Stakgold Chapter 5 Meta Notes PhL 7.4.09

Raw notes for this chapter were 79 pages, these meta notes are 18 pages, 4:1 reduction.

5.1 Introduction (1) **2**

5.2 Test Functions (3) **2**

5.3 Distributions (4) **2**

5.4 Convergence of Distributions relative to a parameter (10) **2**

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

A. The direct product of two distributions. 3

B. The convolution product of two distributions. 3

C. Delta function of a function $\delta[f(x)]$ 3

D. Change of coordinate system (21) 3

5.6 Fourier Transforms (23) **3**

A. Classical Results 3

B. Examples (25) 3

C. One Sided Functions (28) 4

D. Functions of Slow Growth. (29) 4

E. Fourier Transforms of Distributions in 1 dimension. (30) 4

F. Fourier Transforms of Distributions in n dimensions. (36) 6

5.7 PDE's for Distributions (39) **7**

The Homogeneous Equation Lu=0 (42) 8

Theorem 1: 8

Theorem 2: 8

Example 1: 8

Example 2: 10

Exercises (47) 11

5.8 Fundamental Solutions (48) **11**

Case (1). Laplacian (49) 12

Case (2): Helmholtz equation (53) 13

Case (3): The Heat Equation (58) 13

Case (4): The Wave Equation (61). -- undamped 14

Case (5): The Wave Equation (61). -- damped 14

Summary of the Five Cases Above: 15

Exercises (69) 16

5.9 Classification of PDEs (72) **16**

First Order PDE's in R_2 17

Second Order PDE's in R_2 18

Hyperbolic Example 1: $L = \partial_{xy}$ and $Lu = 0$ 19

Hyperbolic Example 2: $L = \partial_t^2 - c^2\partial_x^2$ and $Lu = 0$ 19

5.1 Introduction (1)

We are now in n dimensions, $x = \{x_i\}$, $||x||$, dx and σ as a surface with dS.

In n dim, a **LI function** on bounded region R in R_n is one that has a finite integral over any ball within R, or alternatively, over any surface patch σ in R on an **n-1 dimensional surface**.

The closure of the set of values where a function is not zero is the **support** in R_n.

A general n-dimensional derivative D^k is described fully by a **multiindex** $k = \{k_i\}$. The **order** of a derivative is Σk_i. An L in n dimensions of order p has the form $L = \Sigma_k a_k D^k$ which is a sum over multiindices which have order ≤ p where $a_k(x)$ in general are C^∞ on R_n.

5.2 Test Functions (3)

Stricter requirement now is that, for φ in K_n, all $D^k\varphi$ exist for any order. And φ_m is a **null sequence** if not only $\lim \varphi_m = 0$ but also $\lim D^k\varphi_m = 0$ for any k. Of course φ has finite domain of support in R_n.

5.3 Distributions (4)

In R_n, every LI function is associated with a (**regular**) distribution $\langle f,\varphi\rangle = \int dx f(x)\varphi(x)$, but non-LI functions also define distributions (**singular**) such as f = δ. Generalization of **Heaviside** $H_R(x)$ which has value 1 on some R in R_n and is 0 elsewhere. **Delta** in n dimensions, **dipole** idea, **translation** of delta, **scaling** rule, **simple layer**. **Product** of a distribution and a function. **Differentiation** of a distribution. The **adjoint** L* of L is given from $\langle Lu,\varphi\rangle = \langle u,L^*\varphi\rangle$. For general $L = \Sigma_k a_k D^k$ get simple result for L*. Same notion of "formally self adjoint" L = L* such as for ∇^2.

A singular distribution generally does not have a "**value**" at all points in R_n but you can usually find regions where the value is clearly 0.

5.4 Convergence of Distributions relative to a parameter (10)

Distribution which has a **parameter** α converges to some t (that is, $t_\alpha \to t$) if $(t_\alpha , \varphi) \to (t, \varphi)$ for ANY test function φ. Lots of examples given. Notion that $dV_n = r^{n-1}dr\, d\Omega_n$ and $S_n(1) = \int d\Omega_n$. It is possible to **replace multiple deltas** such as $\delta^3(\mathbf{x})$ with a *single* delta such as $\lim_{\alpha\to 0} [\ \pi^{-3/2} \exp(-r^2/\alpha^2) / \alpha^3\ ]$, where dimensions are cm^{-3} on both sides. A whole general family of limits using radial g(r) functions exists to make $\delta^n(x)$. Key point: all these forms for delta functions are *limits* of expressions which depend on a parameter involved in that limit. So these expressions are "delta sequences" so to speak and they converge to the resulting symbolic function δ(x).

5.5 Additional Properties of Distributions (17)

A. The direct product of two distributions.

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

B. The convolution product of two distributions.

$<f_1 * f_2, \varphi> = < f_2(x), <f_1(z), \varphi(x+z) >>$ if f_1 vanishes outside some finite interval

Example 1: $\delta * t = t$
Example 2: $(L\delta) * t = Lt$
Example 3: $(Ls)*t = s*(Lt) = L(s*t)$ where L is an ODE operator with *constant coefficients*

Lemma: $L_x = L_{x+b}$ if coefficients are constants. Proof:

C. Delta function of a function δ[f(x)]

D. Change of coordinate system (21)

$\delta^3(\mathbf{x}) = 1/[r^2 \sin\theta] * \delta(r-r')\, \delta(\theta-\theta')\, \delta(\varphi-\varphi')$ // the usual example, Jacobians

Exercises deal with sphere of charge, ring of charge, etc.

5.6 Fourier Transforms (23)

A. Classical Results

Notation is that the FT of f(x) is f^(u), sign and constant conventions noted. The projection goes one way, coming back is the recovery, inversion or expansion. On page 24 we have 4 Parseval formulas, one of which is this, where I include the derivation

$$\int dx\, \hat{f}(x) g(x) = \int dx\, g(x) \{ \int dz\, e^{ixz} f(z) \} = \int dz\, f(z) \{ \int dx\, g(x)\, e^{ixz} \} = \int dz\, f(z)\, \hat{g}(z)$$

which says you can move the ^ from one to the other. The classical recovery formula is along the real axis in the u space as in 5.34. Stak writes u + iv = ω to allow a horizontal contour shifted vertically by an amount v. If you then consider $F(x) = e^{-vx} f(x)$ as candidate x-space function, the recovery formula can be written as in p 25B which looks like the recovery just for f(x), but contour is shifted by v. So we have the notion that the recovery integral is really a contour integration in the ω plane and he does various things with this idea.

B. Examples (25)

He considers f(x) = $e^{-|x|}$ then f(x) = -2sinh(x)H(x) as two examples. They have the same FT, but recovery formulas have a different vertical location.

C. One Sided Functions (28)

Notation here is $f_+(x)$ for a function that is 0 on the left, etc. and you can write $f(x) = f_+(x) + f_-(x)$ and get an interesting recovery 5.48 where the horizontal contour can be adjusted for each piece.

D. Functions of Slow Growth. (29)

f(x) here grows at most as a power and can therefore be "tamed" by e^{-vx} and made to converge if you are talking about a one-sided function! So one approach is to take arbitrary $f(x) = f_+(x) + f_-(x)$ and separately "tame" the two pieces in this manner. We are really leading up to the next section in this approach.

