Stakgold Chapter 5 Raw Notes PhL begun 5.30.09

Preface to this volume: grad students in engineering and physics are his customers. I think this is how this volume is going to pan out.

Chap 5: the multiple variable theory in as general a form as possible (72 pages)
Chap 6: PDE's that don't involve time (potential theory as example) (84 pages)
Chap 7: PDE's that do involve time ("evolution", heat and wave as examples) (115 pages)
Chap 8: Variational and other special methods (48 pages)

CHAP 5: DISTRIBUTIONS AND GENERALIZED SOLUTIONS (1) 3

5.1 Introduction (1) 3

5.2 Test Functions (3) 4

5.3 Distributions (4) 5

5.4 Convergence of Distributions (10) 10

Distributions Depending on a Parameter (10) 10

Exercises: (14) 13

5.5 Additional Properties of Distributions (17) 14

A. The direct product of two distributions. (17) 14

B. The convolution product of two distributions. (18) 15

C. Delta function of a function $\delta[f(x)]$ 17

D. Change of coordinate system (21) 17

Exercises (22) 18

5.6 Fourier Transforms (23) 20

A. Classical Results 20

C. One Sided Functions (28) 22

D. Functions of Slow Growth. (29) 22

E. Transforms of Distributions on the Line (ie, one variable). (30) 23

About the S_1 test function space. 25

Properties of the new S_1 space distributions 25

The Fourier Transform of a distribution is also a distribution 27

Properties of Distributions 29

Examples of Fourier Transforms of distributions: (34) 29

F. Transforms in more than one variable (36) 31

Exercises (37) 31

5.20 A Compute the FT of pf(1/x). 32

5.20 B Compute the FT of sgn(x). 34

5.20 C Compute the FT of log |x| 35

5.21 Compute the FT of x^k 35

Large digression on Spherical Coordinates in n dimensions 36

Resume Exercise 5.22 40

5.7 PDE's for Distributions (39) 41

Claimed Theorem 5.72: 41

Examples: (40) 42

Example 1: 42
Example 2: 42
Example 3: 42
Example 4: 44
(a) Prove p 41 A 44
(b) Question: what is this "transversal vector q" thing? 44
(c) Prove 5.83. 45
Example 4A: 45
Strict and Generalized Solutions of Differential Equations (42) 46
Homogeneous Equation (42) 47
Theorem 1: 47
Theorem 2: 47
Example 1 (p 44) Wave with 1D space. 48
(a) Derive p 44 A 48
(a) Derive p 45 A 49
Example 2 (p 46 bottom) Laplace in n dimensions. 51
Exercises (47) 51
Exercise 5.23. 52
Exercise 5.24. 53
Exercise 5.25. 54
Exercise 5.26. 54
(a) Try if first myself. 54
(b) Look at the Hints 56
Exercise 5.27. 59
Exercise 5.28. 59
5.8 Fundamental Solutions (48) 60
(1) Laplace Equation (49). 61
Comments on the Divergence Theorem 62
(2) Helmholtz Equation (53). 62
(3) The Heat Equation (58). 64
(4) The Wave Equation (61). -- undamped 65
(4) The Wave Equation (65) -- damped 66
Digression on an Integral. 68
Question about those radial solutions. 70
Exercises (69-72) 71
5.9 Classification of PDEs (72) 71
The Cauchy Problem. (73) 71
Introduction (72) 72
First order PDE's 76
Cauchy problem for first order PDE (76) 76
Cauchy problem for second order PDE (79) 77
Examples (81) 78
Hyperbolic Example 1: $L = \partial_{xy}$ and $Lu = 0$ (81) 78
Hyperbolic Example 2: $L = \partial_t^2 - c^2\partial_x^2$ and $Lu = 0$ 79

Chap 5: Distributions and Generalized Solutions (1)

5.1 Introduction (1)

This and following sections are going to generalize the one-variable theory we learned in volume 1, especially in Chapter 1 where we dealt with 1-variable distribution theory. We now have n variables and we need a new and compact notation for things:

$x = x_1, x_2....x_n$	our variable vector now has n, not just 1 component he does not write it bold as **x**, just x.
$x \in R_n$	our variable is in Euclidean n space, not some strange space
$r = \| x \|$	the "radius" of a vector, distance to the origin from vector tip.
dx	the full volume integration, I might have used dV
σ and dS	dS is a patch on a surface of dimension n-1 ONLY. σ is the name for a finite region of such a surface. He calls this kind of surface a hypersurface of dim n-1.

Every single thing we learned earlier comes back here in n dimensions. [I think the importance of surfaces of dimension n-1 is that Green's theorems related them to volumes of dimension n.]

def: **a LI function** is one which can be integrated dx (=dV) over any bounded region R in R_n.

Example: $f(x) = 1/r^k$ is LI if $k < n$. The reader knows how to show this when n = 1,2,3 no doubt, but not obvious how you express dV in terms of "spherical coordinates" in n dimensions. I recall doing this somewhere once. [more on this below]

def: f is **LI on σ** if the dS integral is finite for any bounded region σ' on σ.

So now we have at least two kinds of LI to worry about instead of 1. [but really the same idea]

def: If you examine the set of all points x where $f(x) \neq 0$ and then *close* that set, you have the **support** of the function f(x). Example: f(x) = sin(x) has the entire real axis as its support.

def: **multiindex**. Since we have more dimensions n, there are lots of different derivatives we can construct. A very compact notation is to define a vector $k = (k_1, k_2, ...k_n)$ which vector is called a **multiindex**. A simple first derivative with respect to x_1 would have k = (1,0,0..). Then ∂_{x1}^2 has (2,0...). And then $\partial_{x1}\partial_{x3}^2$ would have k = (1,0,2,0...). The sum of the k_i we can see is somehow the total number of derivatives but he has no word for this concept, so I will invent one: $k = \Sigma_{i=1}^n k_i$ is the **D-order** of the multiindex, whereas n is the **order** of the multiindex. Maybe terms will appear later. A compact notation for the derivative implied by a multiindex k is this: D^k .

description of **differential operator L or order p**: $L = \Sigma_{|k| \leq p} a_k(x) D^k$

So we are now reusing the word "order" to describe the highest order derivative which occurs in our operator L, in the sense that we add the orders of the derivative components as in $k = \Sigma_{i=1}^{n} k_i$. This sum is really a sum over the n variables k_i subject to the condition that the sum is constrained to the lattice in the "first quadrant" which is a "diagonally sliced cube". For n = 3, this is true. For n = 2 the lattice is a triangle in the first quadrant. For n > 3, it is whatever it is, a diagonally sliced hypercube.

The functions $a_k(x)$ are the analogs of our old friends a_o and a_1 and a_2 of Volume 1. But now each of these scalar functions has a multiindex and so is really $a_{k1k2k3..kn}(x)$. There is one function for each term in the sum of course, and there might be a lot of terms.

An **interesting question**: given p and given n, what is the maximum number of terms you could have in the sum for L ? For n = 2 and p = 2, I can count the lattice points of the triangle to get 6. They are

(2,0) (1,1) (0,2) (1,0), (0,1) (0,0)

For n = 3 and p = 2, let's try the count

2,0,0	0,2,0	0,0,2	// always n terms here
1,1,0	1,0,1	0,1,1	
1,0,0	0,1,0	0,0,1	// always n terms here
0,0,0			// always 1 term here

What is the general case solution to this problem?

0,0,...0	1 term
1,0,0...0 + perms	n terms as long as $p \geq 1$
1,1,0...0 + perms	(n,2) terms as long as $p \geq 2$ and $n \geq 2$
2,0,0...0 + perms	n terms as long as $p \geq 2$ and $n \geq 1$
1,1,1,0...0 + perms	(n,3) terms as long as $p \geq 3$ and $n \geq 3$
2,1,0...0 + perms	2*(n,2) terms as long as $p \geq 3$ and $n \geq 2$
3,0....0 + perms	(n,1) terms as long as $p \geq 3$ and $n \geq 1$

You can see that the answer will be some messy expression involving theta functions on p and n. The answer will be simple for very large n and/or p I suspect. Amusing, but not critical right now.

5.2 Test Functions (3)

If φ(x) is a "test function" on R_n every single multiindex derivative D^k of φ(x) must exist no matter how high we go with |k|. That is to say, φ(x) is C^∞ not just separately in each component variable x_i , but also in all combinations of derivatives. In other words, $D^k\varphi(x)$ exists for any $(k_1, k_2....k_n)$ n-tuple you write down. And as expected, φ(x) must have support only in a bounded region of R_n about the origin. For any

bounded region, we know there will be some maximum radius that contains the region and he calls that radius A on page 3. The space of such test functions $\varphi(x)$ is called K_n . This is a linear space. If $\varphi(x)$ is in the space, any derivative of φ is of course in the space as well. A little set of rules for $\varphi(x)$ appears on page 3 lower half. Luckily, we can always construct a specific test function as shown in (5.2) which has the same form as in our 1D work, with variable now set to r. But here he uses a instead of A as the bound.

