Stakgold chap 5 meta meta notes PhL 7.14.09

Text pp 1-87 = 87 p, raw notes 79 p, meta notes 18 p, meta meta notes 11 p, overview 1 p.

5.1 Introduction (1) **2**

5.2 Test Functions in n dimensions (on K_n) (3) **2**

5.3 Distributions in n dimensions (4) **2**

5.4 Convergence of Distributions relative to a parameter (10) **3**

5.5 Additional Properties of Distributions (17) **3**

5.6 Fourier Transforms of Functions and Distributions (23) **3**

5.7 PDE's for Distributions (39) **5**

Theorem 1: 6

Theorem 2: 6

5.8 Fundamental Solutions (48) **6**

Summary of the Five Cases Above: 7

5.9 Classification of PDEs (72) **8**

Overview (1 page, added 3.18.11)

The first order of business is to **extend the theory of distributions** to the case of differential operators of **order p** (instead of order 2) and with **n variables** (instead of 1 variable). The **"multi-index" k** notation handles the fancier terms that can now appear in L. The compact support space of **test functions** is called $\mathbf{K}_n$ and we have a certain rule for null sequences $\varphi_{\mathbf{n}} \to 0$ and for how smooth a test φ needs to be. Stak updates all our **distributional "rules"** for this more general case. He examines the **convergence** of sequences of distributions (the limiting thing need not be integer n→∞, but could be a continuous a→b.) In this section he produces interesting **limits** that represent $\delta^{(\mathbf{n})}(\mathbf{r})$ and I think this was something unfamiliar to Jim Ball when we messed a bit with this in Jan 2011. The **convolution** and **direct product** of distributions are considered.

Anticipating the need to be able to do **Fourier Transforms on distributions**, Stak has a whole section on this idea. It requires a whole new test function φ space called $\mathbf{S}_{\mathbf{n}}$ where now support is no longer compact and the φ all have exponential decay far away. With this completely redefined test function space, we can now do Fourier Transforms on distributions and functions which blow up at worst as a power, and this gives various theorems which get used later. A Fourier Transform of distribution t is called t^ , and the conjugate variables are called x and u.

But before using the Fourier stuff, Stak considers **PDE's in distributions**! Solutions are either **strict** or **generalized**. One shows that u is a generalized solution of Lu=s by showing that <Lu, φ> = <s,φ> for all test functions in $S_{\mathbf{n}}$, where <> means functional, not scalar product. But <Lu,φ> is *defined* as <u, L*φ> so what you really show is <u, L*φ> = <s,φ>. In the case that u is a classical function, we can replace <u, L*φ> = (u, L*φ) where (..) is the scalar product, and this lets us bring in **Green's Theorem** with its surface **current J** (the **parts** term) and then our test in this special case becomes (Lu,φ) – "parts" = <s,φ>? In the case Lu = s = 0, we have to show parts = 0. Stak **demonstrates this idea** with the 1D wave

equation having a possible solution u = H(x-t). He shows that the parts = 0 so that this really is a solution. Since u is not C^2 (it is not even C^1), this is a generalized solution to the equation Lu=0. If u were C^2, it would be a strict solution. The line x = t is an example of a **characteristic**, and along this line u has a constant discontinuity (of 1) just out there in free (x,t) space. Stak shows that for L = Laplace, you cannot have this situation of a discontinuity out in free space. Even a Green's function solving $Lg=\delta$ is continuous everywhere, although it is not C^2 (it is piecewise C^1) so I guess g is a generalized solution. A strict solution of the 1D wave L would be f(x-t) where f is any C^2 function of a single variable (it has to be C^2 because in doing Lu, we need 2 derivatives since p = 2).

The scene suddenly shifts and Stak obtains **spherically symmetric solutions** to a whole set of $Lg = \delta$ equations in varying dimension n and even does one p=4 example. Since there are no BC's, the solutions are called E, not g. Sometimes he gets a solution by requiring $E \to 0$ as $r \to \infty$. These E functions are the "response" to a localized "source". See summary below for the 5 cases he treats. He uses the Fourier Transform work done above as a tool to derive a lot of these **fundamental solutions** E.