E. Fourier Transforms of Distributions in 1 dimension. (30)

We know how to associate a distribution "f" with a function f(x), namely, $\langle f,\varphi \rangle$. We now want to find a distribution "f^" associated with f^(u), the FT of f(x). Our obvious candidate is this:

$$\langle \hat{f}(u), \varphi(u) \rangle = \langle \left(\int dx\, f(x)\, e^{iux} \right), \varphi(u) \rangle = \int dx\, f(x) \langle e^{iux}, \varphi(u) \rangle = \int dx\, f(x) \int du\, e^{iux} \varphi(u)$$

$$= \int dx\, f(x)\, \hat{\varphi}(x) = \langle f(x), \hat{\varphi}(x) \rangle$$

The problem is that if φ(u) is a K_1 test function, φ^(x) might NOT be a test function, so the last item above is not a legal distribution thing, since the second argument must be a test function. When we were dealing with *derivatives* and $\langle f'(x), \varphi(x) \rangle = -\langle f(x), \varphi'(x) \rangle$, we didn't have this problem since φ'(x) is a test function if φ(x) is a test function. I show a specific example where φ^(x) is not a test function.

The fix here is to come up with a **new test function space called S_1 to replace K_1**. In this new space, if φ(u) is S_1, so is φ^(x), and so the above becomes $\langle \hat{f}(u), \varphi(u) \rangle_{S1} = \langle f(x), \hat{\varphi}(x) \rangle_{S1}$ and we then have a well-defined meaning for "f^".

What is this new space S_1 ? Instead of having bounded support as in K_1, a test function in this new space extends to infinity, but must decay faster than any power, which means φ basically has some kind of expo decay. Stak shows on page 29 in a theorem that φ in S_1 => φ^ in S_1. Moreover, the integral implied by $\langle f(x), \hat{\varphi}(x) \rangle$ converges for f(x) having at worst a power behavior for large x, and such an f(x) is called a **function of slow growth**.

So at this point we now have a distribution "f^" which is associated with any f(x) "of slow growth", and it is given by

$$\text{" } \hat{f} \text{ "} = \langle \hat{f}, \varphi \rangle_{S1} = \langle f, \hat{\varphi} \rangle_{S1} = \int dx\, f(x)\, \hat{\varphi}(x) \qquad \text{the new animal}$$

$$\text{"} f \text{"} = \langle f, \varphi \rangle_{K1} = \langle f, \varphi \rangle_{K1} = \int dx\, f(x)\, \varphi(x)$$

where for comparison I have shown the distribution f associated with f(x) from our earlier work. We had no concerns there about the large argument behavior of f(x) since φ(x) in K_1 had bounded support.

Stak talks a bit about **the S_1 space**, and the fact that any $D^k\varphi$ is a test function, and the notion of a **null sequence** in this space. Then he comes up with four powerful **"properties"** of these S_1 test functions:

(5.53)	$[\varphi^{(k)}]\hat{}(u) = (-iu)^k \varphi\hat{}(u)$	diff, then FT
(5.54)	$\partial_u{}^k [\varphi\hat{}(u)] = [(ix)^k\varphi(x)]\hat{}(u) = [\varphi\hat{}(u)]^{(k)}$	FT, then diff
(5.55)	$(\varphi\hat{})\hat{}(x) = 2\pi\,\varphi(-x)$	FT twice in a row
(5.56)	$[\varphi(x-a)]\hat{}(u) = e^{iua}\,\varphi\hat{}(u)$	translation

Here super k means the kth derivative (or D^k when we later extend to S_n). Stak then shows that, using these properties of S_1 test functions, you can obtain a similar set of rules for an arbitrary distribution "t" as follows:

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. $[t(-x)]\hat{}(u) = t\hat{}(-u)$

Here "t" is a distribution (symbolic function) and " t^" is the FT of that distribution. Note that the FT of a distribution is itself a distribution, just based on this simple relation,

$$(t\hat{}, \varphi)_{S1} = (t, \varphi\hat{})_{S1}$$

since φ^ is a test function if φ is. Here are some **examples** of FT's of distributions:

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

Here we conclude that $\delta\hat{} = 1$, so the FT of δ is "1". Here is another example

$$(1\hat{}, \varphi) = (1, \varphi\hat{}) = \int dx\, \varphi\hat{}(x) = \left[\int dx\, \varphi\hat{}(x) e^{ixy}\right]_{y=0} = \varphi\hat{}\hat{}(0) = 2\pi\varphi(0) = (2\pi\delta, \varphi)$$

from which we conclude that the FT of "1" is 2πδ. In this case, the thing we start with "1" is a classical function, whereas its FT is a symbolic function. We then find that $1\hat{}\hat{} = 2\pi 1$ and $\delta\hat{}\hat{} = 2\pi\delta$. Earlier in the chapter we arrived at these same conclusions without the formal S_1 framework we now have. Another example is that

$$[\delta(x-a)]\hat{}(u) = e^{iua}\, \delta\hat{}(u) = e^{iua}\, 1.$$

A final example is more complex and he derives it several ways, this being the result

$H\hat{}(y) = i\,pf(1/y) + \pi\,\delta(y) \quad \Rightarrow \quad -iH\hat{}(y) = pf(1/y) - i\pi\,\delta(y) \quad = 1/(u+i\varepsilon)$

The FT of the Heaviside function is the symbolic function shown, having two pieces. This result certainly has some mystery to it.

F. Fourier Transforms of Distributions in n dimensions. (36)

This section gives a generalization of the above single variable S_1 test function space method to an arbitrary number of variables. Here are the key changes:

• The definition of a test function with n variables involves the space S_n instead of S_1. The condition here involves a multiindex "power" (p 36 A) and a separate multiindex "derivative" (p 36 B, this we saw earlier). The vanishing condition for test functions $\varphi(x_i)$ is then shown in C, where k and $\boldsymbol{\ell}$ are now arbitrary finite multiindices.

• The FT now takes the obvious form shown page 37A where we can put i**x•u** in the exponent. The inversion formula is now 37B. The inversion formula now has $(2\pi)^n$ instead of 2π, just multiplying them all together. This is all very familiar to me for n=3 of course.

• The rule now for doing the FT of a distribution is shown in C, where the implied "inner product" is of course the multidimensional integral. All our results above can be redone. He quotes some of them as shown in D and E.

$1\hat{}(x) = (2\pi)^n\delta^n(x) \qquad [\delta^n(x)]\hat{}(u) = 1$

At this point we have a bunch of good exercises, and I will just quote some of the results learned.

5.16 $t_n\hat{} \rightarrow t\hat{}$ if $t_n \rightarrow t$

5.19 The Laplace Transform $f_+\hat{}(is) = \int_0^\infty dx\ f_+(x)e^{-sx} = \int_0^\infty dx\ f(x)e^{-sx} = f\tilde{}(s)$

5.20
A.	$[pf(1/x)\hat{}](y) = +i\pi\,sgn(y)$	// derived by me and web checked
B.	$sgn\hat{}(x) = 2i\,pf(1/x)$	"
C.	$[\log\|x\|]\hat{}(u) = -\pi\,sgn(u)/u$	// derived by me

5.21 $[x^k]\hat{}(u) = (-i)^k\,2\pi\,\delta^{(k)}(u)$

At this point, I insert a very long digression on "Spherical Coordinates in n dimensions". This is needed for the next exercise:

5.22 Here t involves a "simple layer a(x)" on surface σ and we are supposed to compute t^ . The result is shown in 5.68 (p 39). We then specialize the result to a = 1 and σ being a sphere in n dimensions with result 5.69, and then we compute this thing for n = 2 and n = 3 with results shown in 5.70 and 5.71.