Notion of a null sequence of test functions $\varphi_m(x)$. The idea here is that as n increases, any derivative gets small so that $\lim_{m\to\infty} D^k\varphi_m = 0$. Of course the φ_m all have some common bound $r \leq A$. This subject was mentioned in passing in volume 1. No doubt he will make use of such a "test function null sequence" in proving things later in this chapter.

5.3 Distributions (4)

He uses his old notation $\langle t,\varphi\rangle$ for a functional [I prefer to use (t,φ) and reserve $\langle t,\varphi\rangle$ for the inner product in a Hilbert Space -- this is what I did in my earlier notes .] If we restrict to a continuous linear functional, and if φ is an arbitrary test function in K_n , then this thing (t,φ) gets the name "a distribution" and sometimes one refers to the "action of t on φ". The "continuous" element here is important. It means that if you consider *any* test function null sequence φ_m, then we should have $\lim_{m\to\infty} (t,\varphi_m) = 0$. So good, right off the bat we have a use for this null sequence we just defined.

Theorem 1 (no surprise). Every LI function f(x) in R_n can be associated with a distribution "f" . All we have to do is identify (t,φ) with the inner product $\langle t,\varphi\rangle$ in $L_2^{(n)}$, the space of square integrable functions in n Euclidean dimensions. As in vol 1, we are starting I think with real functions. The notation back on page 1 implies that, for the moment, the components of x_i are real, R_n is real Euclidean space, etc. At this point he does not mention $L_2^{(n)}$, he just shows the integral that is the inner product. The proof of linearity is easy, and so is the proof of continuity of the linear functional so defined by the inner product.

def: If a distribution "f" is associated with a LI function, then it is **regular**, else it is **singular**. These words also appeared in volume 1. If singular, we can go ahead and use the $\langle t,\varphi\rangle$ inner product and represent t as a **generalized (symbolic) function**, again words from volume 1.

Two LI functions which are "equal almost everywhere" produce the same distributions.

Page 5 top talks about **adding** and **scaling** distributions and defines the **zero distribution**.

Examples:

1. Consider $(H_R,\varphi) \equiv \int_R dx\ \varphi(x)$, just the integral over region R. Obviously this is a real number and exists for any finite R. It is linear and no doubt continuous, so it is a distribution. The associated function here is the **n-dim generalized Heaviside function** defined on region R, so that $H_R(x) = 1$ inside R and 0 outside. So write integral as $(H_R,\varphi) = \int dx\ H_R(x)\varphi(x)$ where now we have a full R_n integral. Since this distribution can be associated with an ordinary LI function, namely $H_R(x)$, it must be regular.

2. The **n-dimensional delta** function business. Start with $(\delta_\xi, \varphi) = \varphi(\xi)$. Linearity and continuity are easy to show, so it is a distribution (again, the null sequence idea is employed). There is a corresponding generalized function called $\delta_\xi(x)$ as shown p5 A. This is a "source at $x = \xi$ " .

3. **Translation**. Do the usual shift of integration variable to get the general translation formula shown in (5.4). If we set $t = \delta$ in this general rule, we get p 6 A. The rightmost $\varphi(\xi)$ implies that we know what we mean by this δ distribution, he has not said it explicitly. Normally we would have $\varphi(0)$ on the right but the thing is shifted. Comparing to our earlier p 5 δ_ξ business, we get p 6 B, no surprise here.

4. **Dipole**. I don't much like his presentation here. He uses $\hat{\ell}$ as a unit vector and then defines the derivative $d\varphi/d\ell$ as shown. The charges are $+1/\varepsilon$ and $-1/\varepsilon$ separated by distance ε about point ξ in R_n. I think this derivative should really be written as $\hat{\ell}\bullet\nabla\varphi$ and this is what he means by $d\varphi/d\ell(\xi)$. But we are talking about a functional I will call $(D_\ell, \varphi) = d\varphi/d\ell(\xi)$. Yes, it is linear and it is continuous (again you would invoke the null sequence test function thing, and here we actually have some derivatives). He has made this LOOK LIKE the dipole thing we had in volume 1 which was just $(D,\varphi) = d\varphi/dx(\xi)$. So yes, the functional $(D_\ell, \varphi) = d\varphi/d\ell(\xi)$ is a distribution, and you can call it a **dipole distribution**. If we compare it to the distribution $(\delta,\varphi) = \varphi(\xi)$ which was the Dirac delta distribution, we should really call our new thing the Dirac dipole distribution with respect to direction $\hat{\ell}$. I might go on to write this as

$$(D_\ell, \varphi) = d\varphi/d\ell(\xi) = \hat{\ell}\bullet\nabla\varphi(\xi) = \int dx\ \hat{\ell}\bullet\nabla\varphi(x)\ \delta(x-\xi) = -\int dx[\hat{\ell}\bullet\nabla\delta(x-\xi)\ ]\ \varphi(x)$$

$$= -\int dx\ [\nabla_\ell\ \delta(x-\xi)]\ \varphi(x)$$

so the corresponding dipole generalized function is this thing $\nabla_\ell\ \delta(x-\xi)$. It is the derivative of the delta function in some specific direction $\hat{\ell}$. Stak does not show all this extra detail at this point.

5. **Scaling**. Here we get the general scaling rule for a distribution, using the usual change of variables, but this applies only to a regular distribution where we have a LI function f(x). Notice that all components of x are scaled the same amount, and implied is the fact that $dy = \alpha^n dx$ and as in vol 1, $|\alpha|$ is what is really involved here, see (5.5).

6. **Hypersurface simple layer.** In 5.6 we try to define a distribution (t,φ) as a different kind of integral. This integral is not over all of R_n, it is over some n-1 dim surface in R_n with a weight function a(x) which is LI on σ. Think of a(x) as a 2D surface density of charge on a 2D surface in R_3. So you have a tiny charge near point ξ I would call $dQ(\xi) = a(\xi)dS(\xi)$ in units say of Coulombs. The question then arises: what is the 3D charge *density* we normally write as $d\rho(x)$ due just to this piece of surface ? The claim is this:

$d\rho(x) = a(\xi)dS(\xi)\ \delta(x-\xi)$ charge per cm^3 due *just* to this little piece of surface charge

Notice that $\delta(x-\xi)$ in this case is really $\delta^3(x-\xi)$ which has dimensions cm^{-3}. Now if we sum over all pieces of surface charge, we get what he calls "the total charge density" which I will still write as $\rho(x)$

$$\rho(x) = \int_{\sigma} a(\xi) dS(\xi) \delta(x-\xi)$$

where here the integral volume is dS and we integrate over surface σ. So this is what p 7 A is saying, and my old pencil annotations there are OK.

So let's now get back to the notion of a distribution here. We have:

$$(t,\varphi) = \int_{\sigma} a(\xi) dS(\xi) \varphi(\xi) = \int dx\, \delta(x-\xi) \int_{\sigma} a(\xi) dS(\xi)\, \varphi(\xi) = \int dx\, \rho(x)\, \varphi(x)$$

so in this case, then, our "generalized function" is just ρ(x). In equation p 7 B on the LHS he puts this thing inside the functional form, basically writing there (ρ,φ) (with the large <> symbols). So equation B is just what I have written above. And I have also produced C on the same line.

Question: is this a regular or a singular distribution? Is ρ(x) a regular function? I don't think so. It is more like a delta function. Any time your charge density ρ(x) is not a normal 3D function but has charge constrained on some subsurface within R_3, you have a singular situation. Whether a point, line or surface charge. If you avoid using the symbol ρ, you have to write something like p 7 C. [In general, if sources are distributed on surfaces of dimension less than n and you are working in R_n, then those sources are not functions but distributions.]

Product of a distribution and a function. You get to say (5.7) only if the function f(x) is C^{∞} , though there are ad hoc cases where f(x) can be less restricted. So this is just one of our "rules" for distribution manipulation in n dimensions, like the scaling and translation rules.

Differentiation of a distribution. Page 8 (5.8) gives the *definition* of the derivative of a distribution where we consider k = (1,0,0...) as our sample derivative. The minus sign of course is from the parts integration and the test function makes the actual parts vanish.

Then we want to look at an arbitrary multiindex derivative as in (5.9). We pick up $(-1)^{k1}$ from casting over the $\partial_{x1}{}^{k1}$ derivative by parts. Stak uses the notation |k| to stand for the sum of the k_i. He uses this not because anything might be negative or complex, but just because he needs some notation to represent this sum. He could have called it K or S = Σk_i. So I am happy with (5.9).

