also converge to the solution to the Dirichlet problem with only minor modifications. So, we used a PDE to solve a PDE, ain't that grand? This philosophy can be generalized in a number of interesting directions, such as to finding minimal surfaces via the mean curvature flow (this works best in the curve shortening case, see section 5 of [3]), or to provide a parabolic proof of the Hodge theorem which can then be pushed, with more work, to prove of the Gauss–Bonnet theorem [19].

19. The wave equation and d'Alembert's formula

The last model PDE on the list from section 2 is the wave equation. For time constraints we won't say as much about it as the first two equations, but it and equations like it (hyperbolic PDE) are very important in many fields of math and science (sound, optics, general relativity, etc). Also, the qualitiative behavior of its solutions can be quite a bit different from solutions to the Laplace and heat equations, which in some ways are quite a bit alike, and its good to have a real world concrete example that shows PDE can indeed exhibit a diverse array of phenomena. We can see this rather quickly from the solution to the $n = 1$ case (so, solutions to $u_{tt} - u_{xx} = 0$). More precisely we discuss next the solution to

$$\begin{cases} u_{tt} - u_{xx} = 0 \text{ in } \mathbb{R} \times (0, \infty) \\ u = g, u_t = h \text{ on } \mathbb{R} \times \{t = 0\} \end{cases} \tag{19.1}$$

Note that in comparison with n-th order ODE specifiying the u, u_t at a fixed time should be the "right" initial data to prescribe to get a unique solution, without overdetermining the problem. Our starting point is that by the equality of mixed partials we can factor the wave operator $\partial_{tt} - \partial_{xx}$ into $(\partial_t + \partial_x)(\partial_t - \partial_x)$. Indeed inserting a C^2 function u gives:

$$(\partial_t + \partial_x)(\partial_t - \partial_x)u = (\partial_t + \partial_x)(u_t - u_x) = u_{tt} - u_{xt} + u_{tx} - u_{xx} = u_{tt} - u_{xx} \tag{19.2}$$

The upshot of this is that this lets us solve the wave equation by iteratively solving easy first order PDE – this is clearly something we can't do to the heat equation, and for $n = 1$ the Laplace equation clearly has only linear solutions anyway (theres an interesting but long story about finding the "square root" of the Laplacian in higher dimensions related to quantum physics). Writing $v = u_t - u_x$, we see that we must have $v_t + v_x = 0$. This is an example of what is called a transport equation with constant coefficients, and these can be solved by an easy application of the method of characteristics which we now discuss (in a bit more explicitly geometric way then [5]). Now, if we consider the graph of v a tangent basis for it at a point (x, t) of v is given by $(1, 0, v_x)$ and $(0, 1, v_t)$. Thus the time normal vector for the graph up to scale is given by the vector $(v_x, v_t, -1)$.

On the other hand, because v solves the PDE $v_t + v_x = 0$, we have that the vector $(1, 1, 0)$ is perpendicular to $(v_x, v_t, -1)$ so that the integral curves of this vector, with initial data on the graph of v, will lay on the graph: these are called the characterisitic curves, the namesake of the method – in the general framework where instead of $(1, 1, 0)$ we might have a vector of general functions. Finding these amounts to solving a system of first order ODE to solve apparently called the Lagrange–Charpit equations. The ODE system in this case is $(x_1'(s), x_2'(s), x_3'(s)) = (1, 1, 0)$ so using the initial data $(y, 0, v(y, 0))$ we have

$$\ell(s) = (y + s, s, v(y, 0)) \tag{19.3}$$

all lay on the graph of v and varying y and s we see parameterize a surface which must agree with the graph. Setting $x = y + s$ and $t = s$ in particular the point $(x, t, v(x, t))$ on the graph of v must be the same as the point $(x, t, v(x - t, 0))$ or in other words $v(x, t) = a(x - t)$ for a function a, namely $a(x) = v(x, 0)$. Using the defintion of v then we must have:

$$u_t - u_x = a(x - t) \tag{19.4}$$

This is another transport equation for which we may apply the method of characteristics. The normal to the graph of u is as before $(u_x, u_t, -1)$, but now the equation says the vector $(-1, 1, a(x - t))$ is perpendicular to it. Finding the characterisitics then involves solving a little bit harder system of ODE starting at $(y, 0, u(y, 0))$. These are given by

$$\ell(s) = (y - s, s, u(y, 0) + \int_0^s a(y - 2r)dr) \tag{19.5}$$

Using the fundamental theorem of calculus and that $x = y - s, t = s$. Reasoning as above $u(x, t) = u(x + t, 0) + \int_0^t a(x + t - 2r)dr$. Changing variables to $z = x + t - 2r$ and writing $b(x) = u(x, 0)$ we get that:

$$u(x, t) = \frac{1}{2} \int_{x-t}^{x+t} a(z)dz + b(x + t) \tag{19.6}$$

Now we want to see what these a and b are – they should on principle be in terms of the initial conditions. Plugging in $t = 0$, we see that $b = g$, the zeroth order initial condition. To figure out what a is we recall $a(x) = v(x, 0)$ and v in turn is $u_t(x, t) - u_x(x, t)$. Evaluating at $(x, 0)$ gives $h - g'$, so that

$$\begin{aligned} u(x, t) &= \frac{1}{2} \int_{x-t}^{x+t} h(z) - g'(z)dz + g(x + t) \\ &= \frac{1}{2}(g(x + t) + g(x - t)) + \frac{1}{2} \int_{x-t}^{x+t} h(z)dz \end{aligned} \tag{19.7}$$

This formula is called d'Alembert's formula and was derived assuming that u solved the heat equation. The converse is also true:

Theorem 19.1. *Assume $g \in C^2(\mathbb{R})$, $h \in C^1(\mathbb{R})$, and u as defined by d'Alembert's formula 19.7. Then*

(1) $u \in C^2(\mathbb{R} \times [0, \infty))$,
(2) $u_{tt} - u_{xx} = 0$ *in* $\mathbb{R} \times (0, \infty)$,
(3) $\lim_{(x,t)\to(x_0,0)} u(x,t) = g(x_0)$, $\lim_{(x,t)\to(x_0,0)} u_t(x,t) = h(x_0)$ *for all* $x_0 \in \mathbb{R}$.

In particular (and this is especially clear when $h = 0$) we see that the regularity of solutions does not necessarily improve, at least for the 1-d wave equation. This can be crudely ascribed to that the formula above is more purely in terms of the intial data without convolution against a (mostly) smooth function. Of course this is in stark contrast to the heat equation, where the solution was smooth for positive times even if the initial data was not.

Before we move on for the sequel we want to extend d'Alembert's formula to find solutions to the wave equation just over the half line. Suppose we have a solution to the wave equation with the given initial data:

$$\begin{cases} u_{tt} - u_{xx} = 0 \text{ in } \mathbb{R}_+ \times (0, \infty) \\ u = g, u_t = h \text{ on } \mathbb{R}_+ \times \{t = 0\} \\ u = 0 \text{ on } \{x = 0\} \times (0, \infty) \end{cases} \tag{19.8}$$

Where as before $h \in C^1, u, g \in C^2$ and for technical reasons discussed shortly we'll also suppose that $g''(0) = 0$. Physically this models a vibrating string with one end fixed. Then by so–called odd reflection we craft functions $\widetilde{u}, \widetilde{g}, \widetilde{h}$ from u, g, h by setting

$$\widetilde{u}(x,t) = \begin{cases} u(x,t) \text{ for } x, t \geq 0 \\ -u(-x,t) \text{ for } x \leq 0, t \geq 0 \end{cases} \tag{19.9}$$

$$\widetilde{g}(x) = \begin{cases} g(x) \text{ for } x \geq 0 \\ -g(-x) \text{ for } x \leq 0 \end{cases} \tag{19.10}$$

$$\widetilde{h}(x) = \begin{cases} h(x) \text{ for } x \geq 0 \\ -h(-x) \text{ for } x \leq 0 \end{cases} \tag{19.11}$$

One might wonder why we picked odd reflection specifically to extend u, g, h – why not some other way? The (main) point of course is that $\widetilde{u}$ is C^2 and $\widetilde{u}_{tt}(x,t) - \widetilde{u}_{xx}(x,t) = 0$ in $\mathbb{R} \times (0, \infty)$. Away from $x = 0$ both points are easy to check by linearity of the wave equation, the chain rule, and that $(-1)^2 = 1$ (in particular there are an even number of derivatives taken). Since $u(0,t) = 0$ we see that the odd reflection of u is continuous and u_t is continuous with $u_t(0,t) = 0$. Because we specifically took the odd reflection the left and right limits of $\widetilde{u}_x(x,t)$ agree as $x \to 0$ so $\widetilde{u}_x(x,t)$ exists and is continuous on all of $\mathbb{R} \times (0,\infty)$ giving that $\widetilde{u}(x,t)$ is a C^1 function. Similarly the mixed partials $\widetilde{u}_{xt}, \widetilde{u}_{tx}$ exist and are continuous. Since $u = 0$ on $\{x = 0\} \times (0,\infty)$ we see that $u_{tt}(0,t) = \widetilde{u}_{tt}(0,t) = 0$, so that by the wave equation and u is C^2, approaching $x = 0$ from the right, that $u_{xx}(0,t) = 0$ as well. By the defintion of $\widetilde{u}$ for $x \leq 0$ we see that $\widetilde{u}_{xx}(x,t) = -u_{xx}(-x,t) = 0$ so that $\widetilde{u}_{xx}$ is continuous at $x = 0$ giving that $\widetilde{u}$ is C^2. Because $\widetilde{u}_{xx} = \widetilde{u}_{tt} = 0$ along the set $x = 0$ it also solves the wave equation.

By similar reasoning $\widetilde{g}, \widetilde{h}$ will be continuous since they must be zero at $x = 0$ since u is for $t > 0$ and u is continuous. They will be C^1 since we took the odd reflections, and $\widetilde{g}$ will be C^2 so long as $g''(0) = 0$. So, we can apply d'Alembert's formula to say

$$\widetilde{u}(x,t) = \frac{1}{2}(\widetilde{g}(x+t) + \widetilde{g}(x-t)) + \frac{1}{2}\int_{x-t}^{x+t} \widetilde{h}(z)dz \tag{19.12}$$

Unraveling the defintion of odd reflection this gives that

$$u(x,t) = \begin{cases} \frac{1}{2}(g(x+t) + g(x-t)) + \frac{1}{2}\int_{x-t}^{x+t} h(z)dz \text{ for } x \geq t \geq 0 \\ \frac{1}{2}(g(x+t) - g(t-x)) + \frac{1}{2}\int_{-x+t}^{x+t} h(z)dz \text{ for } 0 \leq x \leq t \end{cases} \tag{19.13}$$

20. Solving the wave equation by spherical means and Huygen's principle

d'Alembert's formula is only for the 1-d wave equation, but it turns out that one can more or less reduce the wave equation in general dimensions on $\mathbb{R}^n$ to the one dimensional case by what is called the method of spherical means. As usual there are other methods one could try to go about this, for instance by separation of variables, but this is an interesting way. Supposing in this section that $n, m \geq 2$ and $u \in C^m(\mathbb{R}^n \times [0,\infty))$ solves the problem

$$\begin{cases} u_{tt} - \Delta u = 0 \text{ in } \mathbb{R}^n \times (0,\infty) \\ u = g, u_t = h \text{ on } \mathbb{R}^n \times \{t = 0\} \end{cases} \tag{20.1}$$

where $g \in C^2(\mathbb{R})$, $h \in C^1(\mathbb{R})$ we will find formula for u in terms of g and h like before, and one can check that conversely a function defined in terms of that formula is a solution to the wave equation with the same initial data. The idea is to start with the spherical averages of u, g, h, which satisfy a simpler PDE and use this. This method works for all n but for the sake of time we will discuss it only for

$n = 2, 3$– the general argument can be found in [5]. With this in hand we will then discuss Huygen's principle which is a very interesting property of the wave equation that is dependent on the parity of the dimension (!) and has interesting real world implications.

Now, for $x \in \mathbb{R}^n$, $r.t > 0$ we define:

$$\begin{cases} U(x;r,t) = \fint_{S(x,r)} u(y,t)dS \\ G(x;r) = \fint_{S(x,r)} g(y)dS \\ H(x;r) = \fint_{S(x,r)} h(y)dS \end{cases} \tag{20.2}$$

Note by taking $r \to 0$ we get the original functions back evalauted at x so we have a way back to what we really care about. Now, with a point x fixed one may derive an equation for U just in terms of r, t called the Euler–Poisson–Barboux equation by calculating much as in the proof of the mean value theorem for harmonic functions:

Lemma 20.1. *For a fixed x and u a solution to the wave equation, $U \in C^m(\overline{R}_+ \times [0, \infty))$ and satisfies:*

$$\begin{cases} U_{tt} - U_{rr} - \frac{n-1}{r} U_r = 0 \ \textit{in } \mathbb{R}_+ \times (0, \infty) \\ U = G, U_t = H \ \textit{on } \mathbb{R}_+ \times \{t = 0\} \end{cases} \tag{20.3}$$

This is pretty close to the 1-d wave equation which would be good because then we could use d'Alembert's formula, but its not quite there. First, we consider the case $n = 3$. There's a nice trick to get the wave equation though: let $\widetilde{U} = rU$, $\widetilde{G} = rG$, and $\widetilde{H} = rH$. The assertion is that $\widetilde{U}$ solves:

$$\begin{cases} \widetilde{U}_{tt} - \widetilde{U}_{rr} = 0 \text{ in } \mathbb{R}_+ \times (0, \infty) \\ \widetilde{U} = \widetilde{G}, \widetilde{U}_t = \widetilde{H} \text{ on } \mathbb{R}_+ \times \{t = 0\} \\ \widetilde{U} = 0 \text{ on } \{r = 0\} \times (0, \infty) \end{cases} \tag{20.4}$$

By the PDE and that $n = 3$ we have:

$$\begin{aligned} \widetilde{U}_{tt} &= rU_{tt} \\ &= r[U_{rr} + \frac{2}{r} U_r] \\ &= rU_{rr} + 2U_r \\ &= (U + rU_r)_r \\ &= \widetilde{U}_{rr} \end{aligned} \tag{20.5}$$

So we can use d'Alembert's formula with reflection if we know that $\widetilde{G}''(0) = 0$. By the calculation in the proof of the mean value formula for harmonic functions, we see that $G'(r) = \frac{r}{n} \fint_{S(x,r)} \Delta g(y) dS$ so in particular $G'(0) = 0$. Since $\widetilde{G}''(0) = 2G'(0) + rG''(0)$ we have $\widetilde{G}''(0) = 0$ so can write $\widetilde{U}$ as:

$$\widetilde{U}(x; r, t) = \frac{1}{2}(\widetilde{G}(x; r+t) - \widetilde{G}(x; t-r)) + \frac{1}{2}\int_{-r+t}^{r+t} \widetilde{H}(x; z) dz \tag{20.6}$$

Of course, this is the formula which is valid for just $r \leq t$ but since we will take $r \to 0$ as alluded to before this will be all we need. Since $\lim_{r\to 0} U(x; r, t) = u(x, t)$ and similarly for G and H we have by the defintion of derivative and the fundamental theorem of calculus:

$$\begin{aligned} u(x,t) &= \lim_{r\to 0} \frac{\widetilde{U}(x; r, t)}{r} \\ &= \lim_{r\to 0} \frac{1}{2r}(\widetilde{G}(x; r+t) - \widetilde{G}(x; t-r)) + \frac{1}{2r}\int_{-r+t}^{r+t} \widetilde{H}(x; z) dz \\ &= \widetilde{G}'(x; t) + \widetilde{H}(x; t) \end{aligned} \tag{20.7}$$

Plugging in the defintions of $\widetilde{G}, \widetilde{H}$ then we have:

$$u(x,t) = \frac{\partial}{\partial t}(t \fint_{S(x,t)} g dS) + t \fint_{S(x,t)} h dS \tag{20.8}$$

As we calculated in the proof of the MVP for harmonic functions,

$$\frac{\partial}{\partial t}(\fint_{S(x,t)} g dS = \fint_{S(x,t)} Dg(y) \cdot (\frac{y-x}{t}) dS(y) \tag{20.9}$$

Which, combining with equation 20.8 gives **Kirchoff's formula**:

$$u(x,t) = \fint_{S(x,t)} th(y) + g(y) + Dg(y) \cdot (y - x) dS(y) \tag{20.10}$$

Note that in the integral a derivative of g is involved in contrast to what we get from d'Alembert's formula, so u might not even be as regular as its initial data! Now for $n = 2$ its natural to try the same route; the Euler-Poisson-Darboux equation still holds for U but the problem is that $\widetilde{U} = r\widetilde{U}$ won't solve the wave equation. Instead of having $\widetilde{U}_{tt} = rU_{rr} + 2U_r$ it will just equal, since $n - 1 = 1$, $rU_{rr} + U_r$ and this isn't $\widetilde{U}_{rr}$. To get around this one may use the simple trick that is refered to in this context as the method of descent, because we increase the spatial dimension by one in an obvious way to find a solution for the $n = 3$ equation for which we have a formula, and this formula "descends" to be a formula for the original problem. For u, g, h as usual, defining $\overline{u}, \overline{g}, \overline{h}$ by setting $\overline{u}(x_1, x_2, x_3, t) = u(x_1, x_2, t)$ and similarly for $\overline{g}, \overline{h}$ its easy to see that $\overline{u}$ still solves the wave equation with intial data $\overline{g}, \overline{h}$. Setting $\overline{x} = (x_1, x_2, 0) = (x, 0)$ of course $\overline{u}(\overline{x}, t) = u(x, t)$, and we may apply Kirchoff's formula to evaluate $\overline{u}(\overline{x}, t)$. After some moderate calculus to turn into a purely 2-d integral in terms of g and h one may find **Poisson's formula**:

$$u(x,t) = \frac{1}{2} \fint_{B(x,t)} \frac{tg(y) + t^2 h(y) + tDg(y) \cdot (y - x)}{\sqrt{t^2 - |y - x|^2}} dy \tag{20.11}$$

As you might guess the denominator is relatd the to the parameterization of the hemispheres of the 2-sphere of radius t centered at x. For general odd dimension $n = 2k + 1$ one may find similar formulae as in the $n = 3$ case by instead of scaling U, G, H by r instead scaling by r^{k-1}, and the method of descent then covers the even cases. An interesting feature is that for u to be merely C^2 we have to assume that g, h are in C^m for a value of m which grows with n. With these formulae we can then solve the nonhomogenous wave equation on $\mathbb{R}^n$ (i.e. when the right hand side of the wave equation is nonzero) by Duhamel's principle.

To wrap up this section let's say something about Huygen's principle, which is from physics but concerns phenoema which are modeled by the wave equation. The principle can actually mean one of a couple things, according to the short survey [23]. Huygen's original principle basically is that to model the fate of a wavefront, think coming from a flash of light or a loud bang, one can treat each of the points along the wavefront at a given time as their own sources (i.e. mini flashes or bangs), which largely cancel each other out, and apply this idea iteratively to try to predict how the disturbance will propogate at later times. This idea was called by J. Hadamard the "major premise" of Hyugen. Huygen's principle in the "narrow sense" (or, the "minor premise") is that an instanatneous signal remains instantaneous for every observer at each later time, and as far as I can tell is what mathematicians (at least [5]) mean by the principle. In 3 dimensions this agrees with our experiences of, for instance, a light being turned on and off again quickly appears as a flash even to people far away. On the other hand a rock thrown into a pond, which has a 2-d surface, disturbs the surface of the pond even when the first wave of the rock passes by. Supposing these phenomena are well modeled by the wave equation, we see that the integral in Kirchoff's law is done over the sphere centered at x with radius t whereas Poisson's formula is over the entire ball (and this generalizes to higher dimensions). The upshot is if g and h are very localized (like a bang, a flash, a dropped rock) $u(x, t)$ in odd dimensions can only be nonzero for a short range of times, and hence the support of

u is "sharp" for a given time, but the opposite is true in even dimensions as long as t is large enough relative to the distance of x to the support of g, h.