5.7 PDE's for Distributions (39)

This section opens with the famous Green's Theorem written as $\int dV\ (vLu - uL^*v) = \int d\mathbf{S}.\mathbf{J}$. and the differential form $(vLu - uL^*v) = \nabla\bullet\mathbf{J}$. We then get at once four examples for L and in each case the current **J** is "quoted". I now understand how to obtain the current (see parts and greens.doc) and have verified each case.

In the fourth example, which uses L = the wave operator, we encounter an interesting vector q which is related to the normal vector n at the boundary of a volume in R_{n+1}. If we write $n = (\mathbf{n}, n_t)$ as our n+1 dimensional "normal vector" at the surface (so that $\mathbf{n}.d\mathbf{s} = 0$ at any point on the surface where ds is any distance along the surface at the point), then this vector $q = (-\mathbf{n}, n_t)$. It is called **the transversal vector** and one can write **J**.d**S** for example 4 (shown in 5.81) in the very simple form shown in 5.83.

We then consider the nature of solutions of Lu = s. Solutions u are classified either as being strict, or as being generalized. A solution is **strict** in R if it is valid (ie, if Lu = s) at every point within some open region R and if u is C^p (recall that p is the order of L). A solution u is a **generalized** solution if $\langle Lu,\varphi\rangle = \langle s,\varphi\rangle$ for all test functions φ in K_n on some region R. We know this is somehow going to involve an integral over our space and so represents a softened notion of a solution u -- allowing for a distributional solution. [I like the word distributional more than the word generalized, but OK.]

Note added 3.19.11. On page 42 Stak says that we have a generalized solution to Lu=s if $\langle Lu,\varphi\rangle=\langle s,\varphi\rangle$ for all test functions φ, where <> is our usual functional distribution concept (not a scalar product). He then says that, when u is a generalized solution, the *definition* of $\langle Lu,\varphi\rangle$ is $\langle u,L^*\varphi\rangle$. The thing on the right, $\langle u,L^*\varphi\rangle$, is the distribution "u" against a test function $\varphi' = L^*\varphi$ which has a well-defined meaning in our distribution theory framework. So $\langle Lu,\varphi\rangle \equiv \langle u,L^*\varphi\rangle$ is the definition of functional "Lu" in the case that Lu is ill-defined at certain points in R because Lu is a distribution and not a function. Just so, the distribution "δ" is ill defined at the point x = 0 within its region R, but $\langle\delta,\varphi\rangle \equiv \varphi(0)$.

Suppose we have Lu = s where both sides are distributions. To show that u was a distributional solution, we would have to show that $\langle u,L^*\varphi\rangle = \langle s,\varphi\rangle$ for all φ. You have to ask: since what we really want to show is $\langle Lu,\varphi\rangle = \langle s,\varphi\rangle$, why can't we just say Lu = s and be done with it? $T_{Lu} = T_s$. This is the Big Confusion in Stak's presentation I think. This makes us ask: what is the meaning of equation Lu=s ? I think this is the answer: Lu = s is an equality between two classical functions when you stay away from singular points, just as Lg = δ says Lg=0 away from singular points and both sides are functions. But if you include the singular points in your analysis, Lu=s is undefined! It is not defined everywhere on your region R of interest. It is not an equality of classical functions. We therefore cannot just say that $T_{Lu}[\varphi] = T_s[\varphi]$ based on the fact that Lu=s which is undefined. We have to SHOW that $T_{Lu}[\varphi] = T_s[\varphi]$ for some function or distribution u. The way we show this is that we have to show that $\langle u,L^*\varphi\rangle = \langle s,\varphi\rangle$ which you could write as $T_u[\varphi'] = T_s[\varphi]$.

Theorem: strict => distributional. Assume u = function. Want to show $\langle u,L^*\varphi\rangle = \langle s,\varphi\rangle$. But now Lu=s is fully defined as an equality of classical functions, so $\langle u,L^*\varphi\rangle = (u,L^*\varphi) = (Lu,\varphi)$ + parts = (s,φ) + parts. Then we have $(Lu,\varphi) = (s,\varphi)$ QED. Alternatively, we want to show $\langle Lu,\varphi\rangle = \langle s,\varphi\rangle$, but since everything is a classical function, all we have to show is $(Lu,\varphi) = (s,\varphi)$ which is true since Lu = s.

Example: Suppose s = 0 and L = L* (formally adjoint). Then to show u is a distributional solution to Lu=0, we need to show that $\langle u,L\varphi\rangle = 0$. We cannot just say $\langle Lu,\varphi\rangle = \langle 0,\varphi\rangle = 0$ as discussed in the paragraphs above. This is what is coming next.

The Homogeneous Equation Lu=0 (42)

Theorem 1: If u is a strict solution of Lu=0, it is also a generalized solution.

Proof: Since a K_n test function φ is C^∞, L*φ is well defined and so is $\int dV\ u\ L^*\varphi$ and we can interpret this to be the meaning of <u,L*φ> on some region R. Because u is C^p, we know we can do full "parts integration" on this integral using Green's Theorem. We know that the "parts" will be the integral over the boundary of R of the surface current **J** which is bilinear in u and φ (and well defined). But R is also the support region for test functions φ and we know that $D^k\varphi = 0$ on the boundary, so therefore the surface current vanishes everywhere on the boundary. Thus, we have so far shown that

$$<Lu,\varphi> \equiv <u,L^*\varphi> = (u,L^*\varphi) = (Lu,L\varphi) + \text{parts} = 0 + 0 = 0 = <0,\varphi>$$

According to 5.85, since we have thus shown that <u,L*φ> = <0,φ>, we have shown that the strict solution u is also a distributional solution.

Theorem 2: If u is a generalized solution of Lu=0 *and* is C^p, it is also a strict solution.

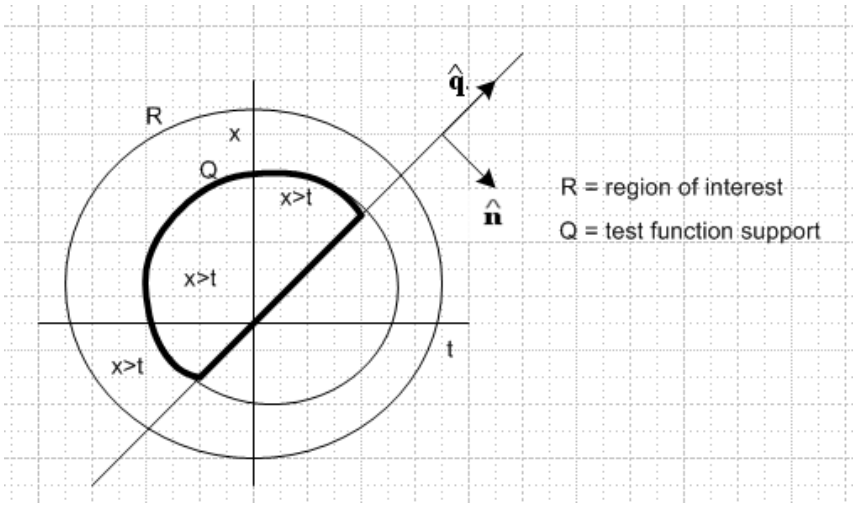
Corollary: If u is a generalized solution of Lu = 0 and is NOT C^p, then it is not a strict solution.