Comment: Notice in (5.9) that we are able to describe the action of D^k on distribution t in terms of the action of D^k on an arbitrary test function in the space K_n.

Examples: (8) After each dose of "theory", Stak wants to bring us back down to earth.

1. This is an example of a distribution. The distribution is named by Stak $t = -\partial_{x1}\delta(x-\xi)$, but really this is the name of the generalized function, I might say t = "$-\partial_{x1}\delta(x-\xi)$" . We do a single parts swing to get the result that $(t,\varphi) = \partial_{x1}\varphi(\xi)$. So this is an x_1 axis dipole distribution, we saw more general earlier.

2. The distribution of this example is t = "f(x)δ(x-ξ)", which is just the delta multiplied by any continuous function, and we find (t,φ) = f(ξ)φ(ξ). This leads to obvious rule (5.10).

3. Now we combine that last two and have t = "f(x) $\partial_{x1}\delta(x-\xi)$" as our new distribution. If we write this out in generalized function language, we have

$$(t,\varphi) = \int dx\, f(x)\, \partial_{x1}\delta(x-\xi)\, \varphi(x) = -\int dx\, \delta(x-\xi)\, \partial_{x1}\, [f(x)\, \varphi(x)\,] = -(\partial_{x1}\, [f(x)\, \varphi(x)\,])\,|^{x=\xi}$$

$$= -\, f(\xi)\, \partial_{x1}\varphi(\xi) \; - [\partial_{x1}\, f(\xi)]\, \varphi(\xi) \qquad (*)$$

So one would use this last line as the "definition" of the distribution t = "f(x) $\partial_{x1}\delta(x-\xi)$" and it works for f(x) being any function that is continuous and differentiable wrt x_1 for all x_1 values in space R of interest. Notice that R in R_m replaces the interval (a,b) of Volume 1.

The next step to get (5.11) is not obvious! Here is what 5.11 claims:

$$f(x)\, \partial_{x1}\delta(x-\xi) = f(\xi)\, \partial_{x1}\delta(x-\xi) \; - [\partial_{x1} f(\xi)]\, \delta(x-\xi) \qquad (5.11)$$

It is claimed this is true with everything being a GF. To see whether this is really true, add back the integral and the test function:

$$\int dx\, f(x)\, [\partial_{x1}\delta(x-\xi)]\, \varphi(x) = \int dx\, f(\xi)\, [\partial_{x1}\delta(x-\xi)]\, \varphi(x) \; - \int dx\, [\partial_{x1} f(\xi)]\, \delta(x-\xi)\, \varphi(x) \qquad (?)$$

The first term on the RHS can be written

$$\int dx\, f(\xi)\, [\partial_{x1}\delta(x-\xi)]\, \varphi(x) = f(\xi) \int dx\, [\partial_{x1}\delta(x-\xi)]\, \varphi(x) = -\, f(\xi) \int dx\, \delta(x-\xi)\, [\partial_{x1}\, \varphi(x)]$$

$$= -\, f(\xi)\, [\partial_{x1}\varphi(\xi)]$$

The second term on the RHS of (?) can be written

$$-\int dx\, [\partial_{x1} f(\xi)]\, \delta(x-\xi)\, \varphi(x) = -\, \partial_{x1} f(\xi)\, \varphi(\xi)$$

Thus, we have shown that the equation (?) can be written

$$\int dx\, f(x)\, [\partial_{x1}\delta(x-\xi)]\, \varphi(x) = -\, f(\xi)\, [\partial_{x1}\varphi(\xi)] - \partial_{x1} f(\xi)\, \varphi(\xi)$$

But this is the same as equation (*) which I have already proven to be true. Thus, equation (?) is true, and thus in terms of GF's (5.11) is true. I cannot find an obvious faster way to prove 5.11.

Comment: Keep in mind: to test any claimed equation in symbolic functions, you "add back" the integral against an arbitrary test function $\varphi(x)$ in K_n and "see what happens".

4. For this example of a distribution, we have t_1 = Lt where L is our full $L = \Sigma_{|k|\le p}\, a_k(x)\, D^k$, the most general multi-variable differential operator of degree p, where D^k has the meaning discussed earlier. As

we "swing over" each derivative $\partial_{x_i}^{k_i}$ we pick up $(-1)^{k_i}$ and we "hit" both $\varphi(x)$ and $a_k(x)$ [you would just write out the integral and do the parts swings one variable at a time]. When you are all done, the sign is then $(-1)^{|k|}$ where $|k|$ has that funny meaning $\Sigma_i\ k_i$ and nothing is negative or complex as noted earlier. If we think of equation A as saying $(Lt,\varphi) = (t,L^*\varphi)$, then $L^*\varphi$ must be as shown in 5.12. And if you are lucky enough to get $L = L^*$, your L is "**formally self adjoint**". [We have not yet brought in the subject of boundary conditions, so we are not yet talking about the domains of L and L* etc etc.]

The **Laplacian** example is interesting. It is the sum of single second derivatives, so sign is +. The various involved functions $a_k(x)$ are all unity. Thus, $L^*\varphi = \Sigma\ D^k\varphi$ which is same as $L\varphi$ and so that Laplacian is an example of a formally self adjoint differential operator in n-space. Maybe this fact explains why this thing appears in so many physical places.

Values of a GF. First, Stak notes that we already know that you can have two LI functions which given the same distribution but the functions differ at some isolated points. As mentioned on page 4, this occurs when f_1 and f_2 are "equal almost everywhere" which means they can differ at some points but the integral of $|f_1-f_2|$ is zero. These two LI functions are examples of GF's which are also regular functions, and even in this mild case, we conclude that knowledge of the distribution (t,φ) cannot tell us the exact value of the GF at every point. So we have to give up on the idea that a GF has a well-defined value everywhere. We certainly know this for $\delta(x)$ which has no value at $x = 0$. [shades of uncertainty principle]

But we can grasp onto one little idea as follows. Suppose in some region R we have $(t,\varphi) = 0$ for all test functions supported within R. If this is true, it is reasonable to say that the GF t(x) vanishes at all x in this region R.

Example 0: suppose $(t_1-t_2,\varphi) = 0$ for all test functions φ supported in R. Then we claim that the two generalized functions $t_1(x)$ and $t_2(x)$ are "equal", where we are defining this word equal by this fact. Consider $f(\mathbf{x})\delta(\mathbf{x})$ and $f(\mathbf{0})\delta(\mathbf{x})$ which are two generalized functions and I bold things to emphasize the fact we are in n-space. They don't "look" equal. But if we let R be the entire R_m except the origin, we know that $(t_1-t_2,\varphi) = 0$ for all test functions φ supported within R. In this example, in fact, $(t_1,\varphi) = (t_2,\varphi) = 0$ (because $\varphi(0) = 0$ for φ supported on this R). So our two GF's here are then "equal" for such R. But we can see they are also equal at the origin! So they must be "equal" over all of R_m.

Example 1: I certainly agree with equation p 10 A given that R excludes the origin. Thus, the two generalized functions $\delta(x)$ and $D^k\delta(x)$ are 0 for any such region, which means are zero for $x \neq 0$.

Example 2: Here we have a charge distributed on a surface σ within our space R_m. Here if we select as our R all regions excluding that surface (as we excluded $x = 0$ above), then we get $(t,\varphi) = 0$ for all test functions with support in R, and we conclude that the GF for this distribution vanishes off the surface, that is to say, for x not on σ. Here the GF is in fact the dS integral, as shown p 7 B,

$$t(x) = \int dS_\xi\ a(\xi)\delta(x-\xi) \quad = \text{the GF of this example.}$$

This vanishing of the GF seems pretty obvious, but Stak has put it on a solid ground in this little section. He has given a rigorous meaning to the idea of a GF vanishing away from the "places where the charge is located". This meaning involves the distribution and its test functions on a region, and does not depend on the GF directly. This is good since the GF is only a symbolic object and does not exist as a "point function".

5.4 Convergence of Distributions (10)

Distributions Depending on a Parameter (10) . This is another rehash of Vol 1 stuff, generalized to our n-space. t_α is a distribution depending on a parameter α. It seems pretty reasonable to say that the distribution converges to some t (that is, $t_\alpha \to t$) if $(t_\alpha , \varphi) \to (t, \varphi)$ for ANY test function φ. The claim is that this limiting thing "t" is in fact a distribution. As usual, we consider the case where our parameter α is really an integer m and we are talking about a sequence s_m which is the partial sum of some t_i up to m. You are allowed to do formal termwise differentiation, a Vol 1 topic. Recall that you can differentiate any way you want and the result is another distribution, guaranteed. [a feature of K_n!] In his proof on page 11 he just shows that a single ∂_{x1} derivative is sort of "transparent" to our parameter limit action. You kick the derivative onto the test function, take your limit, then kick it back. This would then be true for any D^k fancy derivative.