The final chapter section is another scene change. **The Cauchy Data** idea is explained (it is specified on some piece of hypersurface σ of dimension n-1), and we learn that this data determines a solution of a PDE as long as there are no "problem points" on the prescribed surface σ. Problem points can be located by plotting the **characteristics** of L and then a problem point occurs whenever a specifying surface σ is tangent to a characteristic. If you select σ that avoids this problem, then the Cauchy data determines your solution in a certain region of space. The characteristics are simply determined from the coefficient functions in L. It turns out that elliptic L have no such surfaces so there are no problem points, but this is not pursued. Hyperbolic L have two sets of characteristics and parabolic L have one set. Reading this, you can see that Stak is touching upon the huge monster subject of PDE general theory (which I have never studied as of 3/11).

Chapter 6 will ignore all this characteristic stuff, but I think it might return in Chapter 7.

5.1 Introduction (1)

Here the technology is provided for us to talk about our Volume 1 topics in R_n instead of R_1. Many concepts are generalized: interval (a,b) is now region R, boundary is now a hypersurface σ of dimension R_{n-1}, new meaning for Local Integrability. The notion of the multi-index k lets us write a general linear order p partial differential operator very concisely.

5.2 Test Functions in n dimensions (on K_n) (3)

The test function φ(x) is now supported on bounded region R in R_n, and has fancier requirements since it has many kinds of derivatives. The notion of a null sequence of test functions and its role are discussed, all as expected. The space of test functions is now K_n instead of K_1.

5.3 Distributions in n dimensions (4)

As in 1D, any LI function f(x) defines a distribution according to $\langle f,\varphi \rangle = \int dx f(x)\varphi(x)$, but now the integral is over R_n instead of R_1. Heaviside $H_R(x)$ is defined on R in obvious manner. All the definitions

and rules of one dimension reappear here in more general form: scaling, translation, product by function, differentiation of distribution, adjoint L*.

5.4 Convergence of a "sequence" of distributions (relative to a parameter) (10)

It is extremely common for a delta function or other singular function to be represented as a **limit** of a function sequence which depends on a *parameter*, perhaps $\delta(x) = \lim_{a\to 5} f_a(x)$ where here a is the parameter. In general we would write $t = \lim_{a\to A} t_a$ where "t" is a distribution, and we are interested in the existence of this sequence limit, hence the word "convergence" in the section heading above. Often the parameter a approaches zero or infinity, sometimes it is an integer as in a partial sum. As an interesting set of examples, one can represent $\delta^{(n)}(\mathbf{x})$ by a limit of a sequence function which depends only on the radial coordinate r. The notion of **spherical coordinates** in R_n is addressed here and there, I should gather up the conclusions. There is in fact a radial coordinate r, an angle $d\Omega_n$, sphere area $S_n(r)$ and volume $V_n(r)$, and there is a way to define all the angles to generalize the θ or R_2 and the θ, φ of R_3. There is only one polar angle.

5.5 Additional Properties of Distributions (17)

This section addresses four random topics. Two of them concern two different kinds of products of distributions. One of these is the **direct product** $t(x_1,x_2) = t_1(x_1)\, t_2(x_2)$

$$< t_1(x_1)\, t_2(x_2), \varphi(x_1, x_2) > \; = < t_1(x_1), <t_2(x_2), \varphi(x_1, x_2) >>$$

where we have two independent R_n spaces with an integration over each space and a test function defined as φ: $R_n x R_n \rightarrow R$. Obviously one could extend this to more than two coordinates. Not much is done with this idea, however. The other product is the **convolution product**

$$<t_1 * t_2, \varphi> \; = < t_2(x), <t_1(z), \varphi(x+z) >>$$

which still involves a double R_n integration, but uses the usual test functions of a single variable. This is the thing that appears in diagonalization discussions, though Stak does not mention that. Some special rules apply if L happens to have constant coefficients, which implies that $L_x = L_{x+b}$.