That the behavior of a PDE can dependent so sharply on dimension is interesting. This isn't the only instance though and often when this occurs (outside some trivial reasons, like the PDE becoming an ODE is small dimensions) points to something deep and mysterious. We have Kirchoff and Poisson's formula, but it isn't all that clear to me from the outset why we should expect them to have roughly the forms they do, although maybe somebody knows the answer. These can be derived using Fourier transforms, and maybe its clearer whats going on from that perspective. Huygen's principle is important in our everyday lives too: one imagines that if we lived in an even number of spatial dimensions sight and hearing (and so communication) would be more complicated. On the other hand some neat ideas out there using it to form hypotheses in cosmology too – if we could find evidence of residual waves from a disturbance (like black holes colliding, or whatever, I'm not a physicist and just spouting off here) it would possibly indicate that our universe actually had an even number of spatial dimensions with some of them normally inperceptable to us, like the "compact dimensions" in string theory.

21. Energy methods for the wave equation

To finish off our lightning tour of the wave equation we discuss energy methods for it. Taking notation from the heat equation, first we discuss uniqueness of the wave equation for solutions in U_T where U is bounded with smooth boundary. Of course we haven't solved the existence problem, but it follows from general theory which we will hopefully at least touch on:

Theorem 21.1. *There exists at most one function $u \in C^2(\overline{U}_T)$ which solves*

$$\begin{cases} u_{tt} - \Delta u = f \text{ in } U_T \\ u = g \text{ on } \Gamma_T \\ u_t = h \text{ on } U \times \{t = 0\} \end{cases} \tag{21.1}$$

Proof: We'll just sketch this because this technique is old news to us now. As usual we suppose that there are two solutions and consider their difference, which has trivial data on the boundary and solves the wave equation. The clever part is to define the right energy to consider, and if one uses

$$E(t) = \frac{1}{2} \int_U w_t^2(x,t) + |Dw(x,t)|^2 dx \tag{21.2}$$

One finds the energy is constant so must be constantly equal to zero, implying w is constant so equal to zero itself. For some motivation for this energy one sees in the course of the proof that, after integrating by parts, that $w_{tt} - \Delta w$ appears which is zero. □

A more interesting application of energy methods is to show that disturbances "propogate" at finite speed, in that a wave caused by a disturbance (think a flash or a scream) can't travel too fast. In other words the speed of sound/light/the object in the model is built into the wave equation by scaling t appropriately. For the following, suppose that $u \in C^2(\mathbb{R}^n \times (0,\infty))$ solves the wave equation and consider the backwards wave cone with apex (x_0, t_0) by

$$K(x_0, t_0) = \{(x,t) \mid 0 \leq t \leq t_0, |x - x_0| \leq t_0 - t\} \tag{21.3}$$

(This looks like a party hat.) Then the claim is the following; note that it also follows from the representation formulae from the last section because the values of $u(x,t)$ only depend on the value of g, h in a ball of radius t but the argument is much easier than deriving those represenations and uses no regularity assumptions on the initial data:

Theorem 21.2. *If $u = u_t = 0$ on $B(x_0, t_0) \times \{t = 0\}$, then $u = 0$ within the cone $K(x_0, t_0)$.*

Proof: We define similar to above the local energy:

$$e(t) = \frac{1}{2}\int_{B(x_0,t_0-t)} u_t^2(x,t) + |Du(x,t)|^2 dx \tag{21.4}$$

where $t \leq t_0$. Its "local" because the integral is only done on the ball above and that ball is in turn a slice of the cone $K(x_0,t_0)$. As we are want to do, we calculate its derivative, with the second term coming from the fundamental theorem of calculus:

$$\begin{aligned} e'(t) &= \int_{B(x_0,t_0-t)} u_t u_{tt} + Du \cdot Du_t dx - \frac{1}{2}\int_{S(x_0,t_0-t)} u_t^2(x,t) + |Du(x,t)|^2 dS \\ &= \int_{B(x_0,t_0-t)} u_t(u_{tt} - \Delta u) dx + \int_{S(x_0,t_0-t)} \frac{\partial u}{\partial \nu} u_t dS - \int_{S(x_0,t_0-t)} u_t^2(x,t) + |Du(x,t)|^2 dS \\ &= \int_{S(x_0,t_0-t)} \frac{\partial u}{\partial \nu} u_t - \frac{1}{2}u_t^2 - \frac{1}{2}|Du|^2 dS \end{aligned} \tag{21.5}$$

Where above we integrated by parts and used that u solved the heat equation. This will be nonpositive, which is what we want (apriori, you always want something to have a sign one way or the other) as long as the first term doesn't overwhelm the second and third. This follows using that $|ab| \leq \frac{a^2+b^2}{2}$. So, because $e(0) = 0$ by assumption it remains zero for all $0 \leq t \leq t_0$. This gives that u_t, Du are zero in $K(x_0, t_0)$ so that u is identically zero in it. $\square$

Note that in above the geometry of K mattered in the boundary term we got when calculating the derivative of e. If instead we considering an expanding set (in t) we would have gotten the wrong sign on the $\int_{S(x_0,t_0-t)} u_t^2(x,t) + |Du(x,t)|^2 dS$ contribution. Physically, there could be disturbances which are supported outside $B(x_0, t_0)$ which can bleed in at later times. Note that this result strongly contrasts with the heat equation, for which we showed infinite propogation speed.

22. Case of the century: Holder v. Sobolev, and the method of continuity

Having spent some time on the three classical PDE above, we now turn to the following important generalization of the Laplace equation, considered on an open bounded subset U of $\mathbb{R}^n$:

$$Lu = -\sum_{i,j=1}^{n} a^{ij}(x) u_{x_i x_j} + \sum_{i} b^i(x) u_{x_i} + c(x) u \tag{22.1}$$

where here the a^{ij} are the coefficients of a symmetric matrix (so that $a^{ij} = a^{ji}$). L will be said to be **(uniformly) elliptic** if there exists a constant $\theta > 0$ such that

$$\sum_{i,j=1}^{n} a^{ij}(x)\xi_i \xi_j \geq \theta |\xi|^2 \tag{22.2}$$

for almost every $x \in U$ and all vectors $\xi \in \mathbb{R}^n$. Of course the regular old Laplacian is elliptic in this case. Now to solve this problem there are a variety of methods, some of which end up being easier to apply depending on the context. In the study of general elliptic problems, there are two major types of function spaces that seem to be considered often, Holder spaces and Sobolev spaces. Not to say these are the only spaces of functions anyone has ever considered, and depending on the problem at hand they might be modified accordingly. For instance, by multiplying by additional functions in their definition as "weights" or to incorporate the boundary data imposed. As a bit of philosophy, to be taken with a grain of salt, the space of functions at hand should be large enough for the PDE to be solvable in some sort of manner, perhaps (often) merely weakly, while being small enough, in a sense depending on how a weak solution is defined, so that one can prove something helpful about the solution. For instance, its often helpful if in a space a solution can be bounded in the appropriate norm by the input data – estimates of this form are generally called "apriori estimates". Exactly how much one needs depends on the idea to actually solve the problem and what you want to do with/know about the solution so its an art, not a science – or at least it seems so to me (a more experienced analyst might have a different opinion). For example in the Perron method above we considered a notion of weakly sub/superharmonic functions and the functions themselves were required to just be continuous, but baked into their defintion they had to satisfy a strong comparison principle with classical harmonic functions.

For the remainder of we'll spend most of our time working with/in Sobolev spaces, but first let's quickly discuss Holder spaces and the method of continuity. Connecting with the above, its a relatively narrow space of functions but we can prove good apriori estimates, and the method of continuity is a way to take advantage of that. The idea of the method of continuity is that to solve a PDE say $Lu = f$, one can try to put it into a family of PDE $L_t u = f$ where $L_0 u = f$ is easy to deal with and $L_1 = L$ is what we care about. Then if we can show the $t \in [0, 1]$ for which L_t is solvable is both open and closed we can solve the original problem since the interval is connected. Its a beautiful idea that generalizes well to nonlinear problems, and is notably the method of proof used by Yau to show the Calabi conjecture (which amongst other results got him a Fields medal). First we define Holder spaces:

Definition 22.1. *(1) Where $u : U \to \mathbb{R}$ is bounded and continuous, the γ-th Holder seminorm of $u : U \to \mathbb{R}$ is*

$$[u]_{C^{0,\alpha}(\overline{U})} = \sup_{x \neq y \in U} \frac{|u(x) - u(y)|}{|x - y|^\alpha} \tag{22.3}$$

and relatedly a function is said to be Holder continuous with exponent α if $|u(x) - u(y)| \leq C|x - y|^\alpha$ for some constant C.

(2) The α Holder norm is

$$||u||_{C^{0,\alpha}(\overline{U})} = ||u||_{L^\infty(\overline{U})} + [u]_{C^{0,\alpha}(\overline{U})} \tag{22.4}$$

(3) Finally, the Holder space $C^{k,\alpha}(\overline{U})$ denotes the space of all functions $u \in C^k(\overline{U})$ for which the norm below is finite:

$$||u||_{C^{k,\alpha}(\overline{U})} = \sum_{|\beta| \leq k} ||D^\beta u||_{L^\infty(\overline{U})} + \sum_{|\beta| = k} [D^\beta u]_{C^{0,\alpha}(\overline{U})} \tag{22.5}$$

is

Some assorted notes: the Holder seminorm isn't a norm for the simple reason that its zero on constant functions. Note that in the defintion of Holder continuous that if $\alpha > 1$ then u is differentiable and whats more its constant, so one almost always restricts to considering $0 < \alpha \leq 1$. By the mean value theorem (the calculus one) we see that if $u \in C^{k+1}(U)$ and u, U are bounded then $D^\beta u$ is Lipschitz continuous for $\beta = k$ and in particular is Holder continuous for any $\gamma \in (0,1)$. So the space $C^{k,\alpha}(\overline{U})$ can be thought of as laying in between C^k and C^{k+1}. Importantly, the space $C^{k,\alpha}(\overline{U})$ turns out to be complete (and hence a Banach space).

With this front matter out of the way, suppose that u is a $C^{2,\alpha}$ solution to the problem

$$\begin{cases} Lu = f \text{ in } U \\ u = g \text{ on } \partial U \end{cases} \tag{22.6}$$

Where U is bounded with a $C^{2,\alpha}$ boundary, and $f \in C^\alpha(\overline{U})$, $g \in C^{2,\alpha}(\overline{U})$. Then for solutions to this problem there are the **Schauder estimates:**

Theorem 22.1. *Refering to the defintion of L above and supposing its elliptic, suppose* $||a^{ij}||_{C^{0,\alpha}(V)}, ||b^i||_{C^{0,\alpha}(V)}, ||c||_{C^{0,\alpha}(V)} < K$ *for some fixed constants* K, α *where* V *is any domain* $V \subset\subset U$*. Then we have there exists a constant* $C = C(U, V, \alpha, n, \theta, K)$ *such that*

$$||u||_{C^{2,\alpha}(V)} \leq C(||f||_{C^{0,\alpha}(U)} + ||u||_{C^{0,\alpha}(U)} \tag{22.7}$$

This is called the interior estimate. The global, or boundary estimate says:

$$||u||_{C^{2,\alpha}(U)} \leq C_1(||f||_{C^{0,\alpha}(U)} + ||g||_{C^{2,\alpha}(U)} + ||u||_{L^2(U)} \tag{22.8}$$

where C_1 *has the same dependencies as* C*. When, refering to* L*,* $c \geq 0$ *we can drop the* L^2 *norm of* u *essentially using the maximum principle to write:*

$$||u||_{C^{2,\alpha}(U)} \leq C_2(||f||_{C^{0,\alpha}(U)} + ||g||_{C^{2,\alpha}(U)} \tag{22.9}$$

As mentioned these are called apriori estimates, because the can be thought of as estimates for the solution in terms of the initial data. There's a number of methods of ways to derive these. One way is via estimates for $L = \Delta$ in terms of the fundamental solution as one would find in [6, 11] which then gives the full Schauder estimates by an approximation argument. An easy "proof by scaling" for the interior estimate came later in [20] – this might be covered in detail discussion (there are also some nice, very short writeups one can find online about it). A brief outline:

(1) Note that harmonic functions $u : \mathbb{R}^n \to \mathbb{R}$ such that $\sup_{B_r(0)} |u| \leq Cr^{3-\epsilon}$ for some $\epsilon > 0$ are quadratic polynomials. To see this first note by the derivative estimates $D^\beta u = 0$ for $|\beta| > 2$. Using that harmonic functions are analytic, with coefficients given directly in terms of $D^\beta u(0)$ (considering the series centered at the origin), the claim follows.

(2) For $u \in C^{2,\alpha}(\mathbb{R}^n)$ there exists a constant $C = C(\alpha, n) < \infty$ such that $\sum_{|\beta|=2} [D^\beta u]_{C^{0,\alpha}(\mathbb{R}^n)} < C[\Delta u]_{C^{0,\alpha}(\mathbb{R}^n)}$ or in shorthand $[D^2 u]_{C^{0,\alpha}(\mathbb{R}^n)} < C[\Delta u]_{C^{0,\alpha}(\mathbb{R}^n)}$. This is a proof by contradiction, where we suppose there is a sequence u_k of such functions such that $[D^2 u_k]_{C^{0,\alpha}(\mathbb{R}^n)} > k[\Delta u_k]_{C^{0,\alpha}(\mathbb{R}^n)}$. Rescaling by $[D^2 u_k]_{C^{0,\alpha}(\mathbb{R}^n)}$ we may take a limit (Holder spaces are complete!) and obtain a function u in such a way which, after some fiddling, has u, Du, D^2u all equal to zero at the origin, is globally bounded, is harmonic, but $D^2 u$ is nonzero elsewhere. The first 3 points combined with (1) above give u is constant, which contradicts the last point that $D^2 u$ is nonzero somewhere. This is arguably the core of the argument and where this method differs from the method using the fundamental solution.

(3) Suppose that L is a general elliptic operator with Holder bounded coefficients. Fixing a point x_0, by a change of coordinates its not so hard to see that for a function $u \in C^{2,\alpha}(\mathbb{R}^n)$ that $[D^2 u]_{C^{0,\alpha}(\mathbb{R}^n)} < C[a^{ij} u_{ij}]_{C^{0,\alpha}(\mathbb{R}^n)}$ (summing over the indices) for some constant C. Then the general Schauder estimate follows for a solution u to $Lu = f$ by "freezing" the coefficients of L at a point and using they are Holder continuous. The Holder continuity of the coefficients gives a quantitative bound on the difference of L to L with frozen coefficients which lets us, with some fairly straightforward estimation using interpolation inequalities, to control the difference well enough to prove the general estimate.

There's also parabolic versions of these estimates, naturally. With this in hand, lets sketch the following using the method of continuity:

Theorem 22.2. *Suppose L is as above for which Schauder estimates apply and that $c \geq 0$. Then for any $f \in C^\alpha(\overline{U})$, $g \in C^{2,\alpha}(\overline{U})$ there exists a unique solution to the problem 22.6 in $C^{2,\alpha}$.*

Proof: Now, one can see this problem is solvable when $L = \Delta$, and this is called Kellogg's theorem. It turns out (chapter 4 in [6]) that one can show the theorem is true in the ball, with a careful analysis of the Green's function on it and then Perron's method gives the claim on more general domains. Alternately, solvability for general L can be considered first on balls and then shown on general domains using a suitable generalization of Perron's method for them (chapter 6 in [6]).

We also note that without loss of generality $g = 0$, because if we let $\overline{f} = f - Lg$ and $\overline{g} = 0$ then a solution $\overline{u}$ to 22.6 with $\overline{f}, \overline{g}$ will give a solution u to the original problem by setting $u = \overline{u} + g$.

Now for general L define $L_t = tL + (1-t)\Delta$. Denoting by $I \subset [0,1]$ the set for which the problem 22.6 can be solved using L_t from above $\{0\} \subset I$ so in particular it isn't empty. It turns out this is a family of uniformly elliptic operators with Holder bounded coefficients independent of t, so we can apply Schauder estimates to all of them. By the same argument as indicated above then I is closed. We wish to show its open which by the connectedness of I will give the claim. By the reduction to $g = 0$ the Schauder estimates 22 show for $t \in I$.

$$||u_t||_{C^{2,\alpha}(U)} \leq C_2(||f||_{C^{0,\alpha}(U)} \tag{22.10}$$

where u_t is a solution to the problem using L_t. Considered as operators from $C^{2,\alpha}(U) \to C^{0,\alpha}(U)$, the solvability of the PDE says that the L_t are surjective; the estimate above says they are injective so in particular L_t^{-1} exists and is a linear continuous operator. Our idea then is, for s apriori not known to be in I, to note after rearranging that solving $L_s u = f$ is equivalent to finding u such that

$$L_t u = f + (L_t - L_s)u = f + (t-s)(L_0 u - L_1 u) \tag{22.11}$$

using the defintion of L_t. Rearranging this using $t \in I$ so L_t^{-1} exists:

$$u = L_t^{-1} f + (t-s) L_t^{-1}((L_0 - L_1)u) \tag{22.12}$$

We can write the right hand side as $F(u)$, so we want to solve the fixed point problem $u = F(u)$! (this is reason for celebration because we've had a lot of luck with these). By the contraction mapping principle, it suffices to show there is some $c < 1$ so that for $u, v \in C^{2,\alpha}(U)$,

$$||F(u) - F(v)||_{C^{2,\alpha}(U)} \leq c||u - v||_{C^{2,\alpha}(U)} \tag{22.13}$$

This isn't bad, with some crude estimating we find

$$||F(u) - F(v)|| \leq |t-s|||L_t^{-1}||(||L_0|| + ||L_1||)||u - v|| \tag{22.14}$$

Now, we can take $|t-s|$ small enough so that $|t-s|||L_t^{-1}||(||L_0|| + ||L_1||$ is less than one, so the contraction mapping principle gives us the fixed point. Since it worked for any s so that $|t-s|$ was sufficiently small, this gives I is open. Since I is open, closed, and nonempty $I=[0,1]$ and we are done.

□

We just solved a pretty general elliptic problem and the method can be applied in other contexts, so why not just stop here? To lay on some more philosophy that you shouldn't take too seriously, one reason is that for variational problems the most obvious method to try to find solutions really is to find minimizers of the associated energy, so, with the direct method in mind, considering complete spaces which are defined in terms of integration are the most natural to use – even when the PDE is nonlinear and not of the form above. A particular example I have in mind that I mentioned in class before is the Plateau problem. More specifically the energy minimizing sequence of functions one considers in the direct method will not necessarily produce a sequence of functions which converge in $C^{k,\alpha}(U)$ even if we supposed they were all in this space. Another reason is that linear PDE correspond to linear operators on function spaces, and with the correct choice of spaces these should be well behaved enough for which to apply linear–algebraic ideas. To do so it can be helpful from experience, if the underlying space is a Hilbert space which has richer structure, which the Holder spaces are not (the parallelogram identity does not hold). And in what is arguably the most obvious inner product to put on functions though, the L^2 inner product, the Holder spaces are not closed with respect to the associated topology or self dual via the associated inner product. In developing the right, or at least good, function space(s) to deal with these considerations we will also develop a framework that can be applied to more general linear PDE than that of the type above. This of course brings us to Sobolev spaces.

23. Basic defintions and properties of Sobolev spaces

Now, because we are dealing with PDE we want the spaces we work with to incorporate the notion of derivatives in some way, either explicitly or by comparison with more regular functions as with the sub/superharmonic functions in the Perron method (see also the notion of viscosity solutions). The point of view here is to consider spaces which have derivatives in a weak sense, inspired by integration by parts.

Referring to the space $C_c^\infty(U)$, the compactly supported smooth functions in a C^1 domain U, as the space of test functions we define a weak derivative of a function by the following:

Definition 23.1. *Suppose $u, v \in L^1_{loc}(U)$ and α is a multiindex. We say that v is the α-th weak partial derivative of u, written $D^\alpha u = v$, when*

$$\int_U u D^\alpha \phi dx = (-1)^{|\alpha|} \int_U v \phi dx \tag{23.1}$$

for all test functions ϕ on U.

This is exactly what we would get if u were smooth from integration by parts; that $u, v \in L^1_{loc}(U)$ is there to ensure that the numbers on both sides of the equation are finite, and in keeping with wall of text above we should be considering functions of spaces complete with respect to integral norms (i.e. the L^p spaces). Clearly if a function has a derivative in the regular "classical" sense it has a weak derivative equal to it, but as one would anticipate there are functions which don't have classical derivatives but do have weak ones. Some tame examples can be found by considering piecewise linear functions, for instance, but there are even functions with weak derivatives which aren't bounded on any domain! The following is almost immediate except perhaps for a standard argument I write out in detail.