Example 1: Differentiating with respect to a parameter. The idea is shown in p 11 A. If the limit shown exists, then you can "differentiate with respect to α" and you can talk about $dt_\alpha/d\alpha$. This may or may not be possible, he points out, but in contrast $\partial_{x1}t$ always exists. Now as our example, we consider the component ξ_1 of ξ to be a "parameter", where ξ is just some point in R_n. He shows in detail that result 5.14 is true as a statement involving GF's. This is typical Stak. The reader knows that 5.14 is true and has used it for decades, but here we have a formal proof that it is true in the context of "distributions which depend on a parameter".

Example 2: Back on page 6 example 6 we dealt with a "simple layer" which I think of as charge on a surface, but described in the language of distributions. On page 7 pre-A we had the charge density dρ(x) due to just a differential piece of our surface σ, then in A we summed over σ to get ρ(x), the charge distribution everywhere in space due to our charged surface.

In this example we have a "dipole layer" which I think of as dipoles lined up like the lipid molecules in a cell wall. There is some function $\hat{\mathbf{n}}(\mathbf{x})$ defined for points x on σ which gives the local normal to the surface. As before, we have to "decode" the Stak notation to make any sense out of this concept. A differential piece of σ at x = ξ yields a dipole moment density I might call $dp_\xi(x)$ which is the magnitude of a vector pointing in the direction $\hat{\mathbf{n}}(\mathbf{x})$, where x lies on σ. If you then integrate over σ, you get some p(x) which again is a magnitude and points in $\hat{\mathbf{n}}(\mathbf{x})$. This p(x) agrees with $p_\xi(x)$ at each point ξ, we are not summing somehow to get a new vector dipole moment. We are just extending $p_\xi(x)$ to include the entire surface σ, and we call that extension p(x). The idea is uncomfortable for me. I would be much happier with some well-defined bolded vector notation in R_3 and THEN go to the generalized notation Stak is using. I'll bet his students in the 1950's had the same problem I am

Here we have the notion that a "distribution" is some charge or dipole or source that is "distributed" on some subsurface in your space R_m, and maybe that is where the word "distribution" comes from. If your sources are distributed smoothly in R_m, you have a non-singular distribution, but if they are distributed on lower dimensional surfaces in R_m, then you have a singular distribution.

Example 3: Another obscure Stakgold example. We start with some arbitrary but positive scalar LI f(**x**) and define a "scaling" sequence f_α based on it, as shown p 12A. The claim is that as $\alpha \to 0$, this becomes δ(**x**). This generalizes something we did in 1D in vol 1. First we can verify the three p 13 top properties:

(a) $\int dx\ f_\alpha(x) = \int dx\ \alpha^{-n}\ f(x/\alpha) = \int du\ f(u) = 1$ by hypothesis of our theorem. So every function in our parameterized sequence has unit volume.

(b) $\int_{|x|>A} dx\ f_\alpha(x) = \int_{|u|>A/\alpha} du\ \alpha^n\ f_\alpha(x) = \int_{|u|>A/\alpha} du\ f(x/\alpha) = \int_{|u|>A/\alpha} du\ f(u)$.
But as $\alpha \to 0$, our integration restriction becomes $|u| > \infty$ and so picks up nothing, answer is 0.

(c) Here same thing but condition is $|u| < \infty$ and picks up everything, so gives unity.

He really wants to show this, so he can make the claim that in GF we have $\lim_{\alpha\to 0} f_\alpha(x) = \delta(x)$:

$\lim_{\alpha\to 0} \int dx\ f_\alpha(x)\ \varphi(x) = \int dx\ \delta(x)\ \varphi(x) = \varphi(0)$ // ie, we fill out the GF meaning with $\int$

but he rephrases this right at the start in A

$\lim_{\alpha\to 0} \int dx\ f_\alpha(x)\ [\varphi(x) - \varphi(0)] = \int dx\ \delta(x)\ [\varphi(x) - \varphi(0)] = \varphi(0) - \varphi(0) = 0$

I guess if we define $\eta(x) = [\varphi(x) - \varphi(0)]$, it is just some "arbitrary" test function which has the property that it vanishes at x = 0, so OK, we want to show this

$\lim_{\alpha\to 0} \int dx\ f_\alpha(x)\ \eta(x) = 0$ for all test functions $\eta(x)$ which have property $\eta(0) = 0$.

He writes this integral as two terms, inside and outside a sphere or radius A. He shows that the LHS can be made arbitrarily small as you drive $\alpha \to 0$, so in the limit it must be 0 and our proof is complete.

Example 4: Well, now finally we come to the issue of "**spherical coordinates**". I think we want to say here that $dV_n = r^{n-1} dr\ d\Omega_n$ where $d\Omega_n$ is some generalized "solid angle" in n dimensions. When n=3 we know that $\int d\Omega_3 = 4\pi$, and when n = 2 we know $\int d\Omega_2 = 2\pi$. In general he calls this integral $S_n(1)$ in B on page 13.

Now we want to go back to Example 3 and set f(x) = g(r). In Example 3 there were lots of functions f(x) we could have used to get a delta function in the limit, and one subset of all these functions are functions which are functions only of the radius r, and he calls this function g(r) as case of $f(x_1, x_2....)$. The first thing we need to make sure about g(r) is that it has the right scale, so we need this to be true:

$\int dV\ g(r) = 1 \quad \Rightarrow \quad \int r^{n-1} dr\ d\Omega_n\ g(r) = 1 \quad \Rightarrow \quad S_n(1) \int r^{n-1} dr\ g(r) = 1$

$\Rightarrow \int r^{n-1} dr\ g(r) = 1/\ S_n(1)$

and for n=3 he writes this in p 14 A. Of course g(r) also has to be non-negative, according to our theorem. For any such g(r) we come up with, we can make a delta function δ(x) as the limit shown p 14 B !

He then comes up with an example for the n=3 world

$$g(r) = \pi^{-3/2} \exp(-r^2)$$

$$\int r^2 dr\, g(r) = \pi^{-3/2} \int r^2 dr \exp(-r^2) = \pi^{-3/2} (\pi^{1/2}/4) = 1/(4\pi) \text{ as required.}$$

This then gives the explicit form p 14 C for δ(x) as a limit of this radial function as shown.

Comment: it is fascinating that we can create a full blown Cartesian delta δ(**x**) in n dimensions by just using a scalar function and taking a limit! We know that somehow δ(**x**) is non-zero only "at" r=0 so that all three Cartesian variables have to be 0. We know that $\exp(-Ay^2)\, y^p \to 0$ as $y \to \infty$ for any power p, and this fact tells us that our limit is 0 away from r = 0. "At" r=0, we have to look at the sifting integral property to see that it is right for δ(x) and that is just what we did in the proof on page 13. Without this general proof, I would have a hard time showing the claimed fact:

$$\lim_{\alpha \to 0} [\, \pi^{-3/2} \exp(-r^2/\alpha^2) / \alpha^3 \,] = \delta(x_1)\delta(x_2)\delta(x_3)$$

This is something brand-new to me, I don't remember seeing it ever before. Notice that dim(α) = cm so that both sides have dimensions cm^{-3}.

Example 5: I am not sure what these are examples of, but in this example he defines u_R to be a function which vanishes outside some region R (think in R_3). We know that ∇^2 is formally self-adjoint so we get the first equality. Then we just write out the $(u_R, \nabla^2\varphi)$ distribution as its integral and integral is then restricted to the region and we can replace u_R with u. Then we use Green's Second identity to finish the line. You can then *interpret* the Green's term as the surface integrals of a charge "simple layer" and of a charge dipole "double layer".

So in this example we have a distribution called "$\nabla^2 u_R$" and we compute "what it is" on line 5.16. This distribution depends on the region R and on the function u. φ of course is just a test function. Any time you have a distribution (something, φ), you always want to compute "what it is". For example, we earlier computed (δ,φ) = φ(0), so we know what it is. Here we just have a more complicated animal. The entire line 5.16 is a continuous linear functional which acts on test functions and depends parametrically on u and R. This section of examples is about distributions that have parameters like this. In the δ(x) stuff above, the parameter was α. Here the parameter is u_R which is a function u and a region R.

A point not to be missed: the distribution $\nabla^2 u_R$ can be *completely described* in terms of a simple layer and a double layer on any surface σ bounding region R. In electrostatics (potential theory), this means that if we know that potential u = 0 outside R, then $\nabla^2 u$ inside R is completely determined by that value of u and its normal derivative on the surface bounding R. Often $\partial u/\partial n = 0$ so we have only the value of u involved. Recall how, in a 1D ODE of second order, u(a) and u'(a) fixes the solution everywhere in an interval (a,b) -- initial boundary values. Or maybe u(a) and u(b). Here we are seeing the generalization of this idea, but he is not saying much about it at this point.