The two other topics treated here are the **delta function of a function** like $\delta(f(x))$, and how you write the delta function in spherical coordinates, or arbitrary **transformed coordinates**, where the **Jacobian** appears.

5.6 Fourier Transforms of Functions and Distributions (23)

Until the very end of this section, we are working in R_1. Stak chooses three FT **conventions**: sign of exponent, size of constant, and f^(u) for the FT designation. The variable conjugate to x is u. He writes not one but four **Parseval formulas**, and talks about the recovery integral as a horizontal contour at an appropriate height in the ω plane where ω = u+iv. He shows that two different functions of x can have the same functional form for f^(ω) but in different non-overlapping regions of the ω plane, an interesting fact. We can decompose any function as $f(x) = f_+(x) + f_-(x)$ which are called **one-sided functions**. When FT is

applied to $f_+(x)$ functions, it is the same as the Laplace Transform. Basically he reviews the ***classical*** (ie, non distributional facts) about FT in 1D.

The main topic of this section is how we should compute the FT not of a function in R_1, but of a *distribution* in R_1. His answer to this question uses one of the Parseval formulas as our definition, so we write $\langle t\hat{}(u), \varphi(u)\rangle = \langle t(x), \varphi\hat{}(x)\rangle$ as our definition of the FT t^. At the same time, we have to change from test function space K_1 to **space** $\mathbf{S}_1$ of decaying expo type functions over all of R_1 instead of bounded support test functions as in K_1. Changing from K_1 to S_1 causes $\varphi\hat{}(x)$ to be a test function if $\varphi(x)$ is, then the theory is self-consistent and useful. For LI functions, $\langle f(x), \varphi(x)\rangle_{S1}$ converges as long as the function f(x) "diverges" at most as a power, and such f(x) are called **functions of slow growth**. Examples of such slow growth functions are 1, or x^k.

Stak derives a set of rules for S_1 test functions and then shows how the definition of t^(u) above makes these rules apply to FT's of distributions in general. Here are those rules:

A. $[t^{(k)}]\hat{} = (-iu)^k\, t\hat{}$
B. $[(ix)^k t(x)]\hat{}(u) = [t\hat{}(u)]^{(k)}$
C. $(t\hat{})\hat{}(x) = 2\pi\, t(-x)$
D. $[t(x-a)]\hat{}(u) = e^{iua}\, t\hat{}(u)$
E. $[t(-x)]\hat{}(u) = t\hat{}(-u)$

Using all this new technology, we can now compute FT's of singular distributions in a clean manner, and here is a famous example which shows why "δ^" = "1", which says that the identity distribution is the FT of the delta distribution. We would just write δ^ = 1 for short. The following is true for all φ in S_1 :

$$(\delta\hat{}, \varphi) = (\delta, \varphi\hat{}) = \int dx\, \delta(x)\, \varphi\hat{}(x) = \int dx\, \delta(x) \int dy\, \varphi(y)\, e^{ixy} = \int dy\, \varphi(y)\, 1 = (1, \varphi)$$

Going the other way, we have $1\hat{} = 2\pi\delta$ and then $1\hat{}\hat{} = 2\pi 1$ and $\delta\hat{}\hat{} = 2\pi\delta$. The physics or engineering student would be used to the idea that $1\hat{} = 2\pi\delta$ based on the expo (limit) representation of the delta function, but now we have a framework in which we can actually prove that such formulas are meaningful. The test function concept is the key to understanding distribution theory.

This section ends with some discussion of the famous $-iH\hat{}(y) = pf(1/y) - i\pi\,\delta(y) = 1/(u+i\varepsilon)$ idea, where now we are taking the FT of the Heaviside function. This FT it turns out is quite singular.