We are now going to look into strict and generalized solutions of Lu=0 with two examples.

Example 1: Wave Equation in 1 spatial dimension, so only t and x are involved.

We first find that any wave f(x-t) is a solution, velocity = 1. Of course f(x+t) would also be a solution. If we use any reasonable (say LI) function for f, then we have a strict solution for Lu=0.

Are there generalized solutions of this Lu= 0? A nice example to consider is setting f(x-t) = H(x-t). This candidate for u is obviously not C^p. Stak goes through some tricky steps that I outline in detail in my raw notes in which he shows that, in fact, the "parts" vanishes and we get the exact same conclusion as shown in the line above which ends with = 0. But in this example, we are not starting with something we know is a strict solution. Our space is R_{1+1} and the H function causes the volume integral to be over just the region x > t of the test function support. I had to draw this picture just to understand what is going on:



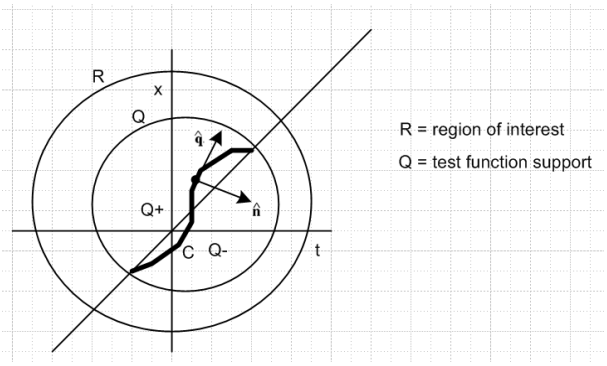
The "volume" of our Green's theorem integral is the "partial disk" with dark perimeter (it is the full disk, but H = 0 on the lower right part). As shown in 5.83, there are two surface integral terms we need to make go away, and they involve that **q** transversal vector. The main boundary piece of concern is the straight line segment where $\varphi \neq 0$. But **q** points along this line and u (=H) is the constant "1" along this line (inside volume), being the H function, so $\partial u/\partial q = 0$ on the line segment, so that term gives nothing. The other term is line integral of $\partial\varphi/\partial q$, but again since **q** is along the line, this is $\varphi(b)-\varphi(a)$ and each of these vanishes since on the support perimeter of φ. Thus, both boundary terms go away and we conclude that

$$\langle Lu,\varphi\rangle \equiv \langle u,L^*\varphi\rangle = (u,L^*\varphi) = \int dV\, u\, L^*\varphi = \int dV\, Lu\ \varphi + \text{parts} = 0 + 0 = 0$$

We are allowed to replace <> with () as shown since u is a classical function in our example (namely H). After we apply Green's Theorem, we can apply Lu = 0 as a classical function inside the volume integral since we are away from singular points there (after all, u = 1 = constant inside). So in this example it is crucial to show that the "parts" vanish, and the parts is the contour integral along the heavy boundary shown in our picture above.

Once again, in this way we are showing that u = H(x-t) is a "generalized" solution of Lu = 0. In passing, note that Lu=0 has a nullspace and f(x-t) is in this nullspace.

Having done this example, Stak asks about the nature of other possible "generalized solutions" we might discover for our 1D wave equation. We imagine that we have some generalized solution u(x,t) which is in some sense singular on a boundary which divides our volume into two pieces. Again, I had to draw a picture:



where the curve C is our imagined singular boundary, like the one we had in the Heaviside case. The question is this: what kind of singularity boundaries are possible? In this more general case where we have non-zero action on both sides of the boundary, the "parts" term becomes this:

$$\text{parts} = \int_C d\ell\, (\Delta u\, \partial\varphi/\partial q - \varphi\, \Delta(\partial u/\partial q))$$

where Δ means the discontinuity across the boundary. We must have parts = 0 for u to be a generalized solution! If both discontinuities vanish everywhere on the boundary, we HAVE no discontinuity boundary at all, contradicting our starting assumption. But if $\mathbf{n}.\mathbf{q} = 0$, we can have a situation where the second discontinuity term above vanishes and the first does not. The second vanishes because we get the same $\varphi(b)-\varphi(a) = 0 - 0 = 0$ idea as in the Heaviside example. By fiddling the first term, Stak shows that the only possible Δu you can have is Δu = constant.

He has thus shown the following: the only possible generalized (non-strict) solutions of our 1D wave equation are those which have a boundary which (1) is flat ($\mathbf{n}.\mathbf{q}=0$) , and (2) has a constant discontinuity. The first requirement means the boundary must have the form $x = \pm t$ + constant, and these two forms are called the **characteristics** of the problem. This conclusion appears underlined in red on page 46. Here we are anticipating the general theory of solving PDE's which always involves these characteristics.

Example 2: Laplace in R_n. Here we think again about a possible boundary of discontinuity out in the middle of our region, but in this case the current surface terms don't involve that "**q**" vector, just the **n** vector. Because of this fact, if you repeat the previous analysis, both of the surface terms must separately vanish (both involving the normal vector **n** as in p 47 A), and in that case, you find that you *cannot have* any discontinuity out there in the middle! There are no "characteristics" for the Laplace problem! The upshot is that for Laplace, every generalized solution is a strict solution and is therefore C^2. You cannot have a Heaviside type solution to Laplace Lu=0, for example, but you can for the wave equation. The big difference is in the form of the surface current J for these two operators! I think we shall find that this fact is true for all Elliptic differential operators L.

Example Summary: So now compare the wave equation with x and t to the Laplace equation with x and y. Both equations are in a space we can call R_2. The only difference in the two L's is the relative *sign* between the two second derivatives. As we have seen in our worked-out exercises above, this small sign change makes a big difference in the form of the "parts" part of Green's theorem. For the minus sign case, the parts term involves the vector q = negate-spatial-part-of(n), For the plus sign case, we just have the normal vector n. In the first case (wave) where the parts involves q, a small class of discontinuity contours

are possible out in the middle of some region, called characteristics. For the Laplace case, there are no characteristics! This is a pretty amazing fact from a simple sign change.

Exercises (47)

One would be unwise to skip these exercises!

Exercise 5.23. This is a pretty important one. We generalize the previous wave Example 1 to the case of R_{1+n} and here are our conclusions: (we include a velocity parameter c now)

- a general strict solution would have the form $u(\mathbf{r},t) = f(\mathbf{r}\bullet\hat{\mathbf{k}}-ct)$
- the characteristics are $\mathbf{r}\bullet\hat{\mathbf{k}} \pm ct$ = constant where $\hat{\mathbf{k}}$ is an arbitrary spatial unit vector
- the only generalized solutions (non-strict) will have a constant discontinuity across this characteristic.