For me, this was a rough Stakgold section.

Exercises: (14)

5.1 **Volume and Surface Area of an n-dimensional sphere.** Stak shows how you can derive these formulas, and all results are stated and we re-derive page 13 B as (5.19).

5.2 **Other δ(x) radial limit formulas n = 2 and n=3**. Recall that in general we can use $f_\alpha(x) = g(r/\alpha)/\alpha^n$ where g is any positive function which is scaled so integral is as shown page 13 B. In this exercise we seem to have

n = 3 $\quad f_\alpha = \alpha\,(1/\pi^2)/\,(r^2+\alpha^2)^2 = \alpha^{-3}\,(1/\pi^2)\,(\,[r/\alpha]^2 + 1)^2$

=> $g(r) = (1/\pi^2)/\,(r^2+1)^2$ which must be normalized properly $S_3(1) = 4\pi$

$\int dr\, r^2\,(1/\pi^2)/\,(r^2+1)^2 = 1/4\pi = (1/\pi^2)\int dr\, r^2/\,(r^2+\alpha^2)^2$

=> $\int dr\, r^2/\,(r^2+1)^2 = \pi/4$ as Maple verifies

n = 2 $\quad f_\alpha = \alpha\,(1/2\pi)/\,(r^2+\alpha^2)^{3/2} = \alpha^{-2}\,(1/2\pi)\,(\,[r/\alpha]^2 + 1)^{3/2}$

=> $g(r) = (1/2\pi)\,(\,r^2 + 1)^{3/2}$ which must be normalized properly $S_2(1) = 2\pi$

$\int dr\, r\,(1/2\pi)/\,(r^2+1)^{3/2} = 1/2\pi = (1/2\pi)\int dr\, r\,/\,(r^2+1)^{3/2}$

=> $\int dr\, r/\,(r^2+1)^{3/2} = 1$ as Maple again verifies

5.3 A δ(x) radial limit for general n.

n = n $\quad f_\alpha = \alpha^{-n}\,(1/\,\pi^{n/2})\exp(-r^2/\alpha^2)$

=> $g(r) = (1/\,\pi^{n/2})\exp(-r^2)$ which must be normalized properly $S_n(1) = n\,\pi^{n/2}\,/\,(n/2)!$

$\int dr\, r^{n-1}\,(1/\,\pi^{n/2})\exp(-r^2) = (n/2)!/\,[\,n\,\pi^{n/2}]$

=> $\int dr\, r^{n-1}\exp(-r^2) = (n/2)!/\,n$ which Maple verifies as $(1/2)\Gamma(n/2)$

$(1/2)\Gamma(n/2) = (1/2)\,\Gamma(n/2+1)/(n/2) = \Gamma(n/2+1)/n = (n/2)!/\,n$ so checks.

5.4 This seems to be just another way to formulate a general delta limit. Here we insist that the integral be zero in any finite shell but be 1 inside any finite sphere.

5.5 This 1D example does not fit our general formula that $f_\alpha(x) = g(r/\alpha)/\alpha^n$ because the thing in the expo does not scale right,

n = 1 $f_\alpha(r) = \alpha\,(1\sqrt{4\pi}\,)\,/\exp(-\alpha^2/4r)/r^{3/2}$

But using the approach of 5.4, he shows this also gives δ(x). I am not doing this one.

5.6 and 5.7. In these two examples, we combine our limit with an integration against an arbitrary function u(x) in the n=1 world. The limit thing is more or less acting like a delta function. The two cases here have slightly different conditions on f(x). The first has it being even so that int(0,∞) = 1/2. In the second, f(x) could still be even and you could set A = 1/2, and the answer then seems to conflict with the result for 5.6. I will not do this problem unless its result gets used later on.

5.8. Here we could think of α = 1/R and cast the famous sine δ(x) limit in our radial g(r) context.

n = 1 $f_\alpha(r) = \alpha^{-1}(1/\pi)\sin^2(x/\alpha) / (x/\alpha)^2$

=> $g(r) =(1/\pi)\sin^2(r) / r^2$ which must be normalized properly $S_1(1) = 2$

$\int dr\,(1/\pi)\sin^2(r) / r^2 = 1/2$

=> $\int dr\,\sin^2(r) / r^2 = \pi/2$ which Maple verifies

5.5 Additional Properties of Distributions (17)

A. The direct product of two distributions. (17)

This is pretty obvious, you have to have separate variables, then you get things like this:

$< t_1(x_1)\, t_2(x_2), \varphi(x_1, x_2) >$ = the obvious double integral

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

where the test function for t_1 is itself a distribution with distribution parameter x_1 being the argument of test function Φ. So this gives a pretty clean meaning to $t(x_1,x_2) = t_1(x_1)\, t_2(x_2)$. Notice that we do NOT have a meaning for something like this: $d(x_1) = t_1(x_1)\, t_2(x_1)$ where we have the same variable.

Examples 1,2: δ(x) in R_2 $= \delta(x_1)\,\delta(x_2)$ and similarly in n-dimensions

Example 3: Discussion here of direct product of t_1 = δ and t_2 = function of two variables. We can interpret δ(z)f(x,y) as a "source density" (source is a more general word I think than charge) which is distributed on the x-y plane at z = 0 with f(x,y). So it is a delta only in the z-sense, and is a regular

function in the other two variables. Fine. This is a typical situation that this chapter is addressing: there are lots of fancy situations involving "partial delta" actions in distributing sources for a problem.

B. The convolution product of two distributions. (18)

He first does this with two functions and defines the usual convolution integral as in 5.22. The bottom line in distribution language is this:

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

and in the reverse order if f_2 does the vanishing. The vanishing is required so that we can interpret this as a test function

$$\psi(x) \equiv \langle f_1(z), \varphi(x+z) \rangle = \int dz\, f_1(z)\, \varphi(x+z) = \int dz\, f_1(z-x)\, \varphi(z)$$

If f_1 is localized, there is a limit on how far it can "reach" in terms of $|z-x|$ and $\psi(x)$ will then also have a limited support and is thus a viable test function. He then says if you maintain this limited range idea, you can treat either or both f_1 and f_2 as a generalized function. Notice that the test function φ has a single variable which is shifted by a parameter "x". The distribution $\langle f_1(z), \varphi(x+z) \rangle$ is then a test function in variable x (which started as a parameter in a distribution). In contrast, in the direct product situation, the test function φ has two variables.

Using symbolic functions, we can say

$$(f_1 * f_2)(x) = h(x) = \int f_1(x-y) f_2(y)\, dy = \int f_2(x-y) f_1(y)\, dy$$

where we see the usual "convolution integral" $h(g_1) = \int dg\, f_1(g)\, f_2(g_1 g^{-1})$ in group notation. Notice that the thing is symmetric in the two functions (or distributions as they may be).

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*

It is not at all obvious to me why you need constant coefficients, given the definition of $L*\varphi$ shown in 5.12 of page 9, but I am sure Stak is right. I would have to write out all the steps in full notation to discover why this is so, and I suspect I will meanwhile learn this anyway soon because it will come up again. It probably has something to do with the fact that there are two variables floating around here.

Note added 6.25.09 to show why constant coefficients is needed here. The key fact as we shall see is this:

Lemma: $D^k_x = D^k_{x+b}$ for any complicated derivative in n-dimensions, if coefficients are constants.

This follows from the fact that $\partial_{xi} = \partial_{xi+b}$ for each partial.