Finally, Stakgold **generalizes** the FT from R_1 to $\mathbf{R}_n$ and basically nothing dramatic happens, everything goes through in an obvious fashion. The test function space is called S_n, and we then have certain multiindex properties of such test functions. The exponent is now $i\mathbf{x}\bullet\mathbf{u}$ but later he likes $i\mathbf{x}\bullet\boldsymbol{\alpha}$.

Why does Stakgold introduce all this FT business? This is not made clear at this point, but later in the Chapter when we look at solutions of LG = δ equations, we see that the FT is a nice tool for finding solutions for these Green's Functions (fundamental solutions) which he will call E, not G. Most students would be familiar with the idea that the Laplace or Fourier Transform converts a constant-coefficient ODE Lu = f into a simple polynomial equation in the transformed space variable s which gives you u^(s) in terms of f^(s) and various poles and zeros in s, then you can use the inversion formula to get u(x) with a contour integral.

5.7 PDE's for Distributions (39)

In Chapter 1 we studied ODE equations like Lu = 0 and Lu = δ and we talked there about the fact that solutions u might be regular functions, or they might be symbolic functions. If the solution u and all "p" derivatives are regular functions, the solution is called a **classical solution**. If u is a function but some of these derivatives of u are singular, it is a **weak solution**, and if even u itself is singular, it is called a **distributional solution**.

In the PDE world these same terms are not used. Instead, a solution is either **strict** or **generalized**. Strict means that Lu=s is valid at each point in some region R, whereas generalized means that Lu = s is only true in a blurred sense over R which our new technology clarifies: $\langle Lu, \varphi \rangle \equiv \langle u, L^*\varphi \rangle = \langle s,\varphi \rangle$ for any φ in $K_{\mathbf{n}}$. Here, L* is the very well defined formal adjoint of L. In other words, if processing u against L*φ gives you the same result as processing s against φ, then u is a distributional solution. Warning: you must not think of these distributions as integrals, until such time you can show that to be true.

Stakgold does not have any interest at this point in the inhomo equation, but instead worries only about the **homo equation Lu=0**, in which case we must have $\langle Lu, \varphi \rangle \equiv \langle u, L^*\varphi \rangle = 0$ if u is a distributional solution.

At this point we are made aware of the **generalized Green's Theorem in $\mathbf{R}_n$** which says this:

$$\int dV\,(\varphi Lu - uL^*\varphi) = \int d\mathbf{S}.\mathbf{J}$$

or

$$(Lu,\varphi) - (u,L^*\varphi) \;=\; \int d\mathbf{S}.\mathbf{J} \qquad = \text{"parts"}$$

where **J** is the "surface current" associated with differential operator L (I show **J** is not unique). When we are trying (as we do many times) to show that some solution u is in fact a generalized solution to Lu=0, we have to show that

$$\langle Lu, \varphi \rangle \equiv \langle u, L^*\varphi \rangle = \langle s,\varphi \rangle$$

and if u is a classical function (but is not a C^p function), we have then need to show that

$$(u, L^*\varphi) = \langle s,\varphi \rangle$$

which means we have to show that

$$(Lu,\varphi) \;-\; \text{"parts"} \;=\; \langle s,\varphi \rangle$$

In the case that s = 0, to show that Lu=0 as a distribution, we have to show that "parts" = 0.

As one might expect, the surface current **J** depends on the form of L, and Stak considers various famous cases. Usually the $d\mathbf{S}\bullet\mathbf{J}$ integrand involves only the normal vector **n** to the surface S, but in the case of the 1D wave operator, it also involves a related vector **q called the transversal vector**, and this fact has implications for the ability of a PDE to have solutions with discontinuities "floating in space", such as a Heaviside shock wavefront in the space x-t.

Stak presents a few theorems, to wit:

Theorem 1: If u is a strict solution of Lu=0, it is also a generalized solution.
Theorem 2: If u is a generalized solution of Lu=0 *and* is C^p, it is also a strict solution.
Corollary: If u is a generalized solution of Lu = 0 and is NOT C^p, then it is not a strict solution.