Comment: Well, this has all been very formal, but I think there is a physical interpretation. The wave equation of course has waves as solutions, and you can probably make any wave you want to superposing Heaviside planar wave fronts, or delta function ones. It is natural to have a moving "wall" of discontinuity in a wave situation, a shock wave so to speak. The Laplace equation does not involve time and a solution is just a static thing, and we don't really imagine it could support some discontinuous wall sitting quietly in space. A Laplace solution is smooth everywhere.

Exercise 5.24. Here we have the unnamed spatial operator $L = \partial_{x1} + \partial_{x2}$. We learn what the surface current is, and we find that there is a characteristic which is $x_1 = x_2$ + constant, and we learn that the Heaviside $H(x_1 - x_2)$ is a generalized solution to this problem! So although this looks more like Laplace that Wave, it is more like Wave in our conclusions. Is there some physical problem with Lu=0 ? The strict solutions are of the form $f(x_1-x_2)$ for any f. If you plot z = f(x-y), you always get a plane! The reason is that $\nabla z = f'(\hat{\mathbf{x}} - \hat{\mathbf{y}})$. The 2D gradient always points "up hill" but that direction never varies here, so the surface must be flat.

Exercise 5.25. The operator $L = \nabla^4$ has the same nature as ∇^2 in terms of generalized solutions.

Exercise 5.26. Here we show that $\delta(x\pm t)$ are also generalized solutions of the 1D wave equation. I expended a lot of raw note space verifying this. The proof is quite different from the $H(x\pm t)$ case and we don't use Green's theorem at all, do work directly with a volume integral, etc, see notes.

Exercise 5.27. Now we have $L = \partial_{x1}\partial_{x2}$ and we show that $\delta(x_1 - a)$ is a generalized solution.

Exercise 5.28. Here we show that the heat conduction operator $L = \partial_t - \nabla^2$ (in 1D) behaves like the Laplace operator in that there are no non-strict generalized solutions. I did not do this exercise, but we do want to know the conclusion. Physically this might say that the diffusion process does not have shock wave solutions!

5.8 Fundamental Solutions (48)

In the previous section we studied Lu = 0, the homo equation. We found both strict and generalized solutions such as f(x-t), H(x-t) and $\delta(x-t)$ for the 1D wave equation, and we looked at a few other L

examples. Here we look at $Lu = \delta(\mathbf{x})$ where $\delta(\mathbf{x})$ is a spike at the origin of R_n (or R_{n+1} with time), but we use E instead of u, so $LE_0 = \delta(\mathbf{x})$ { or $LE_\xi = \delta(\mathbf{x}-\xi)$ } .

Most L's of interest to us in this book have constant coefficients, and in that case $L_x = L_{x+b}$ and $E_\xi(x) = E_0(x-\xi)$, so we need only deal with the E_0 problem driven by $\delta(\mathbf{x})$. E_0 is of course called "the fundamental solution".

At the end of my raw notes, realizing that $\delta(\mathbf{x})$ is really of the form f(r), I understood why solutions E must be functions only of the radial variable r in R_n. You do not get $Y_{\ell m}(\theta,\varphi)$ with $\ell \neq 0$ in these solutions (assuming rotationally invariant boundary conditions, which are not mentioned yet).

We proceed now to actually solve the "Green's Equation" $LE_0 = \delta(\mathbf{x})$ for 5 different L's. Not one word is said in this chapter about "boundary conditions". I presume that if we have some non-spherically symmetric BC's, the solutions given here will no longer apply. For example, the potential of a point charge next to a grounded plane we know is non-central about that point.

Case (1). Laplacian (49)

$L = -\nabla^2$, $\nabla^2 E = -\delta(\mathbf{x})$ $E = E(r)$ Everything is done on this single line! (valid for n = 1,2,3....)

$$-1 = \int dV\, \nabla\bullet(\nabla E) = \int d\mathbf{S} \bullet \nabla E = \int (\partial E/\partial r)\, dS = \int (\partial E/\partial r)\, r^{n-1} d\Omega_n = S_n(1)\, (\partial E/\partial r)\, r^{n-1}$$

Luckily we get to use the divergence theorem right away, and we assume a spherical surface. The above tells us *at once* that the solution is $E_n(r) = C_n / r^{n-2} + A$ for n > 2 and we usually set A = 0. The C_n are specific constants for each n (not arbitrary constants like A), and in fact this result is valid for n = 1 as well where we get $E_1(r) = C_1\, r + A$. The size of the first constant is completely determined by $-1 = S_n(1)\, C_n\, (2-n)$, so our single line above gives us both form and scale of the solution. For n = 3 we get $E = 1/4\pi r$, an incredibly famous result (Jackson E&M). For n = 2 the solution is $E = -\ln(r)/4\pi + K$, and for n = 1 the solution is $E = -r/2 = -|x|/2$. In the above, we have to know that $dS = r^{n-1} d\Omega_n$ in R_n and that $\int d\Omega_n = S_n(1)$, something we have dealt with plenty of times earlier in this chapter.

For n = 3 we get this same result by considering $\nabla_r^2 E = 1/r^2 \partial_r (r^2 \partial_r E) = 0$ away from r=0. For the n = 2 case we have $\nabla_r^2 E = 1/r \partial_r (r \partial_r E) = 0$ (cylin) and we repeat its same result. In 5.111 Stak derives the form of ∇^2_r in R_n and then uses that to replicate the n = n solution. We don't have to do things this way (at least for Laplacian), because the above way provides all solutions

Recall our claim earlier that solutions of the Laplace equation Lu = 0 are always "strict" and thus are real genuine functions. We see that here for all values of n when we are away from r = 0 where Lu = 0.

Recall also our earlier claim that $\langle Lu,\varphi\rangle = \langle u,L^*\varphi\rangle = \langle s,\varphi\rangle$ for any generalized solution. Here we have $s = \delta$, so we are claiming then that $\langle E,L^*\varphi\rangle = \langle\delta,\varphi\rangle = \varphi(0)$ for all test functions. Stak explicitly shows this to be true for the n = 3 case on pages 51-52. He of course uses Green's Theorem.

When we have $LE = \delta$ and integrate over volume we get $\int dV\, LE = 1$. In the n = 1 case with L being of second order with coefficients a_i, this led to $1 \approx \int_a^b dx\, [\, a_2(x)\, D^2 E(x)\,] \approx a_2(0)\, E'(x)|^{0+}_{0-}$ which we referred to as the jump condition on the Green's function then called g. So here we see exactly how this jump condition generalizes when we go from R_1 to R_n.

In my raw notes, I talk about how one interprets the divergence theorem for n = 1.

$$\int d\mathbf{S}\bullet\mathbf{A} \;= \int dS\ \mathbf{n}\bullet\mathbf{A} = \Sigma_{\mathbf{p}}\ \mathbf{n(p)}\bullet\mathbf{A(p)} \;= \hat{\mathbf{x}}\bullet\mathbf{A(b)} + \;(-\hat{\mathbf{x}})\bullet\mathbf{A(b)} \;= A(b) - A(a)$$

See " Proof of the Divergence Theorem.doc" for more details.

Case (2): Helmholtz equation (53)

$L = -\nabla^2 - \lambda = -\nabla^2 + \mu^2$. This is really the wave equation where we have replaced $-\partial_t^2$ with μ^2 ($-\omega^2$) which is the λ term. But here, we are just "doing the math".