Lemma 23.1. *A weak α-th partial derivative of u, if it exists, is uniquely defined up to a set of measure zero.*

Proof: Assume that $v_1, v_2 \in L^1_{loc}(U)$ both satisfy

$$\int_U u D^\alpha \phi dx = (-1)^{|\alpha|} \int_U v_1 \phi dx = (-1)^{|\alpha|} \int_U v_2 \phi dx \tag{23.2}$$

for all $\phi \in C_c^\infty(U)$. Then

$$\int_U (v_1 - v_2) \phi dx = 0 \tag{23.3}$$

for all test functions ϕ. This implies $v_1 = v_2$ almost everywhere. This is a standard argument/fact but its not quite as straightforward as in earlier occurances because v_1, v_2 aren't necessarily continuous. Suppose that $v_1 \neq v_2$ on a set of positive measure. in particular, one of the sets $I_+ = \{x \mid v_1(x) > v_2(x)\}$ or $I_- = \{x \mid v_1(x) < v_2(x)\}$ has positive measure; since $v_1, v_2 \in L^1_{loc}(U)$ these functions are measurable so these sets are. Supposing the measure of I_+ is positive as the irrationals or fat cantor set show its not a given that it has an interior, so we can't necessarily pick a positive test function supported totally on it which would give a contradiction to 23.3. In the following without loss of generality this set is bounded: we can consider its intersection with $B_R(0) \cap U$ for R sufficiently large so the intersection has positive measure. Now, while I_+ might not have an interior we can approximate I_+ in measure as well as we want by open sets V_ϵ containing it, i.e. so that $m(V_\epsilon \setminus I_+) < \epsilon$ for any $\epsilon > 0$ we pick. With this in mind denoting by $a = \int_{I_+}(v_1 - v_2)dx$ and $b(\epsilon) = \int_{V_\epsilon}(v_1 - v_2)dx$ we can pick ϵ small enough so that $|a - b(\epsilon)| < a/2$, and in turn we can easily construct a smooth bump function ϕ supported on V_ϵ so that $\int_U (v_1 - v_2)\phi dx > 0$ giving a contradiction. □

So, at least in the eyes of measure theory the weak derivative of a function is well defined. To be good stand ins for classical derivatives its reasonable to also want them to have the same algebraic properties (linearity, etc.) which they do, but for the ease of stating those we'll postpone it briefly. The point is that these are promising enough to be worth the trouble and we are now ready to formally define **Sobolev spaces** and their norms. Below, $1 \leq p \leq \infty$ and k is a nonnegative integer.

Definition 23.2.

(1) *The Sobolev space $W^{k,p}(U)$ consists of all locally integrable functions $u : U \to \mathbb{R}$ such that for each multiindex α with $|\alpha| \leq k$, $D^\alpha u$ exists in the weak sense and belongs to $L^p(U)$.*

(2) *If $p = 2$, one often writes $H^k(U)$ for $W^{k,2}(U)$. The H is for Hilbert and these will be the most important Sobolev spaces for us.*

(3) *If $u \in W^{k,p}(U)$ we define its norm (its easy to check its indeed a norm) by*

$$||u||_{W^{k,p}(U)} = \begin{cases} (\sum_{|\alpha| \leq k} \int_U |D^\alpha u|^p)^{1/p} \text{ for } 1 \leq p < \infty \\ \sum_{|\alpha| \leq k} \text{ess sup}_U |D^\alpha u| \text{ for } p = \infty \end{cases} \tag{23.4}$$

And in so doing these define the topology of the Sobolev spaces. Finally, we define a set of functions which essentially corresponds to those which are zero along ∂U:

(4) *$W_0^{k,p}(U)$ denotes the closure of $C_c^\infty(U)$ in $W^{k,p}(U)$, and similarly we define $H_0^k(U)$.*

The final defintion above will make more sense when we discuss the density properties of smooth functions in Sobolev space. Because the Sobolev spaces are composed of functions which have finite L^p norm it makes working in them in the noncompact case a bit trickier because for instance constant functions aren't even in L^p – I've heard it said that for this reason using Holder spaces (and associated techniques) in problems over noncompact spaces can be easier if possible whose norms are pointwise defined. Now as already mentioned weak derivatives satisfy many of the same formal properties classical derivatives do. Its easy to check the following list:

Theorem 23.2. *Assume $u, v \in W^{k,p}(U)$, $|\alpha| \leq k$. Then*

(1) $D^\beta(D^\alpha u) = D^\alpha(D^\beta u) = D^{\alpha+\beta}u$ for all multiindices α, β with $|\alpha| + |\beta| \leq k$
(2) For each $a, b \in \mathbb{R}$ $au + bv \in W^{k,p}(U)$ too and $D^\alpha(au + bv) = aD^\alpha u + bD^\alpha v$ for all $|\alpha| \leq k$.
(3) If $V \subset U$ is open then $u \in W^{k,p}(V)$
(4) If ϕ is a test function then $\phi u \in W^{k,p}(U)$ and

$$D^\alpha(\phi u) = \sum_{\beta \leq \alpha} \frac{\alpha!}{\beta!(\alpha - \beta)!} D^\beta \phi D^{\alpha-\beta} u \tag{23.5}$$

As we've discussed time and again, it desirable for a number of reasons that our spaces we work with are complete. We take limits of functions a lot for instance, and also much of functional analysis is built atop complete spaces. Indeed, Sobolev spaces are complete:

Theorem 23.3. *For each positive integer k and $1 \leq p \leq \infty$ the Sobolev space $W^{k,p}(U)$ is a Banach space.*

Proof: As we said its easy to see the length defined in item (3) of definition 23.2 is really a norm using Minikowski's identity and the triangle inequality for the L^p norms, so we discuss completeness. If we consider a Cauchy sequence of functions $u_m \in W^{k,p}(U)$ then u_m and each of its weak derivatives are Cauchy sequences in $L^p(U)$ by the defintion of the Sobolev norm so individually converge to functions in u_α in $L^p(U)$ by the completeness of $L^p(U)$, indexed in the obvious way. Now for a fixed test function ϕ, and multiindex α with $|\alpha| \leq k$, since $u_m \to u$ in $L^p(U)$ $u_m D^\alpha \phi \to u D^\alpha \phi$ in $L^p(U)$ as well. This implies:

$$|\int_U (u - u_m) D^\alpha \phi dx| \leq \int_U |(u - u_m) D^\alpha \phi| dx = ||(u - u_m) D^\alpha \phi||_{L^1(U)} \leq ||u - u_m||_{L^p(U)} ||D^\alpha \phi||_{L^q(U)} \tag{23.6}$$

Where q is the Holder conjugate to p. Since ϕ is a test function $||D^\alpha\phi||_{L^q(U)}$ is finite so that, because $u_m \to u$ in $L^p(U)$, we have

$$\int_U uD^\alpha\phi dx = \lim_{m\to\infty}\int_U u_m D^\alpha\phi dx \tag{23.7}$$

By the defintion of weak derivative and the same reasoning for u_α we have:

$$\lim_{m\to\infty}\int_U u_m D^\alpha\phi dx = \lim_{m\to\infty}(-1)^{|\alpha|}\int_U D^\alpha u_m\phi dx = (-1)^{|\alpha|}\int_U u_\alpha\phi dx \tag{23.8}$$

In particular, the limits of the weak derivatives of u_m, which were taken individually, are indeed weak derivatives of the limit function u. Hence $u \in W^{k,p}(U)$ so it is a complete normed space as claimed. □

As mentioned, in the space of Sobolev functions there can be some pretty wild functions. It can be good to know for proofs that a relatively tame subset of functions is dense in it, and this is the content of the following few theorems:

Theorem 23.4. *Assume $u \in W^{k,p}(U)$ for some $1 \le p < \infty$ and set as usually $u^\epsilon = \eta_\epsilon * u$ in U_ϵ (the set of points ϵ distance from the boundary). Then*

(1) $u^\epsilon \in C^\infty(U_\epsilon)$ for each $\epsilon > 0$, and

(2) $u^\epsilon \to u$ in $W^{k,p}_{loc}(U)$ as $\epsilon \to 0$

Proof: The first claim is familiar to us from our study of harmonic functions and is just because differentiation passes through the integral in the defintion of convolution and lands on η, which is smooth. For the second claim we first show that if $|\alpha| \le k$, then $D^\alpha u^\epsilon = \eta_\epsilon * D^\alpha u$ in U_ϵ or in other words convolution commutes with weak derivatives:

$$\begin{aligned} D^\alpha u^\epsilon(x) &= D^\alpha\int_U \eta_\epsilon(x-y)u(y)dy \\ &= \int_U D^\alpha_x\eta_\epsilon(x-y)u(y)dy \\ &= (-1)^{|\alpha|}\int_U D^\alpha_y\eta_\epsilon(x-y)u(y)dy \text{ (change of variables)} \\ &= (-1)^{|\alpha|+|\alpha|}\int_U \eta_\epsilon(x-y)D^\alpha_y u(y)dy \text{ (for } x \text{ fixed, } \eta(x-y) \text{ is test function)} \\ &= \int_U \eta_\epsilon(x-y)D^\alpha_y u(y)dy \ ((-1)^2=1) \\ &= \eta_\epsilon * D^\alpha u(x) \end{aligned} \tag{23.9}$$

Since generally speaking for a function $f \in L^p(V)$ $\eta_\epsilon * f \to f$ in $L^p(V)$ as $\eta \to 0$ for a fixed set $V \subset\subset U$ (this isn't terribly hard but not worth spelling out here, see appendix C in [5]) we have the result from the defintion of the Sobolev norm in terms of sum, raised to $1/p$, of the L^p norms of the weak derivatives.

□

That the convergence is in $W^{k,p}_{loc}(U)$ and not $W^{k,p}(U)$ can be removed. This point might seem a little subtle (and it sort of is) – the thing is that above we knew that in the proof above the convergence was in any fixed $V \subset\subset U$, but in an exhaustion of U with such sets V_i u^ϵ might not converge to u uniformly in $W^{k,p}(V_i)$ for each i apriori. This isn't bad to deal with though by a partition of unity argument where for each V_i the function constructed is $\epsilon/2^i$ close to u in $W^{k,p}$ (see [5] section 5.3 for more details):

Theorem 23.5. *Assume $u \in W^{k,p}(U)$ for some $1 \leq p < \infty$ where U is bounded. Then there exists functions $u_m \in C^\infty(U) \cap W^{k,p}(U)$ so $u_m \to u$ in $W^{k,p}(U)$*

The last theorem density theorem we wish to discuss here is that approximating smooth functions can be taken to be smooth up to the boundary. The idea is that in the following because the boundary of U is C^1 it can be "straightened out" locally by a reparameterization of $\mathbb{R}^n$. Then we can take a function $u \in W^{k,p}(U)$, translate it up a bit in the direction away from the boundary (which is continuous), mollify, and then translate it back down. The global statement follows by a partition of unity argument similar to what was inidicated above:

Theorem 23.6. *Assume that U is bounded and ∂U is C^1. Suppose $u \in W^{k,p}(U)$ for some $1 \leq p < \infty$. Then there exists functions $u_m \in C^\infty(\overline{U})$ such that $u_m \to u$ in $W^{k,p}(U)$.*

Our next statement, which will be used in the sequel but whose proof doesn't matter too much to us, is the following extension theorem and will be used in conjunction with the approximation results above. Note that extending a function from U to $\mathbb{R}^n$ has to be done in a careful way. For instance, we wanted to extend the function $f = 0$ defined on $\{x \in \mathbb{R} \mid x < 0\}$ to all of $\mathbb{R}$ if we picked some extension other than by zero f would not have a weak derivative; its would have to have a "jump" at the origin which would correspond to some multiple of the Dirac delta which isn't represented by integration against an integrable function. In other words, a more reasonable thing to do is to reflect f at least locally across the origin. By straightening out the boundary of U locally in the following and reflecting u in a clever way one can show the following, again found in [5]:

Theorem 23.7. *Assume U is bounded and ∂U is C^1 and consider a bounded open set V such that $U \subset\subset V$. Then there exists a bounded linear operator $E : W^{1,p}(U) \to W^{1,p}(\mathbb{R}^n)$ such that for each $u \in W^{1,p}(U)$:*

(1) $Eu = u$ a.e. in U
(2) Eu has support within V, and
(3) $||Eu||_{W^{1,p}(\mathbb{R}^n)} \leq C||u||_{W^{1,p}(U)}$ where the constant depends only on p, U, V.

With more differentiability of the boundary one can extend $W^{k,p}$ using so-called higher order reflections. To see how these reflections are designed, say across the half space $x_n > 0$ one can reason that they should preserve poloynomials of x_n up to degree k, and this gives a system of linear equations that can be explictly solved. The final result we wish to discuss is the trace theorem:

Theorem 23.8. *Assume that U is bounded and ∂U is C^1. Then there exists a bounded linear operator $T : W^{1,p}(U) \to L^p(\partial U)$ such that*

(1) $Tu = u |_{\partial U}$ if $u \in W^{1,p}(U) \cap C(\overline{U})$
(2) $||Tu||_{L^p(\partial U)} \leq C||u||_{W^{1,p}(U)}$

for each $u \in W^{1,p}(U)$ with the constant C depending only on p and U.

One then calls Tu the trace of u. Item (1) says essentially the trace could reasonably be called the restriction of u onto ∂U, and the second item says its continuous. It might seem obvious at first that such an operator should exist, but actually there isn't one for merely L^p functions – the reasons is because the compactly supported functions on U are dense in $L^p(U)$. The proof essentially goes by first considering U which is a half space and then extending U and approximating it with a C^1 function. Then the restriction of u can be bounded by its $W^{1,p}$ norm using the divergence theorem. For general U one can then straighten out its boundary locally and use that U is bounded.

Of course, considering the trace of a function u is interesting to us in PDE when prescribing boundary data, which after modifying the input data can often be arranged to be zero. The following theorem, which we'll just state, says that the trace zero functions are exactly those which can be approximated by smooth functions which vanish along ∂U:

Theorem 23.9. *Assume U is bounded, ∂U is C^1, and that $u \in W^{1,p}(U)$. Then $u \in W_0^{1,p}(U)$ if and only if $Tu = 0$ on ∂U.*

24. The Sobolev inequalities

With some generalities about Sobolev spaces out of the way, we now go on to proving some extremely important inequalities and related embedding properties for them with some related facts. Our first goal is to establish an estimate of the form

$$||u||_{L^q(\mathbb{R}^n)} \leq C||Du||_{L^p(\mathbb{R}^n)} \tag{24.1}$$

For functions $u \in C_c^\infty(\mathbb{R}^n)$ and C independent of u – it will then imply corresponding estiamtes for functions in $W^{1,p}(U)$, U bounded, by the density and extension theorems above. First, we note by the fundamental theorem of calculus suggests that an estimate of the form above could be reasonable (and it will be used in the proof) – of course if u doesn't have compact support the result can't be true in full generality considering the constant functions. Assuming an inequality of the above form is true to decide what q could/should be, which will also tell us something about p, we consider reparameterizations of u by scaling: writing $u_\lambda(x) = u(\lambda x)$ we have

$$\int_{\mathbb{R}^n} |u_\lambda|^q dx = \int_{\mathbb{R}^n} |u(\lambda x)|^q dx = \frac{1}{\lambda^n}\int_{\mathbb{R}^n} |u(y)|^q dy \tag{24.2}$$

and

$$\int_{\mathbb{R}^n} |Du_\lambda|^p dx = \int_{\mathbb{R}^n} \lambda^p |Du(\lambda x)|^p dx = \frac{\lambda^p}{\lambda^n}\int_{\mathbb{R}^n} |Du(y)|^p dy \tag{24.3}$$

Of course, if u has compact support so does u_λ, and because we stipulated that C shouldn't have anything to do with u we must have (remembering in L^r norm you raise finally to power $1/r$:

$$\frac{1}{\lambda^{n/q}}||u||_{L^q(\mathbb{R}^n)} \leq C\frac{\lambda}{\lambda^{n/p}}||Du||_{L^p(\mathbb{R}^n)} \tag{24.4}$$

So rearranging we have $||u||_{L^q(\mathbb{R}^n)} \leq C\lambda^{1-\frac{n}{p}+\frac{n}{q}}||Du||_{L^p(\mathbb{R}^n)}$ for any $\lambda > 0$. Now, suppose $1 - \frac{n}{p} + \frac{n}{q} > 0$. Taking $\lambda \to 0$ would then imply every C_c^∞ function has $||u||_{L^q(\mathbb{R}^n)} = 0$, which is obviously a contradiction. Similarly if $1 - \frac{n}{p} + \frac{n}{q} < 0$ we can take $\lambda \to \infty$ to get a contradiction. So we must have $1 - \frac{n}{p} + \frac{n}{q} = 0$ or in other words $\frac{1}{q} = \frac{1}{p} - \frac{1}{n}$. Since we have only ever consider L^r spaces for $r > 0$ this tells us we should restrict to $1 \leq p < n$, and solving for q we see $q = \frac{np}{n-p}$. This number, denoted p^*, is called the Sobolev conjugate of p. Now we prove an inequality of this form, called the Gagliardo–Nirenberg–Sobolev inequality, is true:

Theorem 24.1. *Assume that $1 \leq p < n$. Then there exists a constant $C = C(p, n)$ such that*

$$||u||_{L^{p^*}(\mathbb{R}^n)} \leq C||Du||_{L^p(\mathbb{R}^n)} \tag{24.5}$$

for all $u \in C_c^1(\mathbb{R}^n)$.