At this point, Stak teaches by doing some examples, and there are many trees in this forest. He first considers the 1D wave equation in R_2 which is x-t space. He notes that any f(x±t) is a strict solution (c=1). He then asks whether perhaps H(x±t) is a viable generalized solution. He does a lot of work to show that the surface current term (the "parts") does in fact vanish for such a u. The surface in this case is a certain closed curve C which I have drawn fairly carefully. He notes in passing that in this solution, we are able to "maintain" a constant discontinuity across the line x = t (ie, Heaviside does this) and a line like this we shall later learn is called a characteristic curve. He shows that you cannot maintain a discontinuity over some general curved boundary for the 1D wave equation, only on a boundary like x = 1. The reason is that for a curved boundary, you cannot make those parts vanish unless you assume there is no discontinuity at all across the boundary. So at this very early stage in the chapter, and with much mystery I might add, the reader gets the idea that a solution can only maintain a discontinuity across a boundary in R_2 if that boundary is a characteristic curve for L. Further, the only discontinuity that it can support is Δu = constant. So a characteristic is a place where a solution might have a discontinuity.

Stak's second example is $L = \nabla^2$ whose current **J** does not involve vector q. For this reason, it is not too hard to show that for this L, you cannot have a solution which maintains a discontinuity out in open space. Later we shall learn that this L is "elliptic" which means it has no characteristic curves, and that is the reason you cannot maintain any discontinuities! But that lies ahead, and I can only comment here since I have made two passes already through the chapter.

This section ends with a set of Exercises all of which I have looked at. We take a peek at the wave equation in R_{1+n} space and we identify the characteristics $\mathbf{r}\bullet\hat{\mathbf{k}} \pm ct$ = constant and that is where generalized solutions can have discontinuities. Other examples include $L = \partial_{x1} + \partial_{x2}$ and $L = \nabla^4$. We show that δ(x±t) are also generalized solutions of the 1D wave equation. Then we do $L = \partial_{x1}\partial_{x2}$ and show that $\delta(x_1 - a)$ is a generalized solution.

The Heaviside wave example above would have been called a weak solution in the ODE world because it is a function, but its derivatives are distributions. The exercise showing δ(x-t) is also a generalized solution would be called a distributional solution in the ODE world.

In the meta notes, certain details are clarified that are not covered in this meta meta summary.

5.8 Fundamental Solutions (48)

In this section, we compute the "Green's functions" for LG = δ, similar to what we did in Volume 1, but here we are in R_n space. No comment whatsoever is given about "boundary conditions" -- they are completely ignored here. I think we should assume that either there are none, or if there are some, they are spherically symmetric. For this reason, we consider only spherically symmetric solutions for G. In my raw notes, I give a little proof of why solutions must be function of r only if the problem is rotationally invariant. Since boundary conditions are not included, the term "Green's Function" is not technically correct, so the solutions here are called "fundamental solutions" and of course they will have some undetermined constants as in the 1D case. He uses the letter E instead of G, but in cases where we limit things to t≥0, he uses C for Causal. I don't know what the E stands for!

There are many details here. Stak treats 5 different famous L operators each in turn, but here I will try to list off the tools that he uses and things I learned. For each case, he likes to write down the specific solutions for n = 1,2,3 as well as the general solutions.

- Integrating LE=δ over a small sphere puts a condition on the solution which is akin to the jump condition on g' in the 1D world.

- All 5 examples have constant coefficients so that $E_\xi(x) = E_0(x-\xi)$ so we only deal with E_0 and call it E.

- In some examples, Stak actually proves that the result is a formal distributional solution.

- Sometimes he uses a spatial FT and sometimes a temporal FT in solving for the functions E.

- In the causal cases for evolution equations, he shows an equivalent PDE system with BC's at t = 0. The spatial world is still considered rotationally invariant.

- His notations are: $f\,\hat{}$ = FT, $f\,\tilde{}$ = LT.

- It was in the case of the damped wave equation that we run into the Jim integral I could not do.