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

$L_x = \Sigma_k a_k D^k{}_x = \Sigma_k a_k D^k{}_{x+b} = L_{x+b}$

Armed with these simple facts, we finish off Example 3. First:

$\langle L_x[(s*t)(x)], \varphi(x)\rangle$

$= \langle (s*t)(x), L^*{}_x \varphi(x)\rangle$ // the usual swing over of L_x

$= \langle (s*t)(x), \psi(x)\rangle$ // defining ψ

$= \langle t(x), \langle s(z), \psi(x+z)\rangle\rangle$ // from p 19A

$= \langle t(x), \langle s(z), L^*{}_{x+z} \varphi(x+z)\rangle\rangle$ // install ψ

$= \langle t(x), \langle s(z), L^*{}_z \varphi(x+z)\rangle\rangle$ // const coeff's => $L_x = L_{x+b}$ for any b

$= \langle t(x), \langle L_z s(z), \varphi(x+z)\rangle\rangle$ // swing L_z back to left side of its $\langle\rangle$

$= \langle [(Ls)\bullet t](x), \varphi(x)\rangle\rangle$ // by p 19A

and therefore we have shown that $L(s*t) = (Ls)*t$ which is the first part of 5.26. Here is a proof of the second part of 5.26:

$\langle L_x[(s*t)(x)], \varphi(x)\rangle$

$= \langle (s*t)(x), L^*{}_x \varphi(x)\rangle$ // the usual swing over of L_x

$= \langle (s*t)(x), \psi(x)\rangle$ // defining ψ

$= \langle s(x), \langle t(z), \psi(x+z)\rangle\rangle$ // from p 19B

$= \langle s(x), \langle t(z), L^*{}_{x+z} \varphi(x+z)\rangle\rangle$ // install ψ

$= \langle s(x), \langle t(z), L^*{}_z \varphi(x+z)\rangle\rangle$ // const coeff's => $L_x = L_{x+b}$ for any b

$= \langle s(x), \langle L_z t(z), \varphi(x+z)\rangle\rangle$ // swing L_z back to left side of its $\langle\rangle$

$= \langle [s*(Lt)](x), \varphi(x)\rangle\rangle$ // by p 19B

and therefore we have shown that $L(s*t) = s*(Lt)$ which is the second part of 5.26

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

Only has meaning if f(x) has zeros $f(x_i) = 0$ which all have multiplicity one (simple zeros). A good simple section, and good simple examples are given. The idea is to go to a small neighborhood near where the argument of the delta function vanishes in a simple linear fashion.

D. Change of coordinate system (21)

I went off today and wrote up a separate Jacobian document since that information was strewn all over my curvilinear notes. When the dust clears here, you end up with the sort of inverse Jacobian statement as shown in 5.29, and we get the usual spherical coordinates example where r,θ and φ are u_i so a bit confusing, as in (5.30). But J = 0 so get 1/J = ∞ on the RHS if your point x,y,z lies on the z axis, but Stak has a way around this that is a whole new little subject here I never heard of before.

Before continuing, notice that the usual J does not include an absolute value, but when you get to the delta form, it does and so he writes it as | J |. [Is this really true, or do both have abs value?]

His opening gambit is our usual spherical result which I will write here as

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

[p 22] The somewhat amazing claim made here is that you can write the 3D delta in the following way:

$$\delta^3(\mathbf{x}) = \delta(r)/[4\pi r^2]$$

If I had to verify this claim, I would say

$$\int d^3x\, \delta^3(\mathbf{x})\, f(\mathbf{x}) = \int\int r^2 dr\, d\Omega \;\{ \delta(r)/[4\pi r^2] \}\, f(\mathbf{x})$$

$$= \int d\Omega \int r^2 dr \;\{ \delta(r)/[4\pi r^2] \}\, f(\mathbf{x})$$

But the inner r integral is going to just be $1/4\pi * f(\mathbf{0})$ because really $\delta(r)\, f(\mathbf{x}) = \delta(r)\, f(\mathbf{0})$ and so $f(\mathbf{0})$ is a number F(0,0,0) which cannot depend on the angles Ω. Then we get 4π from the angular integral and we end up with the result $\int d^3x\, \delta^3(\mathbf{x})\, f(\mathbf{x}) = f(\mathbf{0})$ as required. This concept rings a very distant bell in my ancient past.

Now 5.31 makes this claim for two points **x** and **x'** given that θ' = 0 (so **x'** lies on the z axis)

$$\delta(\mathbf{x} - \mathbf{x}') = \delta(r - r')\,\delta(\theta) / [2\pi r^2 \sin\theta]$$

Let's again try a verification. We want to have

$$\int d^3x\, \delta^3(\mathbf{x} - \mathbf{x}')\, f(\mathbf{x}) = f(\mathbf{x}').$$

So write in sphericals as

$$\text{LHS} = \int\int r^2 dr\, d\Omega \; \{ \delta(r - r') \, \delta(\theta) / [2\pi r^2 \sin\theta] \} \; f(\mathbf{x})$$

$$= \int d\varphi \int dr \, \delta(r - r') \int d\theta \; \delta(\theta) / [2\pi] \; f(\mathbf{x})$$

$$= \int d\varphi \int dr \, \delta(r - r') \; 1/ [2\pi] \; f(r,0,\varphi)$$

But $f(r,0,\varphi)$ is really $F(0,0,r)$ in terms of Cartesian coordinates, so does not depend on φ so:

$$= \int d\varphi \int dr \, \delta(r - r') \; 1/ [2\pi] \; F(0,0,r) \;\; = \int dr \, \delta(r - r') \, F(0,0,r) \;\; = F(0,0,r') = f(r',0,0) = f(\mathbf{x}').$$

OK, this is certainly a weird little example.

One thing to notice in these examples: you *can* express $\delta^3(\mathbf{x} - \mathbf{x}')$ in terms of product of *fewer* than 3 delta functions. At first that seems very strange, but you see in $\delta^3(\mathbf{x}) = \delta(r)/[4\pi r^2]$ how the dimensions work out correctly.

Exercises (22)

5.9 This problem puts z' on the positive z axis. But I think this is what we just did above. I don't see why the answer should be different for being on the positive or negative z axis, a mystery to me.

5.10. Here he wants $\delta^n(\mathbf{x})$ in $\mathbf{R}_n$. To write this out as a product of n delta functions, I don't really know where to begin, I don't even know what the coordinates are. However, I think you could write the r-only version in this way: $\delta^n(\mathbf{x}) = \delta(r)/S_n(r)$ where $S_n(r)$ is the surface area of the n-dimensional sphere as given in 5.18.

5.11. Here he treats a **spherical shell of charge**, shell radius is ε. The first integral is his generic way to write a distribution on surface σ of dimension one less than n. The delta just pins 3D points onto this shell. It gives $\rho(x)$. The second form seems more useful. Let's try to verify it. $dA = \varepsilon^2 d\Omega$ so we have

$$\int_\sigma (1/4\pi\varepsilon^2) \, \delta(x-\xi) \, \varepsilon^2 d\Omega \;\; = \;\; (1/4\pi) \int d\Omega \, (1/r^2 \sin\theta) \, \delta(r-\varepsilon)\delta(\theta-\theta')\delta(\varphi-\varphi')$$

$$(1/4\pi) \int d\theta d\varphi \, (1/r^2) \, \delta(r-\varepsilon)\delta(\theta-\theta')\delta(\varphi-\varphi') \;\; = (1/4\pi\varepsilon^2) \, \delta(r-\varepsilon)$$

The total charge on this shell is of course $(1/4\pi\varepsilon^2) \;\; * \; 4\pi\varepsilon^2$ (ie, area) = 1 unit. If you take $\varepsilon \to 0$, I agree that the limit is a unit point charge. Therefore, it would seem that $\delta(r)$ by itself is NOT a unit point charge. But of course $\delta^3(\mathbf{r})$ is a unit point charge, have to be careful!

5.12 Another practical case: we put a **ring of charge** on a sphere or radius r_0 at angle θ_0. Now the charge is on a surface of dimension 2 less than that of the space, so our little dS_ξ formula does not work. Instead, we need some formula like this

$$\rho(x) = \int_C [\text{ linear density }]\, \delta(x-\xi)\, d\ell_\xi$$

where $d\ell_\xi$ is a piece of the curve and ξ is a point on the curve. For our ring in question we have

$$d\ell_\xi = Rd\varphi = r_0 \sin\theta_0 d\varphi$$

and then we would have

$$\int d\varphi\, [\text{ linear density }]\, \delta(x-\xi)\, r_0\sin\theta_0 d\varphi = [\text{ linear density }] \int d\varphi\, \delta(x-\xi)\, r_0\sin\theta_0 d\varphi$$

Now we can put in the usual spherical form for $\delta(x-\xi)$ to get

$$= [\text{ linear density }] \int d\varphi\, (1/r_0^2 \sin\theta_0)\, \delta(r-r_0)\delta(\theta-\theta_0)\delta(\varphi-\varphi')\, r_0\sin\theta_0 d\varphi$$

$$= [\text{ linear density }] \int d\varphi\, (1/r_0)\, \delta(r-r_0)\delta(\theta-\theta_0)\delta(\varphi-\varphi')\, d\varphi$$

$$= [\text{ linear density }]\, (1/r_0)\delta(r-r_0)\delta(\theta-\theta_0)$$

Now the total charge on our ring is $2\pi\, (r_0\sin\theta_0)\, [\text{ linear density }] = 1$ by intention, so we must then have

$$[\text{ linear density }] = 1/[2\pi\, (r_0\sin\theta_0)\,]$$

and our answer is then

$$\rho(x) = 1/[2\pi\, (r_0\sin\theta_0)\,]\, (1/r_0)\delta(r-r_0)\delta(\theta-\theta_0) = 1/[2\pi\, r_0^2 \sin\theta_0\,]\, \delta(r-r_0)\delta(\theta-\theta_0)$$

which agrees with what he gives!