Proof: First we suppose that $p = 1$, so that the Sobolev conjugate $p^* = n/(n-1)$. Since u has compact support the fundamental theorem of calculus says:

$$u(x) = \int_{-\infty}^{x_i} u_{x_i}(x_1, \ldots, x_{i-1}, y_i, x_{i+1}, \ldots, x_n) dy_i \tag{24.6}$$

This implies that $|u(x)| \leq \int_{-\infty}^{\infty} |Du(x_1, \ldots, y_i, \ldots, x_n)| dy_i$ for any i between 1 and n. Since we are interested ultimately interested in this case in bounding $||u||_{n/(n-1)}$ we use this to note:

$$|u(x)|^{n/(n-1)} \leq \prod_{i=1}^{n} (\int_{-\infty}^{\infty} |Du(x_1, \ldots, y_i, \ldots, x_n)| dy_i)^{1/(n-1)} \tag{24.7}$$

Now we start to integrate this with respect to x_1, x_2, etc. which will give us the integral of it on $\mathbb{R}^n$ by Fubini's theorem. Starting with x_1:

$$\begin{aligned} \int_{-\infty}^{\infty} |u(x)|^{n/(n-1)} dx_1 &\leq \int_{-\infty}^{\infty} \prod_{i=1}^{n} (\int_{-\infty}^{\infty} |Du| dy_i)^{1/(n-1)} dx_1 \\ &= (\int_{-\infty}^{\infty} |Du| dy_1)^{1/(n-1)} \int_{-\infty}^{\infty} \prod_{i=2}^{n} (\int_{-\infty}^{\infty} |Du| dy_i)^{1/(n-1)} dx_1 \end{aligned} \tag{24.8}$$

The point of the equality above is that $Du(x_1, \ldots, y_i, \ldots, x_n)$ for $i \neq 1$ depends on x_1, but when $i = 1$ there is no dependence so we could pull it off to the side through the integral in x_1 – one can say x_1 has already been integrated out. We recall now the generalized Holder's inequality: let $1 \leq p_1, \ldots, p_m \leq \infty$ with $\sum \frac{1}{p_i} = 1$, then for functions u_k with $u_k \in L^{p_k}$ we have

$$\int_U |u_1, \ldots, u_m| dx \leq \prod_{k=1}^{m} ||u_k||_{L^{p_k}}(U) \tag{24.9}$$

Here, we are considering $p_k = n-1$ for $k = 1, \ldots n-1$ and $u_k = (\int_{-\infty}^{\infty} |Du| dy_i)^{1/(n-1)}$ (of course since u is a test function these are in all the L^p spaces). So we have, continuing from above

$$\begin{aligned} &= (\int_{-\infty}^{\infty} |Du| dy_1)^{1/(n-1)} \prod_{i=2}^{n} (\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |Du| dy_i dx_1)^{1/(n-1)} \\ &= (\int_{-\infty}^{\infty} |Du| dy_1)^{1/(n-1)} (\prod_{i=2}^{n} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |Du| dy_i dx_1)^{1/(n-1)} \end{aligned} \tag{24.10}$$

We could pull the product inside because all the terms were raised to the same power. We repeat this now integrating with respect to x_2, pulling off to the side the term for which x_2 has already been integrated out:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|u(x)|^{n/(n-1)}dx_1dx_2$$
$$\leq (\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|Du|dx_1dy_2)^{1/(n-1)}\int_{-\infty}^{\infty}\prod_{i=3}^{n}(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|Du|dy_idx_1)^{1/(n-1)}(\int_{-\infty}^{\infty}|Du|dy_1)^{1/(n-1)}dx_2 \tag{24.11}$$

Applying the generalized Holder's inequality once again, with the product of the $n-2$ integrals above and the last term with the same factors as above:

$$\leq (\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|Du|dx_1dy_2)^{1/(n-1)}(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|Du|dy_1dx_2)^{1/(n-1)}\prod_{i=3}^{n}(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|Du|dy_idx_1dx_2)^{1/(n-1)} \tag{24.12}$$

Continuing in this fasion and again using Fubini's theorem we have:

$$\begin{aligned}\int_{\mathbb{R}^n}|u(x)|^{n/(n-1)}dx &\leq \prod_{i=1}^{n}(\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}|Du|dx_1\ldots dy_i\ldots dx_n)^{1/(n-1)}\\ &= (\int_{R^n}|Du|dx)^{\frac{n}{n-1}}\end{aligned} \tag{24.13}$$

Raising both sides to the power $\frac{n-1}{n}$, which is inequality preserving for $x^{\frac{n-1}{n}}$ is an increasing function, gives that $||u||_{L^{\frac{n}{n-1}}(\mathbb{R}^n)} \leq ||Du||_{L^1(\mathbb{R}^n)}$ which is the inequality claimed when $p=1$, with $C=1$ in this special case. Now for the general case that $1<p<n$ the idea/hope/guess is we can get the claim from the $p=1$ by considering $|u|$ raised to a power γ to be selected. Plugging it into the $p=1$ case and seeing what we get gives:

$$\begin{aligned}(\int_{\mathbb{R}^n}|u(x)|^{\gamma n/(n-1)}dx)^{\frac{n-1}{n}} &\leq \int_{\mathbb{R}^n}|D|u|^{\gamma}|dx = \gamma\int_{\mathbb{R}^n}|u|^{\gamma-1}|Du|dx\\ &\leq \gamma(\int_{\mathbb{R}^n}|u|^{(\gamma-1)\frac{p}{p-1}}dx)^{\frac{p-1}{p}}(\int_{\mathbb{R}^n}|Du|^p)^{\frac{1}{p}}\end{aligned} \tag{24.14}$$

Now since there is a $||Du||_{L^p(\mathbb{R}^n)}$ term on the RHS, if we could divide through by the other term which is an integral of u only we should (hopefully) get what we want. To get something clean we would want that other term to be a power of the the u integral on the LHS, so the integrands have to match. Hence we want to pick γ so that $\frac{\gamma n}{n-1} = (\gamma-1)\frac{p}{p-1}$. Solving for γ gives that $\gamma = \frac{p(n-1)}{n-p}$, and plugging this back into $\gamma n/(n-1)$ indeed gives $\frac{np}{n-p} = p^*$. Thus

$$(\int_{\mathbb{R}^n} |u|^{p^*} dx)^{\frac{n-1}{n} - \frac{p-1}{p}} \leq \gamma ||Du||_{L^p(\mathbb{R}^n)} \tag{24.15}$$

Finally, $\frac{n-1}{n} - \frac{p}{p-1}$ with some more easy algebra is equal to $\frac{1}{p^*}$, giving the claim.

□

As already mentioned, we next describe how this inequality extends to functions in $W^{1,p}(U)$ using the density and extension theorem from the last section. Note that below we don't assume or have for free that $u \in L^{p^*}$ but is something we show – considering that $p^* > p$ since $\frac{n}{n-p} > 1$ this means u is in a higher L^p space than we start out with. Considering that in a weak sense the functions in higher L^p spaces are more regular than those in lower ones this says $u \in W^{1,p}$ are more regular than functions just in L^p – no surprise of course, they have weak derivatives! For this reason the following is called the Sobolev embedding theorem, because it says that the inclusion map of $W^{1,p}(U) \hookrightarrow L^{p^*}(U)$ is well defined and continuous. It is also sometimes called the (subdimensional) Sobolev inequality, or conflating it with the above the G–N–S inequality:

Theorem 24.2. *Let U be a bounded, open subset of $\mathbb{R}^n$ with C^1 boundary. Assuming that $u \in W^{1,p}(U)$ then u is in fact in L^{p^*} with the estimate*

$$||u||_{L^p(U)} \leq C||u||_{W^{1,p}(U)} \tag{24.16}$$

where the constant C depends only on n, p, and U.

Proof: First, since ∂U is C^1 there is an extension $\overline{u} = Eu \in W^{1,p}(\mathbb{R}^n)$ such that $\overline{u} = u$, $\overline{u}$ has compact support, and $||\overline{u}||_{W^{1,p}(\mathbb{R}^n)} \leq C_1||u||_{W^{1,p}(U)}$ for some constant C_1 depending just on U, n, p where specifically in the statement of theorem 23.7 we fix a V such that $U \subset\subset V$. Next, since $\overline{u}$ has compact support from the density theorem (of smooth functions in Sobolev space) there exists functions $u_m \in C_c^\infty(\mathbb{R}^n)$ such that $u_k \to \overline{u}$ in $W^{1,p})(\mathbb{R}^n)$.

Now, by the G–N–S inequality we have $||u_j - u_k||_{L^{p^*}(\mathbb{R}^n)} \leq C_2||Du_j - Du_k||_{L^p(\mathbb{R}^n)}$ for all $j, k \geq 1$ for a constant C_2 depending just on n, p. Since the u_k converge to $\overline{u}$ in $W^{1,p})(\mathbb{R}^n)$ they are a Cauchy sequence in that space and hence, by the definition of the Sobolev norm, their weak derivatives are a Cauchy sequence in L^p. The inequality then gives that u_k are a Cauchy sequence in L^{p^*} too and so by completeness of this space their limit, which is still u, is in $L^{p^*}(\mathbb{R}^n)$. Now again by the G–N–S inequality since $u_k \to \overline{u}$ in both L^{p^*} and $W^{1,p}$ we have

$$||\overline{u}||_{L^{p^*}(\mathbb{R}^n)} \leq C_2||D\overline{u}||_{L^p(\mathbb{R}^n)} \tag{24.17}$$

Now, to wrap up because $\overline{u}$ is an extension of u we have its in L^{p^*} too and $||u||_{L^{p^*}(U)} \leq ||\overline{u}||_{L^{p^*}(\mathbb{R}^n)}$. By the extension theorem and defintion of Sobolev norm, $||D\overline{u}||_{L^p(\mathbb{R}^n)} \leq C_1||u||_{W^{1,p}(U)}$. All put together this gives that $||u||_{L^{p^*}(U)} \leq C_1C_2||u||_{W^{1,p}(U)}$. Considering the dependencies of the constants this gives the claim. □

Note that above we didn't assume that $u \in L^{p^*}$ but was something we showed – considering that $p^* > p$ since $\frac{n}{n-p} > 1$ this means u is in a higher L^p space than we initially supposed. Considering that in a weak sense the functions in higher L^p spaces are more regular than those in lower ones this says $u \in W^{1,p}$ are more regular than functions just in L^p – no surprise of course, they have weak derivatives! This inequality has an interesting geometric interpretation, when plugging in functions u that approximate the indicator function for U, with nonzero gradient only nearby the boundary of U. Then clearly the L^q norm of u has something to do with the volume of u, whereas its weak derivative is concentrated nearby the boundary of U. Then we get an inequality relating the volume of U and the area of its boundary – in other words, an isoperimetric inequality. This makes finding the optimal constant C above, called the Sobolev constant, related to the isoperimetric constant and interesting to geometers.

When $u \in W_0^{1,p}$ the extension theorem isn't necessary to use so one can see the following using Holder's inequality, where most importantly the $W^{1,p}$ norm on the RHS is replaced with just the intergral of Du as in the G–N–S inequality:

Theorem 24.3. *Assume U is a bounded, open subset of $\mathbb{R}^n$. Suppose that $u \in W_0^{1,p}(U)$ for some $1 \leq p < n$. Then*

$$||u||_{L^q(U)} \leq C||Du||_{L^p(U)} \tag{24.18}$$

with the constant depending just on p, q, n, U and for each $q \in [1, p^*]$.

An equality of the type above is often called a Poincare inequality since there is just the gradient term on the right – we will prove a similar inequality below that is probably more commonly understood to be the Poincare inequality. One might wonder what happens for functions in $W^{1,n}$ and in particular if u will be bounded considering that $p^* \to \infty$ as $p \to n$ and that functions in L^∞ are bounded but there are actually unbounded functions in $W^{1,n}((-1,1))$. This doesn't contradict the Sobolev inequality because in the G–N–S inequality γ, which we found to be the constant C, tends to infinity as $p \to n$. However, such functions will be of bounded mean oscillation. The BMO norm, given by a supremum over integrals of the form $\fint_{B(x,r)} |u(x) - u(y)| dy$, is bounded as in the Sobolev inequality which means basically that a function won't deviate too far from its average in any ball.

Now, you are probably wondering what happens for $W^{k,p}$ spaces for $k > 1$. By considering the $k-1$ derivatives of $u \in W^{k,p}$ and so on we can show the following:

Theorem 24.4. *Let U be a bounded open subset of $\mathbb{R}^n$ with a C^1 boundary. If $1 \leq k < n/p$ then $u \in L^q(U)$, where $1/q = 1/p - k/n$ and we have the estimate:*

$$||u||_{L^q(U)} \leq C||u||_{W^{k,p}(U)} \tag{24.19}$$

the constant C depending only on k, p, n, U.

Proof: As suggested, since $D^\beta u \in W^{1,p}(U)$ for any $|\beta| \leq k-1$ and $k < n/p$ implies $p < n$ we have

$$||D^\beta u||_{L^{p^*}(U)} \leq C_1(n, p, U)||u||_{W^{k,p}(U)} \tag{24.20}$$

In particular because this is true for any $|\beta| \leq k-1$ this implies that $u \in W^{k-1,p^*}(U)$ and that the (linear) inclusion map $W^{k,p}(U) \hookrightarrow W^{k-1,p^*}(U)$ is well defined and continuous. Recalling that $p^* = \frac{np}{n-p}$ and $kp < n$ where $k \geq 2$, we have $\frac{p}{n-p} < 1$ so that $p^* < n$. Hence we can iterate the argument to see $u \in W^{k-2,p^{**}}(U)$ with the following holding for any $|\gamma| \leq k-2$ (below of course the constant depends really on p, since p^* does):

$$||D^\gamma u||_{L^{p^{**}}(U)} \leq C_2(n, p^*, U)||u||_{W^{k-1,p^*}(U)} \tag{24.21}$$

Where p^{**}, the Sobolev conjugate of p^*, satisfies $\frac{1}{p^{**}} = \frac{1}{p^*} - \frac{1}{n} = \frac{1}{p} - \frac{2}{n}$ and we have for any γ with $|\gamma| \leq k-2$. Rearranging we have $p^{**} = \frac{np}{n-2p}$ and generally the ℓ-th Sobolev conjugate continuing in this manner (formally) is $\frac{np}{n-\ell p}$. So long as $\frac{p}{n-\ell p} < 1$ we can invoke the Sobolev inequality once more; now because $kp < n$ as long as $\ell < k$ we can invoke the inequality again so that we can safely use it k times. In particular $u \in L^q(U)$ for $q = \frac{np}{n-kp}$. The upshot is like above we have a chain of continuous inclusions from $W^{k,p}(U) \hookrightarrow W^{k-1,p^*}(U) \hookrightarrow \cdots \hookrightarrow W^{0,q}(U) = L^q(U)$. As the map $\iota : W^{k,p}(U) \to L^q(U)$ is continuous and linear its norm bounded by a constant C independent of u which may only depend on the parameters in the description of the spaces i.e. n, p, U (one may also be able to bound C in terms of the C_i from the $k = 1$ case of the Sobolev inequality above). □

Of course, this inequality didn't depend at all on u solving a PDE and was just for a general Sobolev function. On the other hand the "reverse" (being intentionally vague) inequality can be shown for a solution to a PDE sometimes, and this can be combined with the Sobolev inequality to get bounds of u in an L^p space then it was initially assumed to belong by a process that can be iterated to eventually show that u is pointwise bounded. This is in broad strokes the idea of the important technique known as Moser iteration. Now as discussed things go a little haywire in the subdimensional Sobolev inequality when we try to plug in $p = n$, and the fear might be that not much can be said when $p > n$ (or, $k/p > n$) but actually we have good results! One way to look at this is that actually when $p = n$ the theory doesn't completely break down, we are just lead by experience to expect u should be in L^∞ instead of the related space BMO of bounded mean oscillation functions. These functions are still not the worse, because as the name suggests they can't oscilate to wildly, so as $p > n$ one might expect that functions in $W^{1,p}$ are even better. This turns out to be the case. The following is called Morrey's inequality – at first it just involves C^1 functions but like before will be extended to general Sobolev functions in the expected way and is clearly quite a nice estimate:

Theorem 24.5. *Assume that $n < p \leq \infty$. Then there exists a constant $C = C(p, n)$ for which*

$$||u||_{C^{0,\gamma}(\mathbb{R}^n)} \leq C||u||_{W^{1,p}}(\mathbb{R}^n) \tag{24.22}$$

for all $u \in C^1(\mathbb{R}^n)$ where $\gamma = 1 - n/p$.

Proof: The first step is to get an estimate of the average oscillation of u, $⨍_{B(x,r)} |u(y)- u(x)|$, in terms of Du, for each ball $B(x, r)$. This can be thought of as a natural place to start because the average oscillation of a function, which we know from our discussion on the $p = n$ case should be controlled in terms of the sobolev norm of u, is akin to the Holder norm in that it measures in a quantitative way how much u can vary about a point. With this in mind we have by the fundamental theorem of calculus, where $w \in S(0, 1)$ and $0 < s \leq r$:

$$\begin{aligned} |u(x + sw) - u(x)| &= |\int_0^s \frac{d}{dt} u(x + tw)dt| \\ &= |\int_0^s Du(x + tw) \cdot wdt| \\ &\leq \int_0^s |Du(x + tw)|dt \end{aligned} \tag{24.23}$$

Integrating this on $S(0,1)$ and switching integrals gives:

$$\begin{aligned}\int_{S(0,1)} |u(x+sw)-u(x)|dS(w) &\leq \int_0^s \int_{S(0,1)} |Du(x+tw)|dS(w)dt \\ &= \int_0^s \int_{S(x,t)} \frac{|Du(y)|}{t^{n-1}} dS(y)dt \text{ (where } y = x+tw) \\ &= \int_{B(x,s)} \frac{|Du(y)|}{|x-y|^{n-1}} dy \\ &\leq \int_{B(x,r)} \frac{|Du(y)|}{|x-y|^{n-1}} dy\end{aligned} \tag{24.24}$$

At the last step using $s \leq r$. On the other hand, setting $z = x+sw$ gives $\int_{S(0,1)} |u(x+sw) - u(x)|dS(w) = \frac{1}{s^{n-1}} \int_{S(x,s)} |u(z)-u(x)|dS(z)$. Putting these together gives:

$$\int_{S(x,s)} |u(z)-u(x)|dS(z) \leq s^{n-1} \int_{B(x,r)} \frac{|Du(y)|}{|x-y|^{n-1}} dy \tag{24.25}$$

Integrating this from $0, r$ then gives:

$$\int_{B(x,r)} |u(y)-u(x)|dy \leq \frac{r^n}{n} \int_{B(x,r)} \frac{|Du(y)|}{|x-y|^{n-1}} dy \tag{24.26}$$

Dividing through by r^n and multiplying by the correct normalizing constants (in terms of area of the unit ball) then gives for some constant C_1 just depending on dimension:

$$⨍_{B(x,r)} |u(y)-u(x)|dy \leq C_1 \int_{B(x,r)} \frac{|Du(y)|}{|x-y|^{n-1}} dy \tag{24.27}$$

This inequality can be phrased in terms of so–called Riesz transforms and in fact the Sobolev embedding theorem above, when $p < n$ can be approached by this inequality (the original approach was similar). To get a hold of the Holder norm of u, we use this to get pointwise bounds of u – this is reasonable because the average of a constant function is itself and the integral above is over y. So, writing $u(x) = u(x) - u(y) + u(y)$ we have by the triangle inequality that:

$$|u(x)| \leq ⨍_{B(x,r)} |u(y)-u(x)|dy + ⨍_{B(x,r)} |u(y)|dy \tag{24.28}$$

The second term above, writing $u(y) = 1 \cdot u(y)$, can be estimated in terms of the L^p norm of u using Holder's inequality and 24.27:

$$|u(x)| \leq C_1 \int_{B(x,r)} \frac{|Du(y)|}{|x-y|^{n-1}} dy + C_2 ||u||_{L^p(B(x,r))} \tag{24.29}$$

where C_2 is the factor we mentioned from Holder's inequality combined with the volume of the ball (so just depends on n, p). By Holder's inequality again and since $B(x,r) \subset \mathbb{R}^n$ we can further estimate:

$$|u(x)| \leq C_1 ||Du||_{L^p(\mathbb{R}^n)} \Big(\int_{B(x,1)} \frac{1}{|x-y|^{(n-1)\frac{p}{p-1}}} dy\Big)^{\frac{p-1}{p}} + C_2 ||u||_{L^p(\mathbb{R}^n)} \tag{24.30}$$

Now, for $p > n$ we have $(n-1)\frac{p}{p-1} < n$. Since $n-1$ spheres of radius r have area on the order of r^{n-1} we see by the coarea formula, which gives integral over the ball centered at x can be done integrating over nested spheres centered at x, that $\int_{B(x,1)} \frac{1}{|x-y|^{(n-1)\frac{p}{p-1}}} dy$ is finite since the integral $\int_0^r x^{-\gamma}$ is finite for $\gamma < 1$. We can estimate this explicitly so for future reference we have (done for general r, not just $r = 1$):

$$\begin{aligned} \Big(\int_{B(x,r)} \frac{1}{|x-y|^{(n-1)\frac{p}{p-1}}} dy\Big)^{\frac{p-1}{p}} &= \Big(\int_0^r \int_{S(x,s)} \frac{1}{|x-y|^{(n-1)\frac{p}{p-1}}} dS(y) ds\Big)^{\frac{p-1}{p}} \\ &= \Big(\int_0^r \frac{C_3 s^{n-1}}{s^{(n-1)\frac{p}{p-1}}} dS(y) ds\Big)^{\frac{p-1}{p}} \\ &= \Big(\int_0^r \frac{C_3}{s^{(n-1)\frac{p}{p-1}-(n-1)}} dS(y) ds\Big)^{\frac{p-1}{p}} \\ &\leq C_4 \big(r^{n-(n-1)\frac{p}{p-1}}\big)^{\frac{p-1}{p}} \\ &= C_4 r^{1-n/p} \end{aligned} \tag{24.31}$$

Where C_3 above is just the area of the unit sphere in dimension n and C_4 is it divided by $1 - n/p$ all raised to the power $(p-1)/p$ so just depends on n, p. Bounding all these constants and their multiples with one big constant C_5, we thus have

$$|u(x)| \leq C_5 ||u||_{W^{1,p}(\mathbb{R}^n)} \tag{24.32}$$

Which is a nice start, and of course is needed to estimate the Holder norm because the α Holder norm is given by $||u||_{L^\infty} + [u]_\alpha$. Emboldened by this, with the defintion of the Holder norm in mind, we want to try out bounding $|u(x) - u(y)|$ for two points $x \neq y$. Denoting by $W = B(x,r) \cap B(y,r)$, where $r = |x-y|$, and writing

$u(x) - u(y) = u(x) - u(z) + u(z) - u(y)$ then as before

$$|u(x) - u(y)| \leq \fint_W |u(x) - u(z)| dz + \fint_W |u(y) - u(z)| dz \tag{24.33}$$

With our estimate for the sup of u in mind along with the estimate for the multiplicative term coming from the $\frac{1}{|x-y|^{(n-1)\frac{p}{p-1}}}$ integral, we have

$$\fint_W |u(x) - u(z)| dz, \ \fint_W |u(y) - u(z)| dz \leq C_5 r^{1-n/p} ||Du||_{L^p(\mathbb{R}^n)} \tag{24.34}$$

Using that $r = |x-y|$ and we thus have $|u(x) - u(y)| \leq C_5 |x-y|^{1-n/p} ||Du||_{L^p(\mathbb{R}^n)}$ giving that, along with the pointwise bound 24.28 above (bounded in terms of the Sobolev norm of u), that u is in $C^{0,1-n/p}$ with $[u]_{1-n/p} \leq C_5 ||Du||_{L^p(\mathbb{R}^n)} \leq C_5 ||u||_{W^{1,p}(\mathbb{R}^n)}$. This gives the estimate $||u||_{C^{0,\gamma}(\mathbb{R}^n)} \leq C ||u||_{W^{1,p}}(\mathbb{R}^n)$ completing the proof. □

Now, we remember that two functions in L^p space. and hence $W^{k,p}$ space are considered the same if (and only if) they differ by a set of measure zero (i.e. agree almost everywhere/a.e.) – this was a necessary sacrifice to make the L^p spaces normed since their norm is in terms of integration. With this in mind, in the statement below we say that u^* is a version of u if $u = u^*$ almost everywhere:

Theorem 24.6. *Let U be a bounded, open subset of $\mathbb{R}^n$ and suppose ∂U is C^1. Assume $n < p \leq \infty$ and $u \in W^{1,p}(U)$. Then u has a version $u^* \in C^{0,1-n/p}(\overline{U})$ with the estimate*

$$||u^*||_{C^{0,1-n/p}(\overline{U})} \leq C ||u||_{W^{1,p}(U)} \tag{24.35}$$

where the constant C depends just on p, n, U.