Summary of the Five Cases Above:

Recall from our previous work that the wave equation can have "distributional solutions" whereas Laplace and Heat and Helmholtz do not. That is borne out in the results summarized here. Where it is relevant, Stak seeks solutions that don't blow up as $r\to\infty$, or which are "causal", being 0 for $t < 0$.

1. Laplacian

$-\nabla^2 E_n = \delta(\mathbf{x})$

$E_n(r) = C_n / r^{n-2} + A$ n > 2 ODE method

$E_3(r) = 1/4\pi r + A$

$E_2(r) = E = -\ln(r)/4\pi + K$

$E_1(r) = E = -r/2 + K$

2. Helmholtz

$(-\nabla^2 - \lambda)E_n = \delta(\mathbf{x})$ $\lambda = -\mu^2$

$E_n(r; \lambda=-\mu^2) = (1/2\pi)\,(\mu/2\pi r)^{(n/2-1)}\, K_{(n/2-1)}(\mu r)$ n = 2,3....

$E_3 = e^{-\mu r}/4\pi r$

$E_2 = (1/2\pi)\, K_0(\mu r)$

$E_1 = e^{-\mu r}/2\mu$

recursion relation: $\partial_r E_n = -2\pi r\, E_{n+2}$

general solutions shown for n odd and even top of page 58

ODE method Bessel, then spatial FT method $\mathbf{x} \leftrightarrow \alpha$

3. Heat Equation (58)

$(\partial_t - a\,\nabla^2)C_n = \delta(t)\delta(\mathbf{x})$

$C_n(r,t) = (4\pi a t)^{-n/2} \exp(-r^2/4at)$ // using spatial FT $\mathbf{x} \leftrightarrow \alpha$

4. Wave equation undamped (61) $(\partial_t^2 - c^2 \nabla^2)C_n = \delta(t)\delta(\mathbf{x})$ n = # spatial dimensions

cribbed solutions shown 5.148 (no odd) and 5.149 (n even)

$C_3(r,t) = \delta(t-r)/4\pi r$ = pulse

$C_2(r,t) = (1/2\pi)H(t-r)/\sqrt{t^2 - r^2}$

$C_1(r,t) = H(t-r)/2$

used temporal FT (Laplace) method $t \leftrightarrow s$

5. Wave equation damped (65) $(\partial_t^2 + 2\gamma\partial_t - c^2 \nabla^2)u_n = \delta(t)\delta(\mathbf{x})$ $u_n = e^{-\gamma t}C_n$

$(\partial_t^2 - \gamma^2 - c^2 \nabla^2)C_n = \delta(t)\delta(\mathbf{x})$ " telegraphy"

$C_n(r,t)$ not given, but could use cribbed results

$C_3(r,t) = \delta(t-r)/4\pi r + (1/4\pi) H(t-r) I_0'(\sqrt{t^2 - r^2})/\sqrt{t^2 - r^2}$ = pulse + wake

$C_2(r,t)$ = we are supposed to find this in Exercise 5.29

$C_1(r,t) = (1/2) H(t-r) I_0(\gamma\sqrt{t^2 - r^2})$

5.9 Classification of PDEs (72)

The title here leaves out a lot of what this section is about.

We come now (for the first time in this Chapter) to the subject of "boundary conditions" in the world of PDE's in R_n. The general idea seems to be this: if you specify some information about the solution u to Lu = s on a piece of hypersurface σ of dimension n-1 in R_n, then that may constrain the solution u in R_n outside this piece of hypersurface in some region R of R_n. In our 1D experience with ODE's and intervals (a,b), a "hypersurface" means a point in the interval, and we there supplied certain information about the solution u at certain points (information could include derivatives).

Many questions drift into view. I imagine Laplace's equation with potential specified on some non-closed piece of surface, perhaps some capacitor plates, and I know this determines the potential everywhere, so that is a good example of R_3 with a hypersurface of 2 dimensions. What would happen if you specified your information on a hypersurface of dimension n-2? For R_3 we know this does not specify much, but maybe for R_4 and above it does? This subject is not addressed. [But I run into it much later in doc "iris green's function attempt..." where I ponder the Green's Function for a wire.]