What happens if we try to repeat our "one liner" from above, which was:

$$-1 = \int dV\ \nabla\bullet(\nabla E) \;= \int d\mathbf{S}\bullet\nabla E \;= \int (\partial E/\partial r)\ dS \;= \int (\partial E/\partial r)\ r^{n-1} d\Omega_n \;= S_n(1)\ (\partial E/\partial r)\ r^{n-1}$$

We pick up an extra term λV and get $-1 = S_n(1)\ (\partial E/\partial r)\ r^{n-1} + \lambda V_n(r)$ where the last is the volume of the sphere in R_n. I think you could solve the problem in this way, but Stak uses ∇^2_r to get ODE 5.114. He changes this to a Bessel equation and requires that the solution not blow up at $r = \infty$ and thus selects a Hankel $H^{(1)}$ as the solution, 5.116. He then uses the generalized jump condition with a small r approximation for the Hankel to get the normalizing constant, and the final answer is (5.118) for $E_n(r; \lambda)$. He then rewrites this in terms of the modified Bessel function K in (5.119). The Bessel order is (n/2) - 1, arg is μr . He then writes these out explicitly for n = 2,3,1. The n = 2 is a K_0, the n = 3 is Yukawa potential where μ is the pion mass, and n = 1 is just the Yukawa exponent.

Now Stak *resolves* this same problem using the spatial FT method. He applies the FT directly to our PDE as in 5.123. Instead of maintaining λ, he sets $\lambda = -1$ and then later restores it. The $-\nabla^2$ becomes what he calls α^2 (I might call it k^2) and we then find the trivial transformed equation $(\alpha^2+1)E_n\hat{} = 1$ so we know $E_n\hat{}(\alpha)$ and all we have to do is the inversion integral to get $E_n(r; \lambda=-1)$. [The simplicity here reminds us of doing Laplace and Fourier transforms of ODE's] He restores the $\lambda = -\mu^2$ and ends up with $E_n(r; \lambda=-\mu^2)$ written not as a Bessel function, but as the integral representation 5.129. This is a nice form because it "exposes" the dependence on r in a simple manner, $e^{-\mu r z}$. From this we quickly obtain a step-by-2 recursion relation for the E_n as in 5.130 and we can then express our even and odd solutions as at the top of page 58.

Obviously the Bessel solution is more direct, BUT these even and odd solutions will be "cribbed" later on when we encounter $-\nabla^2 E_n + \mu^2 E_n = \delta(x)$ in R_n with something new replacing μ^2.

Case (3): The Heat Equation (58)

The Green's equation here is 5.135: $(\partial_t - a\ \nabla^2)C_n = \delta(t)\delta(\mathbf{x})$ where a is a constant. He uses C_n instead of E_n as the solution function, where C stands for "Causal". We are going to insist that the solution be 0 for t < 0, hence "causal". Thus, the solution is $C(\mathbf{x},t) = H(t)\ u(\mathbf{x},t)$. He first shows (as a Theorem) that u then solves 5.136: $(\partial_t - a\ \nabla^2)u_n = 0$ $t>0$ with $u(\mathbf{x},0+) = \delta(\mathbf{x})$. But having done this, he goes back to 5.135 and solves this Green's equation again by the spatial FT method. The FT solution is trivial, just $C_n\hat{}(\boldsymbol{\alpha}\ ;\ t) =$

exp(-a α^2 t). We then do the FT inversion to get $C_n(r,t) = (4\pi at)^{-n/2}$ exp($-r^2/4at$). All done! In the limit t $\rightarrow$ 0 , this thing becomes $\delta(\mathbf{x})$, an example of our radial forms for $\delta(\mathbf{x})$ as limits.

Reminder: he keeps using Fourier stuff because our PDE's have constant coefficients and Fourier gives diagonalization where solution is just algebra.

Case (4): The Wave Equation (61). -- undamped

This is identical to the heat equation but we replace ∂_t with ∂_t^2 and replace a with c^2. The alternate form of the Green's equation this time is p 61 A which has *two* t=0 conditions since we are now second order in time. We still call the solution C_n for "causal" and write C = H(t)u, just as before. Now we do a temporal FT (Laplace really) which leaves the ∇^2 but converts the ∂_t^2 to s^2 giving us 5.146 which is exactly our cribbing equation mentioned above with μ = s, so we read off the solutions for $\widetilde{C}_n(\mathbf{x}; s)$ as in 5.147, where ~ means Laplace transform. Converting back from s to t is trivial, so then $C_n(\mathbf{x}; t)$ is shown in 5.148,9 for odd,even n. As noted, some of these solutions are generalized solutions, and you get a "wake" only for even n. The cases n=1,2,3 are written out. Stakgold then gives some discussion again to his favorite topic, showing that these solutions really area "generalized solutions", though the meaning of this is now just a little more complicated. He shows details for n = 1 and n =3 only. Then in the end he restores speed c, having set it to 1 for most of our work. I like to think of the initial conditions as $\psi(\mathbf{x}) = 0$ and $\partial_t\psi = \delta(\mathbf{x})$ so that our "wave" starts with zero displacement everywhere but with a delta velocity impulse.

Case (5): The Wave Equation (61). -- damped

We add a γ damping term to get (5.157) for u. We change to u = w $e^{-\gamma t}$ and find 5.158 where the inhomo source is altered to be $e^{\gamma t}q$ which is fine. If we want to solve 5.518, we work with the corresponding Green's equation which is 5.159. Notice that the change gets rid of the linear time derivative. Now if we do the time Laplace just as in the previous example, we get the same result but with s^2-γ^2 replacing s^2, so in theory we have $\widetilde{C}_n(\mathbf{x}; s)$ from the cribbed equations. However, here he just quotes the cribbed equation for n = 1 and writes it in 5.161. A new (but small) technical difficulty now appears because $\mu = \sqrt{s^2 - \gamma^2}$ has a branch cut between $-\gamma$ and γ in the s plane, and we have to consider this when we do the time inverse Laplace contour integration shown in 5.165. [he has c = 1 and γ = 1 for the time being] . If we close to the right for t < r, we get our 0 solution as causality demands -- velocity = 1 and signal has not yet reached distance r (r = |x| in 1D if you like). If we close to the left for t>r, we pick up the discontinuity of the integrand across the branch cut. This leads to the solution $C_1(r,t)$ as a certain integral

$$C_1(r,t) = (1/2\pi) \int_{-1}^{1} ds\, e^{st} \cos(r\sqrt{1 - s^2})/\sqrt{1 - s^2} = (1/\pi) \int_0^{\pi/2} d\theta\, ch(t\cos\theta) \cos(r \sin\theta)$$

I had to get help from Jim finding this in GR, see raw notes for this and all details. The integral of interest is this one

$$C_1 = (1/\pi) \int_0^{\pi/2} d\theta\, ch[t\cos\theta] \cos[r\sin\theta] = (1/2)\, I_0(\sqrt{t^2 - r^2})$$

We can then put this into our cribbed step-by-2 recursion relation to get C_3, and when we reinstall γ the result is shown page 68 D, and then we are supposed finally to multiply by $e^{-\gamma t}$ as in p 68 C. So looking at p 68 D, we see that the result for $C_3(r,t;\gamma)$ replicates our undamped result for $\gamma = 0$ which was given in 5.152, but the damping creates the second term which is a wake behind the δ front. He then quotes the n = 1 result which replicates 5.150 when $\gamma = 0$.