Proof: First, we suppose $p < \infty$. As before, we extend u to $\mathbb{R}^n$: since ∂U is C^1 there is an extension $\overline{u} = Eu \in W^{1,p}(\mathbb{R}^n)$ such that $\overline{u} = u$, $\overline{u}$ has compact support, and $||\overline{u}||_{W^{1,p}(\mathbb{R}^n)} \leq C_1 ||u||_{W^{1,p}(U)}$ for some constant C_1 depending just on U, n, p. Next, we approximate $\overline{u}$ with $u_k \in C_c^\infty(\mathbb{R}^n)$ in $W^{1,p}(\mathbb{R}^n)$. Because $p > n$ and smooth functions are C^1 we have from Morrey's theorem that $||u_i - u_j||_{C^{0,1-n/p}(\mathbb{R}^n)} \leq C_2 ||u_i - u_j||_{W^{1,p}(\mathbb{R}^n)}$ for a constant just depending on n, p, U. Because the u_k converge in $W^{1,p}(\mathbb{R}^n)$ they are a Cauchy sequence in the Sobolev norm, which implies by the bound they are a Cauchy sequence in $C^{0,1-n/p}(\mathbb{R}^n)$ and so converge to a function u^* in this Holder space. Since $u_k \to \overline{u}$ in $W^{1,p}(\mathbb{R}^n)$ as well u^* must be a version of $\overline{u}$ – of course since this is an extension of u, u^* restricted to U is also a version of u. Because $||u_k||_{C^{0,1-n/p}(\mathbb{R}^n)} \leq C_2 ||u_k||_{W^{1,p}(\mathbb{R}^n)}$ for all k $||u^*||_{C^{0,1-n/p}(\mathbb{R}^n)} \leq C_2 ||\overline{u}||_{W^{1,p}(\mathbb{R}^n)} \leq C_1 C_2 ||u||_{W^{1,p}(U)}$ giving the claim. The case

$p = \infty$ follows easily from the $p < \infty$ case using Holder's inequality and that U is bounded. $\square$

The estimate above can be called the superdimensional Sobolev inequality. Important for applications to the regularity theory of PDE, we next show the following for $W^{k,p}$ with general k:

Theorem 24.7. *Let U be a bounded open subset of $\mathbb{R}^n$, with a C^1 boundary, and suppose $u \in W^{k,p}(U)$. If $k > n/p$, then $u \in C^{k-[\frac{n}{p}]-1,\gamma}(\overline{U})$ where*

$$\gamma = \begin{cases} [\frac{n}{p}] + 1 - \frac{n}{p}, & \text{if } \frac{n}{p} \text{ is not an integer} \\ \text{any number in } (0,1) & \text{otherwise} \end{cases} \tag{24.36}$$

Finally, we have the estimate

$$||u||_{C^{k-[\frac{n}{p}]-1,\gamma}(\overline{U})} \leq C||u||_{W^{k,p}(U)} \tag{24.37}$$

where the constant C depends only on k, p, n, γ, U.

Proof: First we consider the case that $\frac{n}{p}$ is not an integer. Considering ℓ such that $\ell < n/p < \ell + 1$ (i.e. $\ell = [\frac{n}{p}]$) we see $\ell < k$, so we may consider $D^\alpha u \in W^{\ell,p}(U)$. Since $\ell < n/p$ we may use the general Sobolev embedding theorem, theorem 24.4, to see for all $|\alpha| \leq k - \ell$

$$||D^\alpha u||_{L^r(U)} \leq C||D^\alpha u||_{W^{\ell,p}(U)} \leq C||u||_{W^{k,p}(U)} \tag{24.38}$$

where $\frac{1}{r} = \frac{1}{p} - \frac{\ell}{n}$ and C depends on n, p, U. This implies that

$$||u||_{W^{k-\ell,r}(U)} \leq C||u||_{W^{k,p}(U)} \tag{24.39}$$

Now, $q = \frac{np}{n-\ell p}$ and in particular is greater than n. Thus we can apply Morrey's inequality, to see $D^\alpha u$ all belong to $C^{0,1-\frac{n}{r}}(\overline{U})$ for all $|\alpha| \leq k - \ell - 1$. Because $\frac{1}{r} = \frac{1}{p} - \frac{\ell}{n}$, $1 - \frac{n}{r} = 1 - \frac{n}{p} + \ell = [\frac{n}{p}] + 1 - \frac{n}{p}$ so that $u \in C^{k-[\frac{n}{p}]-1,[\frac{n}{p}]+1-\frac{n}{p}}(\overline{U})$ with norm bounded in terms of $||u||_{W^{k,p}(U)}$ by the previous theorem.

Now suppose that $\frac{n}{p}$ is an integer. Set $\ell = [\frac{n}{p}] - 1 = \frac{n}{p} - 1$. Again since $\ell < n/p$ we may use the general Sobolev embedding theorem to see $u \in W^{k-\ell,r}(U)$ where now $r = \frac{pn}{n-p\ell} = n$. Using this and Holder's inequality (and that U is bounded) $u \in W^{k-\ell,q}(U)$ for all $q < n$ as well, or in other words so that $D^\alpha u \in W^{1,q}(U)$ for all such q and $|\alpha| \leq k - \ell - 1$. Hence by the Sobolev embedding theorem we have $D^\alpha u \in L^{q^*}(U)$ and so $u \in W^{k-\ell,q^*}(U)$ where, since q is let to vary all the way up to n, q^* can be taken to be arbitrarily large. Applying Morrey's inequality as in the first case we thus have $\gamma = 1 - \frac{n}{q^*}$ can taken on an arbitrary range of values, giving the claim. □

We end this section with a refinement of the Sobolev embedding theorem which, like all the results in this section, are quite important. The Rellich–Kondrachov compactness theorem, below, says the image of bounded sets under the inclusion map in the Sobolev embedding theorem is actually a (pre)compact set similar to the Arzela–Ascoli theorem – this gives us the ability to take converging subsequences from arbitrary sequences which is a very strong tool. The insight perhaps of why such a statement should be true is that functions in $W^{1,p}$ satisfying a uniform bound, so have their first weak derivatives bounded, might be close enough in spirit to C^1 functions with unifromly bounded derivatives to say they are an equicontinuous family of sorts. And indeed, the method of the proof essentially is to show that mollifications of such functions are equicontinuous so that the Arzela–Ascoli theorem can be applied to them, which gives the statement as the mollification parameter tends to zero (I might fill this proof out in a later update):

Theorem 24.8. *Assume U is a bounded open subset of $\mathbb{R}^n$ and ∂U is C^1. Suppose $1 \leq p < n$. Then*

$$W^{1,p}(U) \subset\subset L^q(U) \tag{24.40}$$

for each $1 \leq q < p^*$.

By the usual argument using Holder's inequality and that U is bounded if $u \in W^{1,p}$ for $p \geq n$ it is also in $W^{1,s}$ with $s < n$. Using this and Holder's inequality once more, we have as a corollary:

Corollary 24.9. *Assume U is a bounded open subset of $\mathbb{R}^n$ and ∂U is C^1. Suppose $1 \leq p < \infty$. Then*

$$W^{1,p}(U) \subset\subset L^p(U) \tag{24.41}$$

To remind ourselves on why such theorems are useful we give the following consequence, which is the more "proper" version of Poincare's inequality. Below we denote by $(u)_U$ the average of u over U.

Theorem 24.10. *Let U be a bounded, connected, open subset of $\mathbb{R}^n$ with a C^1 boundary and assume $1 \leq p \leq \infty$. Then there exists a constant C depending only on n, p, U such that*

$$||u - (u)_U||_{L^p(U)} \leq C||Du||_{L^p(U)} \tag{24.42}$$

Proof: Suppose by contradiction there is a seuqence $u_k \in W^{1,p}(U)$ satisfying $||u_k - (u_k)_U||_{L_p(U)} \geq k||Du_k||_{L^p(U)}$ for each $k > 1$. We normalize the sequence by setting

$$v_k = \frac{u_k - (u_k)_U}{||u_k - (u_k)_U||_{L_p(U)}} \tag{24.43}$$

Then the average of all the v_k is zero, their $L^p(U)$ norm is one, and $||Dv_k||_{L^p(U)} < 1/k$. Taking a converging subseqeunce v_{k_j} in L^p by the compactness theorem we have it converges to a function v with zero average and $||v||_{L^p(U)} = 1$. On the other hand for an arbitrary test function ϕ we have

$$\int_U v\phi_{x_i} dx = \lim_{k_j \to \infty} \int_U v_{k_j} \phi_{x_i} = - \lim_{k_j \to \infty} \int_U v_{k_j, x_i} \phi dx = 0 \tag{24.44}$$

Where the last equality is true since $||Dv_k||_{L^p(U)} < 1/k$ (and note we aren't using stronger than L^p convergence). Because ϕ is arbitrary this implies $v = 0$ since its average is zero, contradicting that $||v||_{L^p(U)} = 1$. □

Such arguments as above are known as "compactness–contradiction" arguments and are used quite often (Terrence Tao has a whole book about them, [22]). By using Poincare's inequality on the unit ball and then considering $v = u(x + ry)$ for $U = B(x,r)$, we have the following corollary, the Poincare inequality for a ball.

Corollary 24.11. *Assume that* $1 \leq p \leq \infty$. *Then there exists a constant* C, *depending only on* n *and* p, *such that*

$$||u - (u)_{B(x,r)}||_{L^p(B(x,r))} \leq Cr||Du||_{L^p(B(x,r))} \tag{24.45}$$

In particular, with $p = 1$ we have

$$⨍_{B(x,r)} |u - (u)_{B(x,r)}| dy \leq Cr ⨍_{B(x,r)} |Du| dy \tag{24.46}$$

(We just divide through by the volume of the ball of radius r.) By Holder's inequality then one can see that the BMO norm $[u]_{BMO(\mathbb{R}^n)} = \sup_{B(x,r) \subset \mathbb{R}^n} ⨍ |u - (u)_{B(x,r)}| dy$, of u is bounded in terms of the $W^{1,n}$ norm of u as discussed above.

As a final topic for this section let's mention that under the Fourier transform weak derivatives behave as classical ones do, and so integrability of weak derivatives imply strong integral bounds on the Fourier transform of a Sobolev function, which imply it must decay quickly. More precisely one can show:

Theorem 24.12. *Let k be a nonnegative integer.*

(1) A function $u \in L^2(\mathbb{R}^n)$ belongs to $H^k(\mathbb{R}^n)$ if and only if

$$(1 + |\xi|^k)\hat{u} \in L^2(\mathbb{R}^n) \tag{24.47}$$

(2) In addition, there exists a positive constant C such that

$$\frac{1}{C}||u||_{H^k(\mathbb{R}^n)} \leq ||(1 + |\xi|^k)\hat{u}||_{L^2(\mathbb{R}^n)} \leq C||u||_{H^k(\mathbb{R}^n)} \tag{24.48}$$

for each $u \in H^k(\mathbb{R}^n)$.

For noninteger s this gives a natural way to define Sobolev spaces H^s of course. And in fact, one can approach the Sobolev inequalities when $p = 2$ via the Fourier transform and Cauchy–Schwarz in a very easy way. Doing this right would perhaps require maybe a bit more of a digression on Fourier analysis than our time allows (or at least its not traditionally a part of the course), but we can quickly give an idea. For instance if we suppose that $u \in C^\infty$ is in Schwartz space – this is a space of rapidly decaying functions, essentially for which we may take fourier transforms and inverse fourier transforms of u and its derivatives freely and it includes test functions – its easy to prove a bound of the form $||u||_{C^k(\mathbb{R}^n)} \leq C||u||_{H^{2s}(\mathbb{R}^n)}$ when $2s > n + k$. For instance when $k = 0$ from the defintion of Fourier transform we have

$$\begin{aligned} |u(x)| &\leq \frac{1}{\sqrt{(2\pi)^n}} \int |\hat{u}(\xi)| d\xi \\ &= \frac{1}{\sqrt{(2\pi)^n}} \int_{\mathbb{R}^n} (1 + |\xi|)^{-s}(1 + |\xi|)^s |\hat{u}(\xi)| d\xi \end{aligned} \tag{24.49}$$

which implies by Cauchy–Schwarz, that

$$|u(x)|^2 \leq \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} (1 + |\xi|)^{-2s} d\xi \int_{\mathbb{R}^n} (1 + |\xi|)^{2s} |\hat{u}(\xi)|^2 d\xi \tag{24.50}$$

The first integrand is finite since $2s > n$, and by an equivalent form of the theorem above where $(1 + |\xi|^k)$ is replaced with $(1 + |\xi|)^k$ we have that $\int_{\mathbb{R}^n} (1 + |\xi|)^{2s} |\hat{u}(\xi)|^2 d\xi$ is bounded above by $||u||_{H^{2s}}$ giving the statement after taking square roots of both sides. The case $k > 0$ is similar. Depending on conventions some authors will have this written with s halved and depending on how one defines Sobolev spaces, for instance in terms of the completion of Schwartz space under the Sobolev norm, gives differentiability for s large enough by a density argument. Of course for most pratical applications equivalent this is equivalent to what we get from Morrey's equality, modulo the Holder norm control. A nice exposition on this, and a quick use of elliptic theory in geometry to prove the Hodge theorem, is given in chapter 6 in [24]

– this might be covered in the exercise section. Another good resource is [12], which will be referenced more below.

25. Weak solutions via Hilbert space methods

There's still more one could say about Sobolev spaces that would be of use to us but perhaps its best to cut to the chase and see Sobolev spaces in action with a PDE. Our target, to remind ourselves, is to consider the linear partial differential operator L defined by:

$$Lu = -\sum_{i,j=1}^{n} a^{ij}(x)u_{x_i x_j} + \sum_{i} b^i(x)u_{x_i} + c(x)u \tag{25.1}$$

on a bounded domain U with C^1 boundary (our usual setting for Sobolev spaces with the extension theorem and the Sobolev inequalities in mind). Here L will always be uniformly elliptic, in that $\sum_{i,j=1}^{n} a^{ij}(x)\xi_i\xi_j \geq \theta|\xi|^2$ for some fixed $\theta > 0$ with the a^{ij} symmetric in i and j, just like what we assumed in our discussion of the continuity method. For our methods it will actually be more convenient in this section to represent/consider Lu in "divergence form" (where L in the form above is in nondivergence form):

$$Lu = -\sum_{i,j=1}^{n} (a^{ij}(x)u_{x_i})_{x_j} + \sum_{i} b^i(x)u_{x_i} + c(x)u \tag{25.2}$$

The reason for this is basically to make integration by parts steps work out which, from the defintion of weak derivative, one can imagine is important in the use of Sobolev spaces. On the other hand L in nondivergence theorem is better for showing maximum principles. If the coefficients are all C^1 (which they often are in applications) then the two forms are equivalent in that one can write an elliptic operator L in divergence form as one in nondivergence form and vice versa. In particular we will be interested to consider the following problem:

$$\begin{cases} Lu = f \text{ in } U \\ u = 0 \text{ on } \partial U \end{cases} \tag{25.3}$$

We can consider the case $u = g$ on the boundary for g in, say, $C^k(\overline{U})$ similarly, by subtracting Lg off of f and solving that problem. Our method of approaching this problem in this section will be by the Hilbert space method. To explain what we mean, its easy to see that a classical solution u to $Lu = f$ with L in divergence form will satisfy, using integration by parts:

$$\int_U \sum_{i,j=1}^{n} (a^{ij}(x)u_{x_i})v_{x_j} + \sum_i b^i(x)u_{x_i}v + c(x)uvdx = \int_U fvdx \tag{25.4}$$

Where v is any test function on U. Considering that the notion of weak derivative is defined via integration by parts, we'll define a weak solution $u \in H_0^1(U)$ to $Lu = f$ in this method when the relation above is true for all test functions v – we take $u \in H_0^1(U)$ because these functions are the ones which should consider as vanishing along the boundary by the trace theorem. Since the completion of the test functions on U in the space $H^1(U)$ happens to be $H_0^1(U)$, there is no harm actually, in that its equivalent, to say a weak solution u to 25.2 is a function $u \in H_0^1(U)$ for which 25.4 is true for all $v \in H_0^1(U)$.

Now, what the Hilbert? The spaces $H^k(U) = W^{k,2}(U)$, $H_0^k(U) = W_0^{k,2}(U)$ can all be seen to be Hilbert spaces essentially from the defintion of the norms on Sobolev spaces and that $L^2(U)$ is a Hilbert space. There the inner product we recall is given by $(f, g) = \int_U fgdx$, and similarly the inner product on $H_k(U)$ is given by $(f, g) = \sum_{|\alpha| \leq k} \int_U D^\alpha f D^\alpha gdx$. The point of noting this, and that why we define our weak solutions to lay in the Hilbert space $H_0^1(U)$ (the "least regular" Sobolev space that 25.4 makes sense in – we'll expound on this particular point shortly) instead of $W^{1,p}$ for some other value of p is that we can formulate u being a weak solution neatly as being a function $u \in H_0^1(U)$ for which for all $v \in H_0^1(U)$ we have

$$B[u, v] = (f, v) \tag{25.5}$$

where B is the bilinear form (one can check) defined by the left hand side of the above, and (f, v) is the L^2 inner product of f and v. Then since f is fixed, the right hand side of this is a linear functional in terms of v so, since B reminiscent of an inner product (hey, its bilinear!) we wonder if maybe a theorem like the Riesz representation theorem applies: it says if ℓ is a continuous linear functional on a Hilbert space H with inner product $\langle \cdot, \cdot \rangle$ then there exists a vector v_ℓ, called the representation vector, so that $\ell(w) = \langle v_\ell, w \rangle$. If such a statement is true for B as well, then the representing vector v_ℓ would be our solution u! The Lax–Milgram theorem, below, tells us gives us conditions to check about B for this to be true:

Theorem 25.1. *Where H is a Hilbert space, assume that $B : H \times H \to \mathbb{R}$ is a bilinear map for which there exists constants $\alpha, \beta > 0$ such that*

$$|B[u,v]| \leq \alpha ||u|| ||v|| \tag{25.6}$$

and

$$\beta ||u||^2 \leq B[u,u] \tag{25.7}$$

Then for each bounded linear functional f on H there exists a unique element $u \in H$ so that $B[u,v] = f(v)$.