In general, the "boundary condition" information is called "**initial data**" I think, but we know what is meant. It is some "supplied data" about the solution that we hope determines a unique solution. We are always concerned that we might supply too much data so no solution exists, or not enough data so that many solutions exist. Stak never uses the term "boundary conditions" in this section.

The opening gambit idea for supplying a reasonable amount of "initial data" for a PDE problem is the **Cauchy Problem** idea. Here you supply u and just normal derivatives of u up to order p-1 and you supply this on some piece of hypersurface σ. In most examples he gives, this surface is not a closed surface. This data is called **The Cauchy Data**. Just as we "have a problem" in ODE's when $a_0(x) = 0$, we have a similar and related problem in the PDE world. In the ODE world, if $a_0(x_0) = 0$ at some point x_0 (a point on the hypersurface in this case), we cannot solve the ODE to learn $\partial_x^p u(x_0)$ from the Cauchy data because we cannot divide through by $a_0(x_0)$. This derivative is the one piece that is missing from our knowledge of having all derivatives through order p at a point on σ. Once we have all these derivatives, it seems reasonable that we can sort of continue the solution away from the surface σ out into the volume of R_n.

The main idea here is to imagine a local coordinate system on a "**tangent plane**" to σ at point P where λ is then the normal coordinate. In that coordinate system, we can write $L = b(P)\partial^p_\lambda + L'$ where b is some complicated function of the a_k coefficients in the normal coordinates. If b(P) = 0, we have our problem and we cannot compute ∂^p_λ and fill out our full information near P. We also have a problem because then Lu is completely determined by the Cauchy (and transverse) data which conflicts with the need to have Lu = f.

So we arrive at our big question within this Cauchy Problem approach. How can we tell whether or not we can compute all the derivatives through order p? Stak only attempts to answer this question in R_2 where we have coordinates x,y (sometimes called x,t). And he only attempts an answer in the cases p = 1 and p = 2, first and second order PDEs. In each of these two cases, we find that we can in fact compute all the derivatives as long as, at each point on the specifying hypersurface (now a curve C in R_2), we *avoid* having the slope of C be a certain function of the a_k coefficient functions at that point. So the idea then is to first make some kind of "slope map" of the R_2 space (an arrow at each point, like a wind forecast map), and then restrict yourself to using a specifying hypersurface C which is nowhere tangent to a slope in that slope map. Such a map can in fact be represented by a set of continuous curves since "things are smooth", and these curves are called **characteristic curves**, so the Cauchy initial data game requires picking a surface C which is nowhere tangent to a characteristic curve.

For p = 1 and $L = a\partial_x + b\partial_y$ + rest, the slope map is given by dy/dx = b/a. If a and b are constant coefficients, which is often the case, the slope is a constant so you have a set of parallel lines for your characteristic curves.

For p = 2 and $L = a\partial_x^2 + 2b\,\partial_x\partial_y + c\partial_y^2$ + lower terms, the slope map is $dy/dx = (b \pm \sqrt{b^2-ac}\,)/a$. Assuming as we have been that a,b,c are real functions, this brings in a new feature: the structure of the slope map depends on the sign of b^2 -ac. If this thing is negative, there is no slope map! There are no slopes that you have to avoid with your specifying curve C. This is the **elliptic** case, and the canonical example is the Laplacian in 2D for L. On the other hand, if b^2 - ac > 0, then at each point on C there are not one but *two* slopes you have to avoid being tangent to, so there are two sets of characteristic lines that form some sort of mesh. This is the **hyperbolic** case and examples are ∂^2_{xy} and the 1D wave equation. The last case occurs if b^2-ac = 0 in which case there is just one slope to avoid at each point. This is the **parabolic** case, and the heat/diffusion equation is an example.