5.13 Imagine on page 189 that f_1 and f_2 both vanish for $x < 0$. Then we have

$$\psi(y) = \int_{\text{full}} dz\, f_1(z)\varphi(y+z) = \int_{z>0} dz\, f_1(z)\varphi(y+z)$$

Consider some very large value of y like y = 1000. We know that $\varphi(1000+z)$ vanishes except near z = -1000, but this is not contained in the integral, so for very large y, $\psi(y) = 0$. So $\psi(y)$ is bounded on the positive side. Our second convolution equation is bottom of page 18,

$$\langle f_1 * f_2, \varphi\rangle = \int_{\text{full}} dy\, f_2(y)\psi(y) = \int_{y>0} dy\, f_2(y)\psi(y) = \int_{\text{full}} dy\, f_2(y)[\, \theta(y)\psi(y)\,]$$

so the test function here is really $\theta(y)\psi(y)$ which has finite support and is therefore a legal test function. that is to say, it vanishes for $y < 0$ and vanishes for $y >$ the bound we found above. Therefore, you can correctly talk about a convolution of two positive-only functions or distributions.

5.14. This is a change of variable situation where

$$u_1 = ax_1 + bx_2$$
$$u_2 = cx_1 + dx_2$$

If we write $dx_1\, dx_2 = J\, du_1 du_2$ then $J = \det(\partial x_i/\partial u_j)$. But easier to think the other way

$$du_1 du_2 = J'\, dx_1\, dx_2 \qquad \text{where } J' = \det(\partial u_i/\partial x_j) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad\text{-}bc$$

Then we have

$$\delta(u_1)\delta(u_2) = 1/|J'| * \delta(x_1)\delta(x_2) = |\, ad\text{-}bc\, |^{-1}\ \delta(x_1)\delta(x_2) \qquad \text{QED}$$

5.6 Fourier Transforms (23)

A. Classical Results

The **projection** is stated in 5.32, and I point out these facts:

- the function f(x) is defined only on a real argument, but may be complex
- notation for the FT is f^(u), a notation I have never seen before. I might use F(u).
- chooses a + sign in the expo, I usually choose a – sign
- there is no constant out front, just unity.

He claims without proof that the projection exists for any f(x) that is L_1. I guess this is pretty obvious if you just do a little inequality thing. Somehow I thought this was L_2 convergent, I was wrong.

The **inversion** he insists on writing as an R limit as in 5.34. On the top of page 24 he then shows that the recovery formula recovers the original f(x) provided f(x) is a C^1 function. In this proof, he relies on an exercise from Chap 1 which required this C^1 property of f. In this case, we could move the limit inside the integral and claim that $\lim_{R\to\infty} [\sin(Rz)/\pi z] = \delta(0)$, the bracketed object being a delta sequence. But then he comments that this C^1 property is not really necessary, all you need is piecewise C^0 and piecewise C^1 for f(x) with the proviso show in B.

Stak then quotes not one but four distinct **Parseval Formulas**. This reader usually thinks only of the last one which says that the L_2 norms of the function and the transform are (almost) equal. Had he chosen the constant differently in the projection (square root of 2π on the bottom I think), then the projection is scaled differently and we get that the L_2 norms are exactly the same. But convergence of the projection requires a finite L_1 norm.

Exercise: Show Parseval Formula 5.35:

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

So far, the transform variable has been called u and has been entirely real in the sense that you only need values of the transform for real values of u to recover f(x) as in 5.34. He now writes $\omega = u + iv$ and writes the projection as in 5.39. You can interpret this as just the projection of $[\, e^{-vx} f(x)\,]$ where v goes along for the ride as a fixed parameter. The recovery would then be p 25 A. Of course now you need to require that $[\, e^{-vx} f(x)\,]$ be an L_1 function and this will restrict the allowed range of parameter v. The recovery can be trivially rewritten as p 25 B where we now have a contour integral in the ω plane (just a horizontal line at height iv). Finally gets rid of the R limit and writes 5.40, saying this is a notational abuse, but we know that it implies the limit. He points out that the e^{-vx} factor can only help you on one side of the integral or the other, so not really that helpful.

Example 1: $f(x) = e^{-|x|}$. We compute the projection from p 24A naively, and it is clear either from the interpretation $[\, e^{-vx} f(x)\,]$ or from the convergence of the two integrals that we must have v in (-1,1) which means that $\hat{f}(\omega)$ is only given by the simple formula $2/(1+\omega^2)$ if ω lies in a horizontal strip in the ω plane centered at $\mathrm{Im}(\omega) = 0$ and with width 2. That is to say, $-1 < v < 1$. But I know that once you have this simply formula, you just analytically continue and apply the formula everywhere.

He then looks at the recovery integral 5.40 setting v = 0. This is the real axis contour integral in the page 26 figure. In the usual fashion, we add to this either the upper or lower great semicircle based on whether x is assumed positive or negative, and that semicircle contributes nothing to the integral. We then contract the contour around the single contained pole and this gives our recovery $f(x) = e^{-|x|}$.

Example 2: $f(x) = -2\sinh(x)H(x)$. This is an example I think of a "one-sided function". He computes the FT in the usual way and gets the *exact same result* as in the last example, $2/(1+\omega^2)$. BUT, this time we can see from p 27A that we must have $v > 1$ to get convergence of the projection integral. So this $\hat{f}(\omega)$ is only valid in a different "strip" of the ω plane compared to the previous example, and these strips do not overlap. So in example 2 the recovery formula 5.40 is valid for a horizontal contour which lies above the upper pole! Then when you do the contour close up or down trick, you pick up either no poles or two poles, and this recovers $f(x) = -2\sinh(x)H(x)$.

So this calls into question my throw-away comment above about analytic continuation? I don't think so. We can certainly define the function $g(\omega) = 2/(1+\omega^2)$ over the entire ω plane, it is meromorphic with two poles, fine. But our two examples have different recovery formulas, that is the point!

So this was a fascinating example. Two different f(x) functions have the SAME FT functional form which is $g(\omega) = 2/(1+\omega^2)$, but they are valid in non-overlapping strips of the ω plane. If the strips overlapped, the recoveries would have to be the same if you integrated in the overlap region, so that is why they cannot overlap.

Example 3: Here we have gaussian → gaussian under FT and the integrals converge for all ω, not just some finite strip in the ω plane.

Example 4: Here for f(x) we take a delta sequence he calls $f_\theta(x)$ which becomes $\delta(x)$ in the limit $\theta \to 0$. We know this fact from earlier work. He computes the FT to be a sinc function whose limit is 1. Thus, it appears that FT(δ) = 1. The recovery formula then shows that FT(1) = 2πδ. Note that it is very rare that you get a symmetric situation in the sense that FT(f) = g and FT(g) = f. I think a properly scaled gaussian would be the only function pair for which that symmetric relationship would be true.

C. One Sided Functions (28)

First, let's do the $f_+(x)$ world. We now require that (1) $f_+(x) = 0$ for $x<0$; (2) $f_+(x)$ must grow more slowly than $e^{\alpha x}$ for some α (α is real and could have either sign). Notice that if $f_+(x)$ grows as $x^{0.3}e^{\alpha x}$, it does not meet the requirement. He makes a clear definition here of something being "the order of" something.. Also, note that $f_+(x) = x^{213}$ *does* meet the requirement.

We are led at once to the transform pair 5.42 and 5.43 with the requirement now that v = Imω must be v > α. If this is true, then the projection integral converges everywhere above the line Imω = v, and cannot be infinite in this region. If a function g(ω) is finite and continuous in some region, it is analytic in that region. Our main concern here is that there are no poles above Im(ω) = v, so we can close our contour up and pick up no poles, should we want to do that. So he is right to say that f^(ω) is "analytic" above Imω = v. The recovery integral 5.43 has a horizontal contour which runs through this region of analyticity. I expect there to be a pole somewhere on the line Imω = v, so the recovery contour cannot be lowered further. He showed earlier why this "recovery" integral must have its contour in the region shown. So the main thing to notice in the pair A is that the projector starts at 0, and the **recovery** (or **expansion** or **inversion** I guess you might call it) integral must have v > α. In the recovery, if we close up, the upper great circle vanishes ($e^{-i\omega x} = e^{i[u+iv]x} = e^{iux}\, e^{-vx}$), the region is pole free, so we "recover" the fact that $f_+(x) = 0$ for $x<0$. [Notice that this upper/lower circle business depends on how you defined your FT in the first place. He has no usual footnote on this.]

I think the details for the $f_-(x)$ world are trivial, we just re-apply the above. This time $f_-(x) = 0$ on the right, and grows more slowly than $e^{\beta x}$ with real β. Our condition is now v < β because the danger part of the projection integral is at x = -∞. So this projection is analytic in a lower half plane below the line v = β. So transform pair B seems pretty clear as well.