Proof: The first condition says essentially that B is continuous, and the second says that B is nondegnerate in a sense – these basically give that B is close enough to the inner product on H to get the job done. Now to prove this first we note for a fixed $u \in H$, the map $v \to B[u,v]$ is a bounded linear functional on H (by condition (1)) so the Riesz representation theorem says that $B[u,v] = (w,v)$ for some $w \in H$. Ranging over $u \in H$ we then define a map $A : H \to H$ for which

$$B[u,v] = (Au, v) \tag{25.8}$$

Now, on the other hand by the Riesz representation theorem we know that $f(v) = (w,v)$ for some $w \in H$. If we knew that there was some u so that $(Au, v) = (w, v)$, then we would be done. This u would be unique too, because if there were u_1, u_2 so that $B[u_1, v] = f(v)B[u_2, v]$ for each v then $B[u_1 - u_2, v] = 0$. Letting $v = u_1 - u_2$ by item (2) it follows $u_1 = u_2$. In other words, what we need to know is that A is surjective, which we proceed to do.

By bilinearity of B its easy to see that A is linear, and because $||Au||^2 = (Au, Au) = B[u, Au] \leq \alpha ||u|| ||Au||$ (item (1) again) so that A is bounded with $||A|| \leq \alpha$. Employing the second condition, we have that $\beta ||u||^2 \leq B[u,u] = (Au, u) \leq ||Au|| ||u||$ so that A is injective. This same line of arguing also implies the range of A is closed: if Au_i is a Cauchy sequence in H then by the lower bound u_i is also a Cauchy sequence, so converges to an element u_0. Then since A is bounded its easy to see that $Au_i \to Au_0$. The point of this is that if the range $R(A)$ of A is not all of H, then one can find a nonzero element of $w \in R(A)^\perp$. But on the other hand $0 < \beta ||w||^2 \leq B[w,w] = (Aw, w) = 0$ giving a contradiction. □

Note that if B were symmetric i.e. $B[v,w] = B[w,v]$ then we could use it, under the conditions above, to define a new inner product on H so could just plug directly into the Riesz representation theorem. This doesn't need to be the case in the theorem above though and assuming this wouldn't be good enough for our hopeful applications (for instance, if the b^i are nonzero). The next step of course is to see if we can apply the Lax–Milgram theorem to our B, again defined in terms of 25.4. We next prove the following energy estimates:

Theorem 25.2. *When a^{ij}, b^i, c are all uniformly bounded on U there exist constants $\alpha, \beta > 0$ and $\gamma \geq 0$ such that*

$$|B[u,v]| \leq \alpha ||u||_{H_0^1(U)} ||v||_{H_0^1(U)} \tag{25.9}$$

and

$$\beta ||u||^2_{H_0^1(U)} \leq B[u,u] + \gamma ||u||^2_{L^2(U)} \tag{25.10}$$

for all $u, v \in H_0^1(U)$.

Proof: Bounding the sum of the absolute value of all the coefficients above by, say, C we have

$$|B[u,v]| \leq \sum_{i,j}^{n} C \int_U |Du||Dv| dx + \sum_{i=1}^{n} C \int_U |Du||v| dx + C \int_U |u||v| dx \tag{25.11}$$

Using the Cauchy-Schwarz inequality on each of these terms gives that $|B[u,v]| \leq \alpha ||u||_{H_0^1(U)} ||v||_{H_0^1(U)}$ for some constant α. For the second estimate, by the ellipticity assumption on L (you knew it was coming eventually!) we have

$$\begin{aligned} \theta \int_U |Du|^2 dx &\leq \int_U \sum_{i,j=1}^{n} (a^{ij}(x) u_{x_i}) u_{x_j} \\ &= B[u,u] - \int_U \sum_i b^i(x) u_{x_i} 2 + c(x) u^2 dx \\ &\leq B[u,u] + C \int_U |Du||u| dx + C \int_U u^2 dx \end{aligned} \tag{25.12}$$

By the Cauchy inequality (in this case this is called the Peter–Paul inequality, because we are robbing Peter (dividing by 4ϵ) to pay Paul (multipling by ϵ)) we have

$$\int_U |Du||u| dx \leq \epsilon \int_U |Du|^2 dx + \frac{1}{4\epsilon} \int_U u^2 dx \tag{25.13}$$

where $\epsilon > 0$ – this type of inequality is used constantly in these types of things to "absorb" some terms into others, as we are about to illustrate. Picking $\epsilon > 0$ so that $C\epsilon < \theta/2$. we have from the bound above then that

$$\frac{\theta}{2} \int_U |Du|^2 dx \leq B[u,u] + C' \int_U u^2 dx \tag{25.14}$$

for a constant C which is potentially much larger than C. Since $u \in H_0^1(U)$ we recall that we didn't need the extension theorem to prove the Sobolev embedding theorem, obtaining the version of the Poincare inequality in theorem 24.3. If $n > 2$ we can apply this in our case along with the fact that the Sobolev conjugate of 2 is greater than 2 to see $||u||_{L^2(U)} \leq ||Du||_{L^2(U)}$ to get, the full H_0^1 norm of u is bounded by $\int_U |Du|^2$ up to a constant, completing the statement. We can take care of the corner case that $n = 2$ by bounding $W^{1,p}$ for $p < 2$ by Holder's inequality using U is bounded, and then can proceed as usual to give the result. □

It makes some sense that the ellipticity conditon plays a role here, because it says in a sense that the PDE is nonsingular at the highest order. In the analogy with solving a linear algebra problem, which is the angle we are essentially taking in this endeavour, we should have our "matrix" (i.e. L) be nonsingular in some sort of sense, because its not like we can solve $Ax = b$ all the time for general matrices A. The ellipticity will be used in other parts of the theory, for instance for the maximum principle which actually has a hand (in one route at least) in the existence theory as we'll see in the next section.

Also, note that the energy bounds are in terms of the H_0^1 norm of u – we didn't and probably shouldn't be able to get energy estimates (particularly, the lower bounds) in terms of $H_0^k(U)$ for arbitrary $k > 1$ in complete generality, assuming for the sake of argument u was also in these spaces. The reason this is so is because the next theorem then gives an existence result using Lax–Milgram, and if we could always find weak solutions in H_0^k using it for sufficiently large k have on our hands a very regular weak solution u by Morrey's inequality and subsequent embedding result. But if the coefficients of L are all very regular too this would imply the right side f is differentiable as well. We definitely have circumstances in the theorem below though where we can find a weak solution for f only in L^2 however – so in other words the energy estimates and observations facts dictate we should only be working in H_0^1 spaces for small k when finding a weak solution with $f \in L^2$. When f is more smooth/regular we will later on show our weak solution is more regular, this is basically what people mean by regularity theory.

Now continuing on the energy estimates give us almost what we want, except that γ in the statement above might be nonzero. But it does tell us we can solve the problem 25.3 if we peturb L appropriately:

Theorem 25.3. *There is a number $\gamma \geq 0$ such that for each $\mu \geq \gamma$ and each function $f \in L^2(U)$ there exists a unique weak solution $u \in H_0^1(U)$ of the problem 25.3 with $L_\mu u = Lu + \mu_u$.*

Proof: Denoting the bilinear form for the operator L_μ by $B_\mu[u,v] = B[u,v] + \mu(u,v)$, one can see that $B_\mu[u,u] = B[u,u] + \mu||u||^2_{L^2(U)}$. If $\mu > \gamma$ then, the energy estimates above say that the conditions of Lax–Milgram are satisfied for B_μ. Applying it with the functional $\ell(v)$ given by $\ell(v) = (f,v)$ where f is the RHS in the PDE problem gives the claim. □

This is a really nice start of course, although its not quite a existence theorem for 25.3 for our original operator L. In the next section we dig a bit deeper into our bag of functional analysis to squeeze more out of this result and say something about the problem we were originally asking about, giving a dichotomy which gives (or perhaps more appropriately points a path forward to) a general existence theorem.

26. The Fredholm alternative: existence from uniqueness

In functional analysis for PDE an important class of operators are the compact ones:

Definition 26.1. *A bounded linear operator $K : X \to Y$ is called compact provided for each bounded sequence $\{u_k\}_{k=1}^\infty \subset X$, the sequence $\{Ku_k\}_{k=1}^\infty$ is precompact in Y; that is, there exists a subsequence of $\{Ku_k\}_{k=1}^\infty$ which converges in Y.*

For instance, the Rellich–Kondrachov compactness theorem says that the inclusion map $W^{1,p}(U) \hookrightarrow L^{p^*}(U)$ is compact. Projection onto a finite dimensional subspace will also be compact; of course the identity map is a nonexample. For our purposes an important theorem about compact operators is the Fredholm alternative; first we give the more technical statement and, while all parts of it are useful, we then zero in on what we really need to remember for now:

Theorem 26.1. *Let $K : H \to H$ be a compact linear operator on a Hilbert space H. Then*

(1) $N(I-K)$ is finite dimensional,
(2) $R(I-K)$ is closed,
(3) $R(I-K) = N(I-K^*)^\perp$
(4) $N(I-K) = \{0\}$ if and only if $R(I-K) = H$, and
(5) dim $N(I-K) =$ dim $N(I-K^*)$

The proof isn't particularly hard and can be found in appendix D of [5]. The most important point for us in this section is item (4), restated in plain language:

Corollary 26.2. *For a compact operator $K : H \to H$ exactly one of the following is true:*

1. *For each $f \in H$, the equation $u - Ku = f$ has a unique solution, or*
2. *The homogenous equation $u - Ku = 0$ has solutions $u \neq 0$.*

Now, if we write $A = I - K$, this can be rephrased as $Au = f$ has a unique solution always or $Au = 0$ has nonzero solutions. If we imagine A to be a matrix, this should remind you exactly on the rank nullity theorem! This is how I like to remember it. Now we proceed to apply it to our setting. There's a bit more one can say which can be read off of the general Fredholm alternative as you can find in [5], but its less important. As we'll see after the proof, its actually stronger than it might look at first glance:

Theorem 26.3. *Where the coefficients of L are all uniformly bounded, precisely one of the following statement holds:*

1. *For each $f \in L^2(U)$, problem 25.3 has a unique weak solution, or*
2. *There exists a weak solution $u \neq 0$ of the problem when $f = 0$ (i.e. the homogenous problem).*

Proof: Choosing $\mu = \gamma$ in theorem 25.3 we have for each g a unique function $u \in H_0^1(U)$ such that $B_\gamma[u, v] = (g, v)$ for all $v \in H_0^1(U)$; we write this u as $L_\gamma^{-1} g$. Now, we see that if u is a solution to 25.3 with our original operator L with RHS f, then $B_\gamma[u, v] = (\gamma u + f, v)$ for all $v \in H_0^1(U)$ i.e. if

$$u = L_\gamma^{-1}(\gamma u + f) \tag{26.1}$$

Now, its easy to see that $L_\gamma^{-1}(af + bg) = aL_\gamma^{-1} f + bL_\gamma^{-1} g$ because both B and the inner product are bilinear. Using this we can rewrite the equation above as

$$u - Ku = h \tag{26.2}$$

where $Ku = \gamma L_\gamma^{-1} u$, $K : L^2(U) \to H_0^1(U) \subset L^2(U)$ and $h = L_\gamma^{-1} f$. Then if K is compact we have the Fredholm alternative holds for it which implies our statement. To see this we work backwards: if the first possibility in theorem 26.2 is true then there is always a solution u to $u - \gamma L_\gamma^{-1} u = L_\gamma^{-1} f$, so that $u = L_\gamma^{-1}(\gamma u + f)$, which means that $B_\gamma[u, v] = (\gamma u + f, v)$. Then since $B_\gamma[u, v] = B[u, v] + \gamma(u, v)$ $B[u, v] = (f, v)$ so that by defintion u is a weak solution of 25.3. Of course the second possibility is ruled out by uniqueness. If the second possibility of the Fredholm alternative holds then since $h = L_\gamma^{-1} f = 0$ only if $f = 0$, the second item of the statement du jour holds (and the first doesn't) giving the claim.

To see K is compact (as an operator from $L^2(U)$ to itself) we see from the energy estimates above and our choice of γ, with the proof of theorem 25.3 in mind, that for $u = L_\gamma^{-1} g$ we have

$$\beta ||u||^2_{H^1_0(U)} \leq B_\gamma[u,u] = (g,u) \leq ||g||_{L^2(U)} ||u||_{L^2(U)} \leq ||g||_{L^2(U)} ||u||_{H^1_0(U)} \tag{26.3}$$

In the third inequality using Cauchy Schwarz and the fourth using the defintion of Sobolev space. Since $Kg = \gamma L_\gamma^{-1} g = \gamma u$ this implies that $||Kg||_{H^1_0(U)} \leq C||g||_{L^2(U)}$ for some constant C. If we consider a bounded sequence $f_i \in L^2(U)$ this implies the sequence Kf_i will be bounded in $H^1_0(U)$. Now the Rellich–Kondrachov theorem comes into play, to say that the sequence Kf_i will have a convergent subsequence! Thus K is compact and we are done. □

Now what's so great about this? The point is to solve the problem 25.3 for a general f we just have to check $Lu = 0$, $u = 0$ along ∂U is only solved by $u = 0$, and we have already seen ways to do this: for $L = \Delta$ it follows using the maximum principle. So, if we can show all solutions to the above problem are sufficiently regular to apply a maximum principle and that a maximum principle holds, then the general problem is solvable for L – this is a beautiful facet of the theory: existence of weak solution, uniqueness, and regularity are all intertwined, as if by an occult hand [17] – of course other uniqueness methods which employ less regularity, such as energy methods, may apply which would let us sidestep regularity etc. to get uniqueness. It also turns out that there is a weak maximum princple for weak solution to elliptic PDE (this still qualifies as standard material but probably isn't as well known as the usual maximum principle), as discussed in chapter 8 of [6]. Anyway, as one might imagine though for the regularity theory and maximum principle theorems below to hold for a given elliptic operator L we must impose additional conditions on it, but these will turn out to not be unreasonable.

Going back to the "hard" Fredholm theorem, theorem 26.1, obviously all the statements are phrased not in terms of K persay but $I - K$. Defining (in case this defintion is new to you) the cokernel of an operator $A : X \to Y$ as $Y/R(A)$, we say that an operator A is **Fredholm** if $\dim \operatorname{coker}(A), \dim \ker(A)$ are finite and its range is closed. The difference $\dim \ker(A) - \dim \operatorname{coker}(A)$ is called the index of A. The Fredholm alternative can be seens as an assertion about Fredholm operators, because items (3) and (5) of theorem 26.1 say that for a compact operator K $I - K$ is Fredholm. Now the Fredholm alternative we showed for PDE involving elliptic operators L didn't really go by showing that L itself was a Fredholm operator. It turns out that L, and elliptic operators in a broader sense, can be seen to be Fredholm directly which we'll touch on in the section below on the parametrix method.

27. The parametrix method, and a taster of Atiyah–Singer

Before continuing onto the regularity theory for elliptic PDE via bootstrapping as presented as in [5] we sketch out the parameterix method/viewpoint for elliptic differential operators as well as some consequences of it, for instance as a different way to show regularity from bootstrapping. Of course, this section is sketchy on details and isn't material which will be on the examination, but I think its good to know about. A nice introduction to it is given in [18], and one with more details and generality is given in chapter 3 of [12], which is a book filled to the brim with interesting material. It can be thought of as a generalization of the technique we discussed concerning the Malgrange–Ehrenpreis theorem, where roughly a Fourier analysis approach is taken to find the Green's function of a general constant coefficient partial differential operator – there were some complications, but this is good enough for us right now. The idea specifically was to use Fourier anaylsis to trade derivatives for multiplication by the frequency variable, which turns the PDE into an algebraic equation which we can then invert and then at least formally obtain a solution (and so Green's function) by using the inverse Fourier transform.

Now, for a PDE with varying coefficients this idea isn't quite as helpful, for instance because while the Fourier transform which trades differentiation for multiplication by the frequency variable the mirror statement is true when we consider the Fourier transform of xu. So, for instance where we denote the frequency variable by ξ and u is a single variable function:

$$\mathcal{F}(\partial_x u(x) + xu(x))(\xi) = i(y\mathcal{F}(u)(\xi) + \partial_\xi \mathcal{F}(u)(\xi)) \tag{27.1}$$

The Fourier series here didn't really do anything, so one might be worried the Fourier transform might not really help at all. However, the Fourier transform idea is still of great use with some further assumptions. Recalling our defintion of a linear PDO of order m as a differential operator which can be written as:

$$P(x, D) = \sum_{|\alpha|\leq m} a_\alpha(x) D^\alpha \tag{27.2}$$

Then we say its principal symbol $\sigma(x,\xi)$ is, with the Fourier transform in mind:

$$\sigma(x,\xi) = \sum_{|\alpha|=m} a_\alpha(x)\xi^\alpha \tag{27.3}$$

where ξ is a vector in $\mathbb{R}^n$ and ξ^α is defined by multipication of its components in the obvious way. An elliptic operator, which we define for order even higher than 2, is then one for which its principal symbol is never zero for any x and $\xi \neq 0$; note the defintion one finds in an "official" treatment will likely be a bit more complicated, for instance in [12] there is additionally a growth assumption in ξ on the principal symbol. Loading on some more terminology said another way we wish its inverse to belong to a so-called symbol class for which we can take the inverse Fourier transform safely.

The idea then is we define an "almost inverse" of P via the inverse Fourier transform of $\frac{1}{\sigma(x,\xi)}$ or something close to it, and use it to study P. This might seem naive but turns out to actually be quite a strong method. Here the ellipticity gurantees we can take the inverse transform, after cutting out the origin, and a nice facet of this approach is its very clear how it plays an important role – of course when L is elliptic as defined in the previous section it will also be elliptic in this sense. This is called the/a **Parametrix** for P; a more precise defintion can be given of course and one finds that there is some ambiguity allowed in the defintion, up to an infinitely smoothing operator which we'll define shortly.

For instance for a 2nd degree elliptic operator L as we've been dealing with this "almost inverse" can be (thought of being similar to) the fundamental solution $\epsilon(x, x_0)$ for the operator $\sum_{i,j=1}^{n} a^{ij}(x_0)\xi_i\xi_j$ for a fixed x_0, which when $a^{ij} = \delta^{ij}$ is just the fundamental solution for the regular Laplacian. Specializing mometarily to this case we denote its action on functions via convolution by $S_{x_0}(f)$ i.e.

$$S_{x_0}(f)(x) = \int \epsilon(x-y, x_0)f(y)dy \tag{27.4}$$

Because S above was defined via the inverse Fourier transform of a function in ξ that was not a polynomial, which would correspond to a symbol that came from an actual differential operator, it is called a pseudodifferential operator. Denoting $L = L(x,D)$ we then have:

$$Id = L(x,D)S_{x_0} + (L(x_0,D) - L(x,D))S_{x_0} \tag{27.5}$$

Writing Tx_0 for $(L(x_0,D) - L(x,D))S_{x_0}$, we can then rewrite the equation above as $Id - T_{x_0} = L(x,D)S_{x_0}$. Similarly note we can write $Id = S_{x_0}L(x,D) + S_{x_0}(L(x_0,D) -$

$L(x, D)) = S_{x_0}L(x, D) + T'_{x_0}$. As an aside from the first equality we see if we can invert $Id - T_{x_0}$, which is possible when T_{x_0} is sufficiently small (it happens that like for real numbers the geometric series identity $(Id - T_{x_0})^{-1} = \sum T_{x_0}^i$ is well defined and holds when T_{x_0} is small) we can then proceed to obtain the Green's function for $L(x, D)$ – "small" here of course depends partly on the space the operator T_{x_0} is defined on and here it turns out we can make it smaller than one if one restricts to functions in Sobolev space in a sufficiently small domain.

What we want to emphasize more though is that, replacing S_{x_0} with the prescription in [12] and similarly T_{x_0}, T'_{x_0} by T, T', these will be "infinitely smoothing operators" in that functions it outputs will be smooth despite this not necessarily being the case for functions in its domain. This is hopefully easy to imagine because convolution against the kernel $\epsilon(x, x_0)$ of S will be a C^2 function, even when the input is merely C^α as shown in [6] (so regularity improves). Using this with the second identity then we can easily deduce regularity of solutions to $Lu = f$ if f is smooth: $u = SP(x, D)u + T'u = Sf + T'u$ which will be smooth. A downside of this from our perspective is that u should be in H^2 to define Lu sensibly, but our method for producing solutions (by the Hilbert space method) produces solutions which are apriori in H^1_0. This issue can be sidestepped/subsumed using that L is Fredholm from, say $H^2 \to L^2$. This turns out to be a consequence of T' being compact due to the Sobolev embedding theorem, to be compact – the proof of this is actually quite short (see lemma 5.1 in [12]) and mostly follows from some general functional analysis – the hard work is arguably getting L into the arrangement above with its parametrix. This then implies that $Lu = f$ is very often solvable, as long as f isn't in the (finite dimensional) kernel of the adjoint L^* of L. And of course if we can show $L^*u = 0$ is only solvable by $u = 0$ for some problem, we can always solve $Lu = f$. Of course the same is true for more general elliptic operators P.