Summary of the Five Cases Above:

Recall from our previous work that the wave equation can have "generalized solutions" whereas Laplace and Heat and Helmholtz do not. That is borne out in the results summarized here.

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

$E_n(r) = C_n / r^{n-2} + A$ $\quad n > 2$ $\quad$ 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 $(-\nabla^2 - \lambda)E_n = \delta(\mathbf{x})$ $\quad \lambda = -\mu^2$

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Exercises (69)

Did not do any of these, see raw notes comments.

5.9 Classification of PDEs (72) // for example, "hyperbolic"

Here is the concept of "the Cauchy Problem" in a nutshell. Imagine you have Lu=f in R_n space. Imagine that you have some finite piece of hypersurface σ within this space on which the value of u is specified, and on which also all normal derivatives (λ = local normal coordinate) up to order p-1 are specified. The information as to what the surface σ is and what the data is on this surface is called **The Cauchy Data**. The 1D analogy is "initial" conditions u(a) and u'(a) for 2nd order ODE at some point a.

Now consider a particular point P on the surface σ. Because of the Cauchy data, we know all derivatives of u at P except we don't know $\partial_\lambda^p u$ which is the pth normal derivative. [We also don't know any derivatives that contain this or higher λ derivatives.] We don't have any trouble computing transverse partials because those are determined by u on σ in the neighborhood of P.

Next, we come to the issue of what I call **"P on σ being a problem point for L"** and which Stak calls " L is an interior Diff Op at P wrt σ" -- I prefer my phrasing. When you write Lu in "local coordinates" where λ is the normal coordinate, you find that $Lu = b(P)\partial_\lambda^p u$ + lower terms. If it happens that b(P) = 0, then you basically have Lu(P) = lower terms, but these lower terms are completely determined by the Cauchy data at point P, since P lies on σ. On the other hand, Lu(P) = f(P). In general, these are going to conflict with each other, which really says you are not free to specify the Cauchy data as you like. If this situation arises at any point on your surface σ, you are likely to NOT have a unique well defined solution of your problem Lu=f with this particular Cauchy data.

When b(P)=0, if you try to compute $\partial_\lambda^p u(P)$ from Lu=f (just as you would try to do in a computer iterated ODE solution starting at point P = a in 1D) , the conclusion is that $\partial_\lambda^p u(P)$ is either indeterminate or infinite (does not exist). If you *can* compute $\partial_\lambda^p u(P)$ at a point P, then you know that P is not a problem point. But if you can compute $\partial_\lambda^p u(P)$, then you can compute all derivatives through order p in the original coordinate system. Therefore, a test for P being a problem point is to see whether or not you can compute all order-p derivatives. We shall do this explicitly below. [Good.]

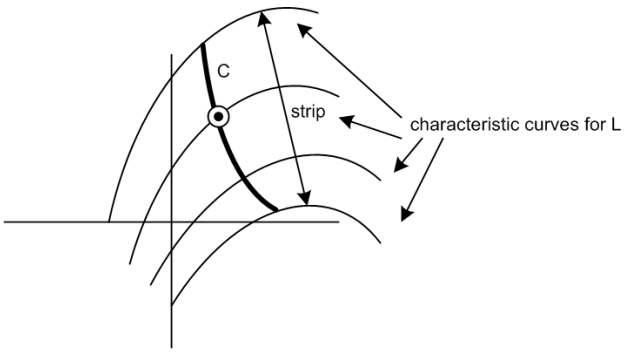
Examples: Section 5.9 opens with an example where $Lu = a_0\partial_x u + a_1 u = f$ in the space R_1. Here p = 1 and the surface σ has dimension 0 which means it is just some point x_0 on the x axis. The full Cauchy data here is $u(x_0)$. Stak shows that you can solve the problem uniquely by a simple iterative (computer) algorithm, as long as $a_0(x_0) \neq 0$. He does not say this, but I think you must have $a(x) \neq 0$ everywhere in your interval of interest (a,b), or the iteration process will be quenched with you get to a problem point. This does not quite fit in with the general theory, which is why he does not mention it. Remember that our subject here is not whether you can form the solution to the problem, but is whether you have a problem point on your hypersurface σ. If you don't have such a problem point, that does not mean the solution exists. If you do have a problem point, you definitely have a problem and very likely a solution does not exist. In this example, we might say that if $a(x_0) \neq 0$, we can create the computer solution at least in some neighborhood of P.

Stak goes on to consider an ODE of order p in R_1 and thinks about computer iteration of a column vector of values u, ∂u, $\partial^2 u$.....$\partial^{p-1} u$. As long as that leading $a_0(x_0) \neq 0$, you will be able to run this algorithm, and you will be able to compute $\partial^p u$ at x_0 from Lu=f and then x_0 will not be a problem point.

First Order PDE's in R_2

I skip now to this section before defining new terms. Let $L = a\partial_x + b\partial_y$ + rest. We have p = 1 and we are now working in R_2. We then ask the question: can we compute $\partial_x u(P)$ and $\partial_y u(P)$? If we can, then we know that P is not a problem point (P is somewhere on σ which here is a curve C in the x-y plane). It is very easy to write two linear equations for these two unknowns, which equations include dx and dy separating an adjacent point Q on σ from P, and you conclude that P is not a problem point as long as a certain determinant does not vanish. When you write out det = 0, it comes out being dy/dx = b/a. As just noted, dy/dx is the slope of C at a point P on C.

This is extremely instructive, so let's have lots more words. In general, a and b are functions of (x,y) and we are saying that a point P is a problem point if the slope of the curve C happens to be b(P)/a(P). How might you check your curve C to see if it has any problem points? Well, one way is to compute the quantity **dy/dx = b/a** at every point in the x-y plane for starters. How might we display "a slope" at every point in the plane? We could do one of those Maple arrow plots. But since everything is presumed continuous, we could instead plot a dense set of curves more or less filling the x-y plane, which curves have a slope which is equal to dy/dx at each point on the curve. You might then have a picture like this, where I show a σ surface C which has no problem points,



The curves just mentioned have a name -- they are called the **characteristic curves** of the problem, and for p = 1 we see they are completely determined by the functions a and b in L. If you first draw up these curves and then draw your candidate surface C, you can see by inspection whether there are any problem points P. At a problem point P, the curve C will be tangent to one of the characteristic curves! Very simple. In this case of R_2 and p = 1, Stak quotes a theorem without proving it. He says that if C has no problem points (as in the above picture), your problem has a solution in the curved strip I have marked between the bounding characteristic curves. Furthermore, you are allowed to have a discontinuity in your Cauchy data (value of u along C), and you will then find that your general solution has a discontinuity all along the characteristic curve which passes through the point on C where you specified a discontinuity. I have marked this with a circled • in the above picture. You could have lots of points like this.

In terms of our past discussion, the fact that you can have such discontinuities out in the middle of nowhere means that you can have "generalized solutions" of Lu=f. We saw examples of this where we had a Heaviside wall along a characteristic curve for the 1D wave equation. For p = 1 in R_2, there are always characteristic curves, so you can always have generalized solutions!!