So this is then our introduction to the "classification of PDEs". We only really are doing this for R_2 and for p = 1 and p = 2. In the p = 2 case, we get the elliptic, hyperbolic and parabolic cases.

Stak closes out the chapter with two hyperbolic examples (p 851). For each example he considers boundary condition specification methods which are NOT of the Cauchy problem type as well as which are of the Cauchy problem type. His first example is $L = \partial^2_{xy}$ and Lu = 0, while his second example is that 1D wave equation $L = \partial_t^2 - c^2\partial_x^2$ and Lu = 0 which, in fact, is related by a coordinate transformation to the first example. Here he derives something called **d'Alembert's Formula** for the solution when your boundary condition data is supplied only at t =0, to which I provide a light cone interpretation. He then shows how one can systematically solve a complicated 1D wave problem (string with ends tied down) using the **method of characteristics**.

Stak claims but does not prove that a solution to Lu=s can only have discontinuities that run along characteristic curves. We saw this earlier in the chapter in the 1D wave equation case, but Stak has not really made a solid connection IMHO. If your specified data on C has a discontinuity, that discontinuity

propagates away from the curve along the intersecting characteristic curve. In the hyperbolic case, I guess this means it propagates off in 2 directions at once. Again, this is just a claim, not proven or even discussed much.

And so ends this volume's discussion of what we might call "the general theory of PDE's". I think that from now on, he is going to focus on specific equations like Laplace and Wave and Diffusion. These are the famous classical equations and there is plenty to be learned about just these equations. Somehow we are going to go beyond R_2 I am sure. I don't think he will spend any time talking about cases where the coefficients are functions, not constants. In R_1 this is the world of Special Functions.

I think I can see how Stakgold had accumulated a set of lectures over the early energetic years he taught and then attempted to organize them into chapters of a book. In many ways, this subject defies organization, and it is sometimes a little hazy what goes where.

Some questions I have at this point:

(1) does the p = 2 classification extend beyond R_2 ? If so , how so?

(2) is the general theory different somehow for something like Minkowski space SO(3,1) ?

(3) is there a lot more "general PDE theory" that he did not mention, or which has been developed in the last half century since he wrote these volumes? Surely there are many books on PDE's and they must address many topics he skips over. Let's take a contents peek right now. OK, lots of books, but Stakgold's topics still predominate!

The web provides this possible answer to question (1):

If there are n independent variables $x_1, x_2, \ldots, x_n$, a general linear partial differential equation of second order has the form

$$Lu = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{i,j} \frac{\partial^2 u}{\partial x_i \partial x_j} \quad \text{plus lower order terms} = 0.$$

The classification depends upon the signature of the eigenvalues of the coefficient matrix.

1. Elliptic: The eigenvalues are all positive or all negative.
2. Parabolic : The eigenvalues are all positive or all negative, save one which is zero.
3. Hyperbolic: There is only one negative eigenvalue and all the rest are positive, or there is only one positive eigenvalue and all the rest are negative.
4. Ultrahyperbolic: There is more than one positive eigenvalue and more than one negative eigenvalue, and there are no zero eigenvalues. There is only limited theory for ultrahyperbolic equations (Courant and Hilbert, 1962).

For n = 2 we can write $L = a\partial_x^2 + b\partial_x\partial_y + b\,\partial_y\partial_x + c\,\partial_y^2$ and the matrix in question is $\begin{pmatrix} a & b \\ b & c \end{pmatrix}$. The eigenvalues of this matrix $Mv = \lambda v$ occur at $\det(M-\lambda I) = 0$ which is $(a-\lambda)(c-\lambda)=b^2 = 0$ etc. It is certainly not obvious to me why for elliptic the eigenvalues for n=2 have the same sign, but I guess I could show it.

Note added 3.19.11. I noticed today that the entire 800 page Volume 2 of Courant & Hilbert is on this very subject, partial differential equations. I have various other downloaded PDE books and there are a lot more where they came from. It is a big subject.