This fact also has special importance in the Atiyah–Singer index theorem, which is a major focus of [12]. One can "globalize" the discussion above to consider elliptic operators on manifolds, and these will continue to be Fredholm. One may then define the analytic index of a elliptic operator as just its index as a Fredholm operator. On the other hand one can define the topological index of it, which is not obviously related the the analytic index. The index theorem then says that these two indices are the same; one can then get many interesting results just by plugging different elliptic operators into the machine.

28. Regularity theory for elliptic PDE ala bootstrapping

We now discuss, alternate to the parametrix method sketched above, regularity of solutions to $Lu = f$ via "elliptic bootstrapping." Considering just the Laplacian the rough observation is that if $u \in C^2(U)$ and $\Delta u = f$ where $f \in C^1(U)$, then u should actually have 3 derivatives because a second order differential operator applied to it, which we think of as using up two derivatives of u, still has another derivative left to give – of course Δu doesn't involve the full Hessian of u, but this isn't rigorous anyway. This is why we call it bootstrapping, because we are arguing that u should have more derivatives than we initially suppose using the equation. Extending this analogy to weakly differentiable functions then we might expect that if $f \in H^k$ then u should generally be in H^{k+2}, which is exactly what happens and basically what we'll spend our time showing. Similar to how it helps with the Sobolev inequalities an arguably slicker way to go about proving regularity, or at least the interior estimates, is via Fourier analysis (so Fourier methods can be useful in at least two different ways, counting the Parametrix method as Fourier analysis). A good source for this is the final chapter of the book [24], mentioned already before.

First we need a way to show a function u (say measurable on a domain U) actually has weak derivatives, and that will be to consider (and estimate) its difference quotient: we define the i-th difference quotient of a function u of size h on a set $V \subset\subset U$

$$D_i^h u(x) = \frac{u(x + he_i) - u(x)}{h} \tag{28.1}$$

for $x \in V$ and $0 < |h| < dist(V, \partial U)$. Similarly we define $D^h u = (D_1^h u, \ldots D_n^h u)$. Then we have the following theorem; the first part is used often below but the second part is more important:

Theorem 28.1.

(1) *Suppose* $1 \leq p < \infty$ *and* $u \in W^{1,p}(U)$. *Then for each* $V \subset\subset U$ *we have*

$$||D^h u||_{L^p(U)} \leq C||Du||_{L^p(U)} \tag{28.2}$$

for some constant C *and all* $0 < |h| < \frac{1}{2} dist(V, \partial U)$.

(2) *Conversely, suppose that for some* $1 < p < \infty$ $u \in L^p(V)$ *and there exists a constant* C *such that*

$$||D^h u||_{L^p(V)} \leq C \tag{28.3}$$

for all $0 < |h| \leq \frac{1}{2} dist(V, \partial U)$. *Then*

$$u \in W^{1,p}(V) \text{ with } ||Du||_{L^p(V)} \leq C \tag{28.4}$$

Proof: Beginning with the first statement, we start off with supposing u is smooth. Then by the fundamental theorem of calculus and change of variables we have:

$$u(x + he_i) - u(x) = h \int_0^1 u_{x_i}(x + the_i)dt \tag{28.5}$$

This obviously implies

$$|u(x + he_i) - u(x)| \leq |h| \int_0^1 |Du(x + the_i)|dt \tag{28.6}$$

By Cauchy–Schwarz applied to the right hand side, using as usual $|Du(x + the_i)| = 1 \cdot |Du(x + the_i)|$, we then have (writing $D_i^h u(x)$ for the LHS above):

$$|D_i^h u(x)| \leq C(\int_0^1 |Du(x + the_i)|^p dt)^{1/p} \tag{28.7}$$

for some constant C which can be bounded above independent of $|h|$. Since for $p \geq 1$ $x \to x^p$ is increasing we may raise both sides of the inequality to the power p and preserve it, so following with integration in x we have:

$$\int_V |D_i^h u(x)|^p dx \leq C \int_V \int_0^1 |Du(x+the_i)|^p dt dx = C \int_0^1 \int_V |Du(x+the_i)|^p dx dt \leq C \int_U |Du(x)|^p dx \tag{28.8}$$

Where, in the last inequality we used that for a fixed t $\int_V |Du(x+the_i)|^p \leq \int_U |Du(x)|^p$ and the t integral was over the unit interval. Using the definition of $D^h u$ and the triangle inequality gives item (1) for smooth functions, and by the standard extension and approximation argument we have it for $u \in W^{1,p}(U)$.

Moving onto (2), let $\phi \in C_c^\infty(V)$ and not for sufficiently small h (to stay in the domains of defintion of our functions) that by simple change of variables:

$$\int_V u(x) \frac{\phi(x + he_i) - \phi(x)}{h} dx = -\int_V \frac{u(x) - u(x - he_i)}{h} \phi(x) dx \tag{28.9}$$

or, in other words:

$$\int_V u D_i^h \phi dx = -\int_V (D_i^{-h} u) \phi dx \tag{28.10}$$

which is what you expect from the defintion of weak derivative via integration by parts. The assumption that $||D^h u||_{L^p(V)} \leq C$ is bounded uniformly in h and $p > 1$ gives, as a consequence of the Banach–Alaoglu theorem, that there is a function

$v_i \in L^p(V)$ and a sequence $h_k \to 0$ such that $D_i^{-h_k}u \rightharpoonup v_i$ i.e. weakly – this of course doesn't imply $D_i^{-h_k}u$ coverge to anything in particular in the L_p topology. Spelling this out a bit more the Banach–Alaoglu theorem says that bounded balls in the dual of a Banach space are compact in the weak-$*$ topology, and how we are applying it here is that we are considering $D_i^{-h_k}u$ as linear functionals on $(L_p)^*$. i.e. in L_p^{**}, given the natural map $L_p \to L_p^{**}$ by evaluation. Now because we are assuming that $p > 1$, we have $L_p^{**} = L_p$, which is to say its reflexive. The topology on L_p though considered as the double dual of L_p is different from the regular one and there are actually a couple different topologies referenced in the statement of Banach–Alaoglu. First the notion of boundedness in the theorem is in reference to the norm on the dual space given by operator norm. To describe it recall $L_p^* \simeq L_q$, and the map from $\iota : L_p \to L_q^*$ is given by $\iota(f)(g) = \int fg$ – we conflate f with $\iota(f)$ when speaking of it as being in the double dual. Then the norm on L_p as a dual space of L^q is just the L^p norm by Holder's inequality, so the assumption says the difference quotients of u are all bounded thought of as functionals on L_q. The weak–$*$ topology is the topology with respect to weak convergence, so in our setting Banach–Alaoglu and reflexivity say for all $\ell \in (L^p)^*$ $\ell(D_i^{-h_k}u)$ converges as numbers and there is an element $v_i \in L^p$ such that $\lim\limits_{h_k\to 0} \ell(D_i^{-h_k}u) \to \ell(v_i)$ which works for every ℓ. Below we consider functionals ℓ given by integration against a test funtion (or its derivative), which will be in L_q for any $1 < p < \infty$.

Again we used $p > 1$ because L_p is reflexive and what we are trying to prove actually turns out to be false for $p = 1$; it actually isn't too hard to cook up an example and might be a homework problem. Anyway, weak convergence of course isn't quite as nice as regular convergence, but it works for us because the notion of weak derivative is defined in terms of integration against test functions which give linear functionals on L^p in the natural way – that the defintion of weak derivative behaves well with respect to only weak convergence we see now is certainly a notch in its belt. Unraveling the defintion of weak convergence gives:

$$\begin{aligned}\int_V u\phi_{x_i}dx = \int_U u\phi_{x_i}dx = \lim_{h_k\to 0}\int_U uD_i^h\phi dx \text{ (can see this using Cauchy–Schwarz)} \\ = -\lim_{h_k\to 0}\int_V (D_i^{-h_k}u)\phi(x)dx \\ = -\int_V v_i\phi dx\end{aligned} \tag{28.11}$$

Above we swtiched back and forth between V and U using that the support of ϕ was in V along with the fact that its difference quotient might merely be in U. This gives that $v_i = u_{x_i}$ in the weak sense, ranging over $i = 1, \dots n$, and since each of these are in L^p it gives Du both exists weakly and is in $L^p(V)$, so that $u \in W^{1,p}(V)$.

$\square$

With this tool in hand, we now begin our discussion of interior regularity where as usual we suppose that U is a C^1 and bounded domain. Below note we suppose $u \in H^1(U)$ as opposed to $H_0^1(U)$ because the behavior of u near the boundary doesn't matter here. Also, note the similarity of the form of the estimate below to the Schauder estimates – of course we will also show that the inequality makes sense to discuss in the first place.

Theorem 28.2. *Assume $a^{ij} \in C^1(U)$, $b^i, c \in L^\infty(U)$, $f \in L^2(U)$, and $u \in H^1(U)$ is a solution to $Lu = f$ in U where L is uniformly elliptic. Then actually $u \in H^2_{loc}(U)$ and for each open $V \subset\subset U$ we have*

$$||u||_{H^2(V)} \leq C(||f||_{L^2(U)} + ||u||_{L^2(U)}) \tag{28.12}$$

the constant C depending only on V, U, and the coefficients of L.

Proof: Fixing V and an open set W such that $V \subset\subset W \subset\subset U$ we may consider a smooth cutoff function ψ supported in W which is equal to 1 in V; we will use it shortly. Now, since we want to show second weak derivatives of u exist with the difference quotients theorem we showed above in mind we want to control the difference quotients of the u_{x_i}. To do so we seperate the highest order term in $B[u, v] = (f, v)$ out from the other parts to write:

$$\sum_{i,j=1}^{n} \int_U a^{ij} u_{x_i} v_{x_j} dx = \int_U \tilde{f} v dx \tag{28.13}$$

where $\tilde{f} = f - \sum_{i=1}^{n} b^i u_x - cuv$. Here of course since u is a weak solution we are allowed to pick any $v \in H_0^1(U)$. Our choice is to pick, for a fixed $k \in \{1, \dots, n\}$ and $|h|$ sufficiently small (depending on the distance of W to ∂U):

$$v = -D_k^{-h}(\psi^2 D_k^h u) \tag{28.14}$$

This should seem like a reasonable choice because we could try to peal difference quotients off of it and onto u_{x_i} using change of variables. With this choice of v we estimate the LHS of 28.13 to see

$$
\begin{aligned}
&\sum_{i,j=1}^{n}\int_U a^{ij}u_{x_i}v_{x_j}dx = -\sum_{i,j=1}^{n}\int_U a^{ij}u_{x_i}(D_k^{-h}(\psi^2 D_k^h u))_{x_j}dx \\
&= \sum_{i,j=1}^{n}\int_U D_k^h(a^{ij}u_{x_i})(\psi^2 D_k^h u)_{x_j}dx \\
&\text{(using weak derivatives and difference quotients commute, change of variables)} \\
&= \sum_{i,j=1}^{n}\int_U (a^{ij,h}D_k^h u_{x_i} + (D_k^h a^{ij})u_{x_i})(\psi^2 D_k^h u)_{x_j}dx \\
&\text{(using } D_k^h(vw) = v^h D_k^h w + w D_k^h v\text{)} \\
&= \sum_{i,j=1}^{n}\int_U a^{ij,h}D_k^h u_{x_i} D_k^h(u_{x_j})\psi^2 \\
&+ \sum_{i,j=1}^{n}\int_U a^{ij,h}D_k^h u_{x_i} D_k^h(u)2\psi\psi_{x_j} + (D_k^h a^{ij})u_{x_i}D_k^h(u_{x_j})\psi^2 + (D_k^h a^{ij})u_{x_i}D_k^h(u)2\psi\psi_{x_j} \\
&= A_1 + A_2
\end{aligned}
\tag{28.15}
$$

Where the second to last equality was what we get from cracking open $(\psi^2 D_k^h u)_{x_j}$ and above by g^h we mean $g(x+he_k)$. The first term, A_1, can be bounded below using that L is elliptic:

$$A_1 \geq \theta \int_U \psi^2 |D_k^h Du|^2 dx \tag{28.16}$$

This term is good for us because it involves L^2 norms of difference quotients of Du. By doing some really crude estimating and that ψ, a^{ij} are fixed so there pointwise values, derivatives, and difference quotients (by mean value theorem) are bounded by some number independent of h, u we have

$$|A_2| \leq C\int_U \psi|D_k^h Du||D_k^h u| + \psi|D_k^h Du||Du| + \psi|D_k^h u||Du|dx \tag{28.17}$$

The first and second terms above could potentially be problematic because they also involves difference quotients of Du so they might somehow cancel out the A_1 term and leave us with nothing to say. The thing to do in this case is use Peter–Paul inequality and absorb. Applying it for any $\epsilon > 0$ to the first and second terms, with

Cauchy–Schwarz applied to the third and using the support of ψ is in W:

$$|A_2| \leq \epsilon \int_U \psi^2 |D_k^h Du|^2 dx + \frac{C}{\epsilon} \int_W |D_k^h u|^2 + |Du|^2 dx + C \int_W |D_k^h u|^2 + |Du|^2 dx \quad (28.18)$$

for an appropriate constant C. Picking $\epsilon = \theta/2$ and using that there is a constant C for which $||D^h u||_{L^2(W)} \leq C||Du||_{L^2(W)}$ we have

$$|A_2| \leq \frac{\theta}{2} \int_U \psi^2 |D_k^h Du|^2 dx + C \int_U |Du|^2 dx \quad (28.19)$$

(where of course C was adjusted) giving

$$\sum_{i,j=1}^n \int_U a^{ij} u_{x_i} v_{x_j} dx \geq \frac{\theta}{2} \int_U \psi^2 |D_k^h Du|^2 dx - C \int_U |Du|^2 dx \quad (28.20)$$

Phew! But now we need to estimate the RHS of 28.13. Since we have a lower bound for the LHS what we want/what we could use is an upper bound for the RHS. From the defintion of $\tilde{f}$ for a general v we have:

$$|(\tilde{f}, v)| \leq C \int_U (|f| + |Du| + |u|)|v| dx = C \int_U |f||v| + |Du||v| + |u||v| dx \quad (28.21)$$

Our choice of v, $-D_k^{-h}(\psi^2 D_k^h u)$, is like two derivatives of u (two difference quotients, at least) so one can imagine that the Peter–Paul and absorbtion trick will be needed once more to incorporate it with the estimate of the LHS of 28.13. With this in mind, we estimate the integral of $|v|^2$:

$$\int_U |v|^2 dx \leq C \int |D(\psi^2 D_k^h u)|^2 dx \quad (28.22)$$

where we used that $\psi^2 D_k^h u \in H^1$ and the difference quotient estimate. So since the weak derivative commutes with difference quotient, the product rule, and the triangle inequality we have:

$$\int_U |v|^2 dx \leq C \int_W |D_k^h u|^2 + \psi^2 |D_k^h Du|^2 dx \quad (28.23)$$

(The constant C above involves now a bound on the derivative of ψ.) Using the difference quotient estimate again we have:

$$\int_U |v|^2 dx \leq C \int_U |Du|^2 + \psi^2 |D_k^h Du|^2 dx \quad (28.24)$$

Using Peter–Paul now on each of the terms in $C\int_U |f||v|+|Du||v|+|u||v|dx$ we have for any $\epsilon > 0$:

$$(\tilde{f}, v) \leq \epsilon \int_U \psi^2 |D_k^h Du|^2 dx + \frac{C}{\epsilon}\int_U f^2 + u^2 + |Du|^2 dx \tag{28.25}$$

Where we applied it to isolate out the term involving difference quotients of Du. Taking $\epsilon = \theta/4$ then we have (for our choice of v)

$$\begin{aligned} &\frac{\theta}{2}\int_U \psi^2 |D_k^h Du|^2 dx - C\int_U |Du|^2 dx \leq \sum_{i,j=1}^n \int_U a^{ij} u_{x_i} v_{x_j} dx = (f, v) \\ &\leq \frac{\theta}{4}\int_U \psi^2 |D_k^h Du|^2 dx + C' \int_U f^2 + u^2 + |Du|^2 dx \end{aligned} \tag{28.26}$$

Rearranging and using that $\psi = 1$ on V gives:

$$\int_V |D_k^h Du|^2 dx \leq C\int_U f^2 + u^2 + |Du|^2 dx \tag{28.27}$$

Since the right hand side of the above inequality is independent of h we can then say from the theorem from the start of the section that the x_k weak derivative of Du exists and is in L^2 and, ranging over k, that $u \in H_2(V)$. Adding $u^2 + |Du|^2$ to both sides of the inequality above and ranging over k we also have the inequality

$$||u||_{H^2(V)} \leq C(||f||_{L^2(W)} + ||u||_{H^1(W)}) \tag{28.28}$$

which is almost, but not quite, the inequality in the statement because there it only involves the L^2 norm of u. To get this we use now $v = \psi^2 u$ for a new cutoff function equal to 1 on W. Then by the same techniques we can show (homework problem, no doubt!):

$$\int_U \psi^2 |Du|^2 dx \leq C\int_U f^2 + u^2 dx \tag{28.29}$$

This gives that $||u||_{H^1(W)} \leq C(||f||_{L^2(U)} + ||u||_{L^2(U)})$, which finally completes the proof. □

Notice that in the proof above that elliptiicity played a role, although in a more subtle way than in the parametrix method. It makes sense that ellipticity is involved in showing regularity because, as we saw in the wave equation (decidedly nonelliptic), a solution need not be more regular than its input data and in fact can be less regular; of course right now we are on the boundary so to speak of our regularity assumptions to define a weak solution, but we are about to iterate the argument. We've touched on this already but just to reiterate the cutoff was there to ensure all functions above were well defined even when considering their difference quotients for h small enough

($x + he_i$ might not be in U even if x is but h is too big, and an upper bound for $|h|$ that is valid tends to zero as we approach ∂U). We used ψ^2, as opposed to ψ, so that when we took a derivative of ψ^2 there was still a ψ left which was identically equal to 1 on V. This was used towards the end of the proof to use so we could write $\psi^2|D_k^h Du|^2 = |D_k^h Du|^2$ in V and hence conclude $u \in H_2(V)$.

Knowing that $u \in H^2(V)$ for any $V \subset\subset U$ already tells us something pretty interesting: we can now in the sense of weak derivatives directly calculate Lu and by integration by parts see that $B[u,v] = (Lu, v)$ for any test function v. So, if u solves $Lu = f$ weakly one can see that $Lu = f$ almost everywhere which is arguably more satisfying than our notion of weak solution originally gave. But we can often do even better by basically iterating the argument above, which is the content of the following statement:

Theorem 28.3. *Let m be a nonnegative integer, and assume $a^{ij}, b^i, c \in C^m(U)$ and $f \in H^m(U)$. Then if $u \in H^1(U)$ is a weak solution of the elliptic PDE $Lu = f$ in U we have in fact $u \in H^{m+2}_{loc}(U)$ and for each $V \subset\subset U$ we have the estimate*

$$||u||_{H^{m+2}(V)} \leq C(||f||_{H^m(U)} + ||u||_{L^2(U)}) \tag{28.30}$$

with the constant C depending only on m, U, V, L.