So in R_2 with p = 1, the theory is pretty solid. If you look specifically at $L = \partial_x$, then a=1 and b = 0 and you find that dy/dx = b/a = 0 so the characteristics are horizontal lines. The general solution to Lu=0 here is u = F(y) so u is then constant everywhere on horizontal lines. Thus, you know u from your Cauchy

data (data on C) at all points left and right of your curve C. Above and below C you know nothing. C will be OK as long as it does not have a point of zero slope in this example.

Another example is L = $\partial_x + \partial_y$ which has dy/dx = 1/1 = 1. The characteristics are then a dense set of straight lines of slope 1 , so make sure C does not have slope 1 anywhere.

Who first did this analysis do you suppose, and when?

http://books.google.com/books?id=T2z225xBWF8C&pg=PA8&lpg=PA8&dq=gaspard+monge+characteristic+curves&source=bl&ots=5wP_AxrxMI&sig=vTHxIWoC-EWvW6GxR4Acv4lOtr8&hl=en&ei=R01bSvG5IYWEsgOXvOWDCw&sa=X&oi=book_result&ct=result&resnum=10

1.2 Characteristics and Classification

Characteristics were first used by Lagrange in two major papers in 1772 and 1779, in which first-order linear and nonlinear equations were considered. This work appears in Chapter 3. Gaspard Monge (1746–1818) introduced characteristic curves in his work "Feuilles d'analyse appliquie à la géométrie" completed in 1770, but not published until 1795, by which time he had considered second-order equations of the type to be discussed next.

So I think Monge and circa 1795 are the answers to my questions, including the p = 1 case below.

Second Order PDE's in R_2

Now we have p = 2, so we want to know if we can compute $\partial_x^2 u(P)$, $\partial_y^2 u(P)$ and $\partial_x \partial_y u(P)$ at some point P on our curve C. As in the p = 1 case, we come up here with a set of *three* linear equations involving these three unknowns, and involving as before dx and dy separating two points on C. We find that we can compute these three p=2 derivatives as long as a certain det ≠ 0. So P will be a problem point in this family of problems (R_2 and p=2) when the curve C is tangent to a characteristic curve which has the following slope: $\mathbf{dy/dx = (b \pm \sqrt{b^2-ac})/a}$ where $L = a\partial_x^2 + 2b\,\partial_x\partial_y + c\partial_y^2$ + lower terms, and notice that Stak puts 2b instead of the traditional B. Let A = a, B = 2b and C = c, then the problem slope would be written as $dy/dx = (B \pm \sqrt{B^2-4AC})/2A$ which is the traditional form. There are three big cases of interest: $B^2 - 4AC > 0$, =0, and < 0. By way of analogy, if you have $Ax^2 + Bxy + Cy^2$ + lower = 0, the curve is either a hyperbola, ellipse, or parabola based on the sign of $B^2 - 4AC$, so those three names are traditionally applied to our R_2 p=2 PDE situation.

Now let's look at the three cases.

First, if $B^2-4AC < 0$, the radical arg is imaginary and there are no characteristic curves in the real x-y plane!! This is the **elliptic** case such as the Laplacian which has B = 1 and A=C = 1 so that $B^2-4AC = -1 < 0$. Thus, there are never any problem points P on surface C. The fact that there are no characteristics is connected with the fact that a solution cannot maintain a discontinuity out in the middle of nowhere, as we showed in previous discussions for the Laplace case. Because there are no characteristics, the Cauchy Problem approach does not work too well with elliptic PDE's in R_2 (at least I think that is why).

Second, if $B^2-4AC = 0$, we then have dy/dx = B/A which is very similar to the p=1 situation. We would draw the single family of characteristic curves, and then make sure that our C surface is not tangent to any of these curves. This is the **parabolic** case, an example of which is heat flow in 1 spatial

dimension, where we then have t = y and B = 0 and C = 0 hence B^2-4AC = 0. [Easy to remember that this is parabolic using the stupid idea that $y = x^2$ sort of matches $\partial_t = \partial_{x^2}$.]

Third, if B^2-4AC > 0. In this case, we get two families of characteristic curves from the equation dy/dx = $(B \pm \sqrt{B^2\text{-}4AC})/2A$ (from the ± sign), and we then have to make our curve C steer clear of both. This is the **hyperbolic** case and there are two simple examples: first, $L = \partial_{xy}$ which has B = 1 and A=C =0 (this turns out to give dy/dx = 0,∞ as the two families of curves). Another is the wave equation in 1D which has B = 0, A = 1 and C = -1. In this **hyperbolic** case, the Cauchy problem approach works well.

Comment: Our discussion of the Cauchy Problem starts out being in R_n with hypersurface σ and with L of some order p. We have general notions of what it means for P to be a problem point on σ. But we don't have any "general formulas" for characteristic curves except in the case p =1 and p = 2. Also, for these cases p = 1 and p =2 , we don't have any formulas beyond R_2 , ie, everything is for n = 2.

Hyperbolic Example 1: $L = \partial^2_{xy}$ and Lu = 0

Several different explicit solution methods are found for this problem.

In (1), it is treated as a Cauchy Problem as discussed earlier, and data is given on a certain curve C page 82 figure. The solution is given in p 82 C, but you have to compute ∂u/∂x from the Cauchy data ∂u/∂n and u on C. In the special case the C is a 45 degree line, you can use 5.186 which involves ∂u/∂n directly. This formula says that you can know u at any point P in the shaded region of page 82 given the value at Q and R (both on C) and given ∂u/∂n which you have to integrate along your straight line.

In (2) we are just reminded that C cannot be tangent to a characteristic. These are V and H lines for this problem.

In (3) we are given only u data on two pieces of perp characteristics (OB and OA). The exact solution out in the middle is then given by 5.187 where u at P is obtained from u at R,Q and O.

In (4) we are given only u data on one piece of characteristic, and on a noncharacteristic curve C. The solution out in the open is given by the same formula as (3), but location O is different.

The last two methods are NOT Cauchy Problem formulations, but are nevertheless situations that arise and the methods work.

Hyperbolic Example 2: $L = \partial_t^2 - c^2\partial_x^2$ and Lu = 0

A simple transform to coordinates η and ξ converts this problem to $\partial^2_{\eta\xi}$ so we can in theory use results from Example 1. In the x-t space, the characteristic curves form a trellis of diagonal lines shown. Two methods of solution are offered here.

In (1) he derives **d'Alembert's Formula** which gives the wave equation solution u(x,t) in terms of only t=0 data. You need both initial position and initial velocity data (all along your string). The formula gives u at point P in terms of u at two t=0 points Q and R, and in addition you need the integral of the velocity data from Q to R. This formula is very much like 5.186 and I think one could argue exactly why they are similar. Probably you can get 5.189 directly from 5.186. I copied a wiki derivation of d'Alembert which basically does this cribbing operation. The formula builds in causality in effect. To compute u(x,t) at P, the formula only involves data that is in the past light cone of point P! This is very clear from the p 85 figure. The string is infinite I think in this example.

In (2) we get a nice example of "**the method of characteristics**" applied to a finite string problem where we formally tie down both ends, making u = 0 at both ends. The s-t plane of interest is now a vertical strip that is partitioned into diamond and triangle shapes by the characteristics. The idea is that you use d'Alembert in the bottom triangle, then you jump across the characteristics with constant Δu, and you then use techniques just discussed in Hyperbolic Example #1 to deduce the solution in the other regions. Very interesting, but I doubt anyone would do something this way, probably of historic interest.