Proof: The basic idea is that if u solves $Lu = f$, then any valid derivative $D^\alpha u$ of u also satisfies an elliptic equation so the theorem above can be applied. Indeed we proceed by induction, with the case $m = 0$ corresponding to the theorem above. Supposing the statement is true up to $m = k$, if $u \in H^1(U)$ is a weak solution with $a^{ij}, b^i, c \in C^{k+1}(U)$ and $f \in H^{k+1}(U)$, then u will belong to $H^{k+2}_{loc}(U)$. So, if $a^{ij}, b^i, c \in C^{k+2}(U)$ and $f \in H^{k+2}(U)$ we need to show u, which we already know to be in $H^{k+2}_{loc}(U)$, will be in fact be in $H^{k+3}_{loc}(U)$ with the bound above for $m = k+1$. Fix as in the statement below V, W so that $V \subset\subset W \subset\subset U$ and let α be a multiindex with $|\alpha| = k+1$. Considering $\tilde{v} \in C_c^\infty(W)$ write $v = (-1)^{|\alpha|} D^\alpha \tilde{v}$ which, since u is a weak solution to $Lu = f$, satisfies $B[u,v] = (f,v)$. Writing this out fully (and cancelling out $(-1)^{|\alpha|}$) gives:

$$\int_U \sum_{i,j=1}^n (a^{ij}(x) u_{x_i})(D^\alpha \tilde{v})_{x_j} + \sum_i b^i(x) u_{x_i} D^\alpha \tilde{v} + c(x) u D^\alpha \tilde{v} dx = \int_U f D^\alpha \tilde{v} dx \tag{28.31}$$

Because all the coefficients are in $C^{k+2}(U)$, $f \in H^{k+2}(U)$, and $u \in H^{k+2}_{loc}(U)$ we can integrate by parts repeatedly to move D^α in the above off of $\tilde{v}$ and onto u and the coefficents of L. With the product rule in mind we find that $B[D^\alpha u, \tilde{v}] = (\tilde{f}, \tilde{v})$ where

$$\tilde{f} = D^\alpha f - \sum_{\beta<\alpha} C(\alpha,\beta)[-\sum_{i,j=1}^{n}(D^{\alpha-\beta}a^{ij}D^\beta u_{x_i})_{x_j} + \sum_{i=1}^{n} D^{\alpha-\beta}b^i D^\beta u_{x_i} + D^{\alpha-\beta}cD^\beta u] \tag{28.32}$$

In other words, we brought over the terms in the product rule after integrating by parts where the derivatives didn't all fall over onto u to the RHS. The point is that $D^\alpha u$ weakly solves $Lw = \tilde{f}$, where here we are thinking of the lower order weak derivatives of u as functions independent of $D^\alpha u$. We can then apply the theorem above then to say that D^α is actually in $H^2(V)$ so that, ranging over α, $u \in H^{k+3}(V)$. Furthermore we have

$$||D^\alpha u||_{H^2(V)} \leq C(||\tilde{f}||_{L^2(U)} + ||D^\alpha u||_{L^2(U)}) \tag{28.33}$$

Using the boundedness of the coefficients of L and the defintion of $\tilde{f}$ we have by Cauchy–Schwarz that

$$||\tilde{f}||_{L^2(U)} + ||D^\alpha u||_{L^2(U)} \leq C(||f||_{H^{k+2}(U)} + ||u||_{H^{k+2}(U)}) \tag{28.34}$$

By the triangle inequality then, we see

$$||u||_{H^{k+3}(V)} \leq C(||f||_{H^{k+1}(U)} + ||u||_{H^{k+2}(U)}) \tag{28.35}$$

By induction we can in turn bound the second term above by $||f||_{H^k(U)} + ||u||_{L^2(U)}$, completing the argument after suitably adjusting the constant. □

As a corollary of elliptic bootstrapping and Morrey's inequality we have the following upshot, which is what is used the most at the end of the day.

Corollary 28.4. *Assume $a^{ij}, b^i, c, f \in C^\infty(U)$ and $u \in H^1(U)$ is a weak solution of the elliptic PDE $Lu = f$ in U. Then $u \in C^\infty(U)$ and solves the PDE $Lu = f$ in the classical sense.*

We end this section with saying something about regularity up to the boundary. From the work above the bootstrapping gives as many weak derivatives as f and the coefficients of L allow in all of U but the point is we can give bounds that don't deteriorate as we approach the boundary with more assumptions. Similar to the interior case one first shows the following:

Theorem 28.5. *Suppose $a^{ij} \in C^1(\overline{U})$, $b^i, c \in L^\infty(U)$ and $f \in L^2(U)$. Then if $u \in H^1_0(U)$ is a weak solution of the elliptic boundary value problem $Lu = f$ in U,*

$u = 0$ on ∂U where ∂U is C^2 and U is bounded then $u \in H^2(U)$ with the estimate

$$||u||_{H^2(U)} \leq C(||f||_{L^2(U)} + ||u||_{L^2(U)}) \tag{28.36}$$

where the constant C depends only on U and the coefficients of L.

Proof: (sketch) Note that in the above we are now assuming u vanishes along the boundary of U in the trace sense and we need more regularity of ∂U than we've usually been assuming (that its just C^1). First one can work in a half ball, say the unit ball centered at origin intersected with the halfspace $\{x \mid x_n > 0\}$ one which we pick a cutoff function equal to 1 on $V = B(0, 1/2) \cap \{x \mid x_n > 0\}$. Then the difference quotient argument, using essentially the same test function v (although altering the bump function appropriately) goes through to give the bounds:

$$\sum_{k,\ell=1,k+\ell<2n}^{n} ||u_{x_k x_\ell}||_{L^2(V)} \leq C(||f||_{L^2(U)} + ||u||_{H^1(U)}) \tag{28.37}$$

Note the bound above excludes $u_{x_n x_n}$, because this would require difference quotients where u is translated out of the half ball. To overcome this one uses that from the interior estimates $Lu = f$ a.e. and the ellipticity condition to control $u_{x_n x_n}$ by the bound:

$$|u_{x_n x_n}| \leq C(\sum_{k,\ell=1,k+\ell<2n}^{n} |u_{x_k x_\ell}| + |Du| + |u| + |f|) \tag{28.38}$$

This with the estimates above implies $u \in H^2(V)$ with good bounds – bounds that don't depend on the choice of $W \subset\subset V$ which might concievably get worse as $\text{dist}(W, \partial U)$ tends to zero. For the general case we then use that ∂U is C^2 so we can straighten it out by a C^2 function – that is realize U locally as the image of the half ball above under a C^2 mapping. This map can be arranged to be C^2 invertible so one can define L', u', f' by composing with its inverse and find that u' satisfies $L'u' = f'$, this PDE on the half ball happens to be elliptic so we can apply the work we already did, and this implies that the original function u satisfies local versions of the bounds we want. These can be patched together using that U is bounded so ∂U is compact giving the claim. □

As before, this argument can be iterated and a statement similar to theorem 28.3 can be shown, this time additionally assuming that ∂U is C^{m+2} and $u \in H^1_0(U)$. Cutting to the chase we have the following corollary as above in combination with the trace theorem, theorem 23.8 and defintion of H^1_0:

Corollary 28.6. *Assume $a^{ij}, b^i, c, f \in C^\infty(U)$, $u \in H_0^1(U)$ is a weak solution of the elliptic problem $Lu = f$ in U, $u = 0$ on ∂U where U is bounded with smooth boundary. Then $u \in C^\infty(\overline{U})$ and solves the elliptic boundary problem in the classical sense.*

29. Maximum principles for general elliptic operators

The regularity theory above says that oftentimes solutions to PDE will be smooth, and next we show for such solutions that the maximum principle holds. By the Fredholm alternative, this will then lead to a general existence theory for a broad class of PDE to the problem $Lu = f$ in U, $u = 0$ along ∂U by showing there is a unique solution to the problem when $f = 0$, so we have a chain of reasoning where regularity theory implies/lets us apply a uniqueness statement, via the maximum principle, which gives us a general existence statement via the Fredholm alternative. As we mentioned there are alternative ways one may try to show uniqueness which sidestep regularity (which was sort of tedious), such as energy methods. There are also maximum principles for merely weak solutions to elliptic PDE, as discussed in chapter 8 of [6], by a test function argument. Its important to know such statements hold if one finds themselves in such a situation but in practice one often has regularity anyway. Supposing here our functions u in this section are classical solutions, there is no harm writing L in nondivergence form:

$$Lu = -\sum_{i,j=1}^{n} a^{ij}(x) u_{x_i x_j} + \sum_i b^i(x) u_{x_i} + c(x) u \tag{29.1}$$

As opposed to L in divergence form, which was better suited for integrating by parts, this form makes it easier to use the derivative tests much as we discussed already for harmonic functions:

Theorem 29.1. *Assume $u \in C^2(U) \cap C(\overline{U})$ on a bounded domain U and the zeroth order coefficient c of L vanishes. Then if $Lu \leq 0$ in U $\max_{\overline{U}} u = \max_{\partial U} u$. Likewise if $Lu \geq 0$ the same is true for its minimum.*

Proof: First suppose that we have $Lu < 0$ strictly in u, and the maximum of u is attained at a point $x_0 \in U$. By the first and second derivative tests we have $Du(x_0) = 0$ and $D^2u(x_0) \leq 0$. Now, for harmonic functions it was already clear that $\Delta u(x_0) \leq 0$ by the second derivative test but in this case all we know about the coefficient matrix $A = (a^{ij}(x_0))$ is that it is symmetric and positive definite so a bit more care is needed. By these conditions, we recall by the spectral theorem that there exists an orthogonal matrix $O = (o_{ij})$ so that

$$OAO^T = \text{ diag}(d_1, \dots d_n),\ OO^T = I \tag{29.2}$$

where $d_k > 0$ are of course the eigenvalues of A. The point of this is that after a coordinate change A really looks more like the Laplacian and we can apply the second derivative test to say $\sum_{i,j=1}^{n} a^{ij}(x_0)u_{x_ix_j}$ has a sign. Now, write $y = x_0 + O(x - x_0)$ so that $x - x_0 = O^T(y - x_0)$ and so by the chain rule:

$$u_{x_i} = \sum_{k=1}^{n} u_{yk}o_{ki},\ u_{x_ix_j} = \sum_{k,\ell=1}^{n} u_{y_ky_\ell}o_{ki}o_{\ell j} \tag{29.3}$$

Thus at the point x_0:

$$\begin{aligned}
\sum_{i,j=1}^{n} a^{ij}(x_0)u_{x_ix_j} &= \sum_{i,j=1}^{n}\sum_{k,\ell=1}^{n} a^{ij}(x_0)u_{y_ky_\ell}o_{ki}o_{\ell j} \\
&= \sum_{k,\ell=1}^{n}\sum_{i,j=1}^{n} a^{ij}(x_0)u_{y_ky_\ell}o_{ki}o_{\ell j} \\
&= \sum_{k,\ell=1}^{n} d_k u_{y_ky_k} \\
&\le 0
\end{aligned} \tag{29.4}$$

where the third equality is using that $\sum_{i,j=1}^{n} a^{ij}(x_0)o_{ki}o_{\ell j}$ is exactly the $k\ell$ entry of OAO^T and the inequality is by the second derivative test applied to each of the $u_{y_ky_k}$ evaluated x_0. Since $c = 0$ this gives that $Lu \geq 0$, contradiction our assumption that $Lu < 0$ and giving $\max_{\overline{U}} u = \max_{\partial U} u$ in this case.

In the general case, that merely $Lu \leq 0$, we consider essentially as in the harmonic case the function $u^\epsilon(x) = u(x) + \epsilon e^{\lambda x_1}$, where λ will be picked below and $\epsilon > 0$. By the uniform ellipticity condition we have that $a^{ii}(x) > \theta$ considering the vectors

$\xi = e_i$ so that

$$\begin{aligned} Lu^\epsilon &= Lu + \epsilon L(e^{\lambda x_1}) \\ &\leq \epsilon L(e^{\lambda x_1}) \\ &= \epsilon e^{\lambda x_1}(-\lambda^2 a^{11} + \lambda b^1) \\ &\leq \epsilon e^{\lambda x_1}(-\lambda^2\theta + ||\mathbf{b}||_{L^\infty}\lambda) \end{aligned} \tag{29.5}$$

Where $\mathbf{b}$ is the vector given by the b^i. This can be arranged to be less than zero for λ sufficiently large implying from the work above $\max_{\overline{U}} u^\epsilon = \max_{\partial U} u^\epsilon$. Letting $\epsilon \to 0$ then gives the claim – note that λ doesn't depend on ϵ. The other statement, for u with $Lu \geq 0$ (these are supersolutions mirroring the terminology for harmonic functions, and functions with $Lu \leq 0$ are subsolutions) follows from the linearity of L and considering $-u$. □

A somewhat weaker weak maximum principle can be shown when $c \geq 0$. Below, we denote $u^+ = \max(u, 0)$ and $u^- = -\min(u, 0)$. Then:

Theorem 29.2. *Assume $u \in C^2(U) \cap C(\overline{U})$ and $c \geq 0$ in U. Then if $Lu \leq 0$ in U $\max_{\overline{U}} u \leq \max_{\partial U} u^+$. Likewise if $Lu \geq 0$ $\min_{\overline{U}} u \geq -\max_{\partial U} u^-$.*

Proof: First suppose that u is a subsolution, that is $Lu \leq 0$, and consider the open (since u is continuous) set $V \subset U$ defined by $V = \{x \in U \mid u(x) > 0\}$. Supposing first that V is nonempty, then on V the operator Ku defined by $Ku = Lu - cu \leq -cu \leq 0$. K has no zeroth order term so the theorem above implies $\max_{\overline{V}} u = \max_{\partial V} u = \max_{\partial V} u^+$ (drawing a quick picture makes this clear). If V is empty then u is nonpositive so the assertion follows trivially. If $Lu > 0$ the argument follows as before by considering $-u$. □

Another way this result can be phrased is to say that if u is a subsolution and it has a nonnegative maximum on $\overline{U}$ then it is obtained on ∂U, and likewise if it is a supersolution and has a nonpositive minimum on $\overline{U}$ it is obtained on ∂U. Thus if u as in the assumptions satisfies $Lu = 0$ the maximum of $|u|$ is obtained along ∂U. Combining this with the Fredholm alternative and regularity theory gives the following much anticipated fact:

Theorem 29.3. *Suppose U is a bounded smooth domain and L is a uniformly elliptic operator on U with smooth coefficients and $c \geq 0$. Then there is a unique smooth solution to $Lu = f$ in U, $u = g$ along ∂U for any smooth functions f, g on $\overline{U}$.*

Of course, the statement isn't the most optimal that one could piece together from the facts above at all but is a nice clean statement. Its easy to construct counterexamples to the above when $c < 0$. There is also a general strong maximum principle, the starting point being the famous Hopf lemma, where $B = B(0, r)$ is a ball:

Lemma 29.4. *Assume $u \in C^2(B) \cap C^1(\overline{B})$ and is a subsolution for an elliptic operator L with $c = 0$. Then if there exists a point $x^0 \in S$ so that $u(x^0) > u(x)$ for all $x \in B$ we have $\frac{\partial u}{\partial \nu}(x^0) > 0$, where ν is the outer unit normal to B at x^0.*

Proof: The proof can be thought of as a barrier argument. Consider the auxillary function $v(x) = e^{-\lambda|x|^2} - e^{-\lambda r^2}$ defined on our ball $B = B(0, r)$ and λ is to be selected below. By a computing and estimating similar to the proof of the weak maximum principle above one can see that $Lv \leq 0$ in the annulus $R = B(0, r) \setminus B(0, r/2)$ if λ is sufficiently large. This gives by the linearity of L that $u + \epsilon v - u(x_0)$ is a subsolution in R for any $\epsilon > 0$.

On the other hand by the assumption that x^0 is a strict maximum for u in B we have there is an $\epsilon > 0$ so that $u(x^0) \geq u(x) + \epsilon v(x)$ for all $x \in S(0, r/2)$, and because v vanishes on $S(0, r)$ by its design we have the same is true for $x \in S(0, r)$. Hence on ∂R we have $u + \epsilon v - u(x_0) \leq 0$ so, since its a subsolution for L, must be nonpositive in all of R. But this function is precisely zero at $x = x^0$ so at that point we must have $\frac{\partial u}{\partial \nu}(x^0) + \epsilon \frac{\partial v}{\partial \nu}(x^0) \geq 0$. We can actually compute the normal derivative of v though, and we find:

$$\frac{\partial u}{\partial \nu}(x^0) \geq -\epsilon \frac{\partial v}{\partial \nu}(x^0) = -\frac{\epsilon}{r} Dv(x^0) \cdot x^0 = 2\lambda \epsilon r e^{-\lambda r^2} > 0 \tag{29.6}$$

where in the middle equality we used that we know the unit normal of $S(0, r)$ at x^0 is x^0/r. $\square$

There isn't quite as general as what's given in [5] but suits our purposes. The main point is that we know $\frac{\partial u}{\partial \nu}(x^0) > 0$ strictly – of course we have the weak inequality. Because $u(x_0) > u(x)$ for all $x \in B$ this gave us the room to squeeze in the function v, which one could say propped up the normal derivative of u. Its a beautiful proof and quickly gives a strong maximum principle:

Theorem 29.5. *Assume $u \in C^2(U) \cap C^1(\overline{U})$ and is a subsolution for an elliptic operator L with $c = 0$, where U is a bounded connected domain. Then if u attains its maximum over $\overline{U}$ in an interior point, then u is constant within U.*

Proof: Supposing this is the case and writing the maximum of u as M, denote by C the set of points in U where the maximum is taken (by assumption, nonempty). If $C \neq U$, then the open set $V = \{x \in U \mid u(x) < M\}$ is nonempty. Chosing a point y

in V closer to C than to ∂U, let B be a ball in V whose boundary touches C. The Hopf lemma says that there $\frac{\partial u}{\partial \nu} > 0$ so in particular $Du \neq 0$, contradicting the first derivative test. $\square$

There's a weaker statement available in the case $c \geq 0$, as one can find in chapter 6 of [5]. There is also a general Harnack inequality – one has to be careful about the assumptions though, because it implies the strong maximum principle. The proof offered in [5] for the case $b^i, c = 0$ goes by considering for a positive solution u the function $v = \log(u)$ and showing eventually that its gradient is bounded, on sets $V \subset\subset U$, which one can then integrate.

References

[1] Ben Andrews, Bennett Chow, Christine Guenther, and Mat Langford. *Extrinsic geometric flows.*

[2] Sigurd Angenent. *Shrinking doughnuts.* Progress in Nonlinear Differential Equations and their Applications, 7, 21-38.

[3] Otis Chodosh. Mean curvature flow (Math 258) lecture notes. Unpublished notes of a class taught by Brian White.

[4] Tobias Colding and William Minicozzi II. *Generic mean curvature flow I; generic singularities.* Annals of Mathematics. (2) 175 (2012), 755-833.

[5] Lawrence Evans. *Partial differential equations, 2nd edition.*

[6] David Gilbarg and Neil Trudinger. *Elliptic partial differential equations of second order.*

[7] Matt Grayson. *The heat equation shrinks embedded plane curves to round points.* J. Differential Geom. 26 (1987), 285–314.

[8] Seppo Granlund and Niko Marola. *On a frequency function approach to the unique continuation principle.* Exp. Math. 30 (2012), 154-167.

[9] Gerhard Huisken. *Flow by mean curvature of convex surfaces into spheres.* J. Differential Geom. 20 (1982), 237–266.

[10] Nikolaos Kapouleas, Stephen Kleene, and Niels Martin Møller. *Mean curvature self-shrinkers of high genus: Non-compact examples.* Journal für die reine und angewandte Mathematik, 2018(739), 1-39.

[11] Jurgen Jost. *Partial differential equations.*

[12] Blaine Lawson Jr. and Marie–Louise Michelsohn. *Spin geometry.*

[13] John Lee. *Introduction to smooth manifolds.*

[14] Hans Lewy. *An Example of a Smooth Linear Partial Differential Equation Without Solution.* Annals of Math Vol. 66, No. 1 (Jul., 1957), pp. 155-158.

[15] Peter Li and and S.T. Yau. *On the parabolic kernel of the Schrodinger operator*. Acta Math. 156, 3-4 (1986), 153–201.

[16] Carlo Mantegazza. *Lecture notes on mean curvature flow.*

[17] Various authors. *Order of the occult hand.* Wikipedia entry.

[18] Sh. A. Alimov. *Parametrix method.* Encylopedia of math.

[19] Steven Rosenberg. *The Laplacian on a Riemannian Manifold.*

[20] Leon Simon. *Schauder estimates by scaling*. Calc. Var. PDE 5 (1997), 391–407.

[21] Michael Taylor. *Partial differential equations: basic theory.*

[22] Terrence Tao. *Compactness and contradiction.*

[23] Alexander Veselov. *Huygen's principle*. https://web.archive.org/web/20160221215126 /http://www.lboro.ac.uk/microsites/maths/research/preprints/papers02/02-49.pdf

[24] Frank Warner. *Foundations of differentiable manifolds and lie groups.*

[25] N. A. Watson. *A theory of subtemperatures in several variables*. Proc. LMS, 26 (1973), pp. 385–417.