Now go ahead and decompose an arbitrary f(x) as shown into $f_+(x) + f_-(x)$, but we want to maintain the restrictions on large |x| behavior as we had above. So $f(x) \sim e^{\alpha x}$ on the right, and $\sim e^{\beta x}$ on the left. Now we just add the equation pairs A and B to get equation set C. The only advantage I can see here is that you can use tuned horizontal contours in the two parts of the recovery formula. I think it is in fact typical for special functions to have asymmetric left and right behavior so you could in fact select a and b separately. He notes that if β > α, then there will exist a strip in which you can use the same contour for both integrals in 5.48.

So this is all good to know, but it certainly is not clear yet why such a decomposition is really helpful. I am sure we will find out soon.

D. Functions of Slow Growth. (29)

This means a function that at most grows as a power and is LI. A power-like function in general won't have a full FT (it will diverge of course at 0 or ±∞). However, the $f_+(x)$ component can be tamed by e^{-vx} for any v>0, and similarly $f_-(x)$ with v < 0. Here as usual v is Im(ω). Looking at set C p 28, we can create the two projections for appropriate v, then do the 5.48 recovery using appropriate contours in the two

integrals. This is shown in 5.49 where we set v = +ε in the first and v = -ε in the second. In 5.50 this is rewritten putting the ε inside the arguments as shown, so now have a real integral. If $f\hat{}_+(x)$ has a branch cut or some poles, you pass them on the indicated side. Notice the ε also in the two expos. The "hope" of Stak is that we can then define f^(u) as shown in 5.51 and we can think of that as the "regular" FT of the whole original f(x). Note that each integral in 5.49 has ε installed so it converges, so this fact must also be true in 5.50 but is very non-obvious. For example, the first term in 5.50 has $e^{+\varepsilon u}$ sitting there which appears to diverge as u→∞, but somehow this must be tamed by $f_+\hat{}(u)$.

I think this is the point here: if we can ignore the ε expos in 5.50, then the integrand in the bracket there is the thing shown on the left of 5.51. Then we "hope" to do something like the right equality in 5.51 to have a convergent projection. Ie, we sort of put in a temporary converger expo at both ends of the integral, compute the projection, then take ε → 0. I don't yet see any justification for this $e^{-\varepsilon|x|}$ factor in the preceding equations (other students must wonder as I do about this non-implication), but I agree that it might be a good thing to try to "get convergence" of the projection. All very hazy really.

How does this "example" help us? We have f(x) = 1. We compute the two projections shown in 5.46 for the + and - components of f(x), setting ω = u ± iε to get the ingredients of 5.51. Each of the integrals in p 29 A converges since ε has the right sense in each. We add the two pieces to get B. But we then recognize this thing as a δ sequence from somewhere in Chap 1 of Vol I (for example, p 43 A where k = 1/ε), so our ε→0 limit is then 2πδ(u) and this is indeed what we expect as a normal FT of f(x) = 1. In our current chapter, we already got this same result on p 27 C.

Note: for the first time ever, I think, on page 30 top Stak makes a reference to a page number! This has been a fault IMHO in the book to this point, because the reader has to struggle to find each reference. Sure, it requires some work on the part of the author, but all he needs do is keep a list and make sure it is correct as things shuffle with the print setup. Probably easier to do now than in the 1960's.

Main point of this section: When a FT does not converge in the classical sense, such as when f(x) = 1, it may still "converge" in the sense that it is equal to a distribution (symbolic function) like δ(x), and this is perfectly well defined in our distribution theory. You have some parameter like ε and you are taking a limit of something (a FT) which has as its limit a distribution.

E. Transforms of Distributions on the Line (ie, one variable). (30)

Motivation:

Our desire is to somehow associate f^(u) with a "distribution". We already know how to associate f(x) with a distribution on test function space K_1, as long as f(x) is LI. But getting a distribution associated with f^(u) is not so easy and requires more work and in fact requires a whole new test function space called S_1. Before talking about this new space, let's walk down the avenue of failed plans in search of a viable plan.

Plan A: If φ(x) = e^{iux} were a K_1 test function, we could write

$$T_f[e^{iux}] = (f, e^{iux})_{K1} = \int dx\, f(x)\, e^{iux} = f\hat{}(u)$$

$T_{\delta u}[\varphi] = (\delta_u,\varphi)_{K1} = \int dx\ \delta_u(x)\varphi(x) = \varphi(u)$ // just for comparison

Then we could say that we have associated the FT f^(u) of f(x) with a distribution where we have just taken a certain special test function.

The problem is this: $\varphi = e^{iux}$ is not a test function in $K_1 = D$. We can restrict e^{iux} to some finite interval (-A,A). That makes φ have finite support, but then φ won't be infinitely differentiable at the end points of the interval! That is why we always end up using that weird expo function for our prototype φ(x) with finite support.

[Notice in passing that $\varphi = e^{iux}$ is not a test function in space S_1 either (because it does not decay faster than any power, such as the power $x^0 = 1$. Ie, we don't have $|\varphi| \le 1$ for large x.)]

Plan B: Let's try this way to associate a distribution f^ with the FT of f(x):

$(f^, \varphi)_{K1} = \int du\ f^(u)\varphi(u) = \int dx\ f(x)\varphi^(x) = (f,\varphi^)_{K1}$ // using Parseval 5.35.

Now if we could arrange to have $\varphi^(u) = e^{iux}$, then we would have

$(f^, \varphi)_{K1} = \int du\ f^(u)\varphi(u) = \int dx\ f(x)\varphi^(x) = (f,\varphi^)_{K1} = \int dx\ f(x)\ e^{iux} = f^(u)$

Now from page 34 Example 3 we know that if $\varphi^(u) = e^{iux}$, then φ(x) = δ(x-u). But this φ is certainly not a test function in any space whatsoever! So things are looking grim for Plan B.

Plan C: So let's *forget about* trying to make e^{ikx} be a test function (or the FT of a test function). Instead, let's think more about this Parseval idea:

$(f^, \varphi)_{K1} = (f,\varphi^)_{K1}$??

Is this equality true for test functions φ in K_1? I think it is not true! Here is a counterexample. Suppose we take our famous $\varphi = \exp(1/[x^2-1])$ on (-1,1). Then we have

$\varphi^(u) = \int_{-1}^{1} dx\ \exp(1/[x^2-1])\ e^{iux} = \int_{-1}^{1} dx\ \exp(1/[x^2-1])\cos(ux)$

Maple gives us this integral as an infinite sum of ${}_0F_1$ functions which are really Bessel functions.

```
> int(exp(1/(x^2-1))*cos(u*x),x=-1..1);
```

$$\sqrt{\pi}\left(\sum_{_k1=0}^{\infty} \frac{(-1)^{-_k1}\,\Gamma(-_k1+1)\,\mathrm{hypergeom}\left([\,],\left[\frac{3}{2}-_k1\right],-\frac{1}{4}u^2\right)}{\Gamma(1+_k1)\,\Gamma\left(\frac{3}{2}-_k1\right)}\right)$$

$$J_n(x) = \frac{\left(\frac{1}{2}x\right)^n}{n!}\,{}_0F_1\left(;n+1;-\frac{1}{4}x^2\right)$$

The connection to the Bessel is n = 1/2-k and x = u so a term has $(u/2)^{-n}/n! * J_n(u)$. For any value of u, this is going to have some weird value, and there is no reason to expect φ^(u) to vanish outside some finite range!

Thus we have an example of a test function φ(x) in K_1 such that φ^(u) is NOT in K_1.

It is at this point that Stak defines the new space S_1 and, for test functions in that space, as he shows p 31 Theorem D, if φ is in S_1, then φ^(u) is also in S_1. So this is the big new difference from the K_1 world. So now we can write

$(f^\wedge, \varphi)_{S1} = (f, \varphi^\wedge)_{S1}$

where φ and φ^ are both S_1 test functions and this equality at least is consistent with itself!

About the S_1 test function space.

So, the idea is to invent a whole new space of test functions φ called S_1 which, although supported on the full infinite range, decay faster than any power! And the derivatives of φ must have this same property. This requirement can be stated as 5.52 where you can pick any finite integers k and ℓ you want. They note that $\varphi = e^{-x^2}$ on (-∞,∞) would be in S_1 but of course not in K_1 since support is infinite.

With this newly defined space of test functions, we can now look at *candidate* 31 A for $(f,\varphi)_{S1}$. Ie, this is a candidate for our definition of a distribution " f " in this new space. Recall that when we dealt with $(f,\varphi)_{K1}$, we only needed f to be LI, we never worried about convergence of the integral at ±∞ because the K_1 φ(x) always had finite support. Now in S_1 we don't have this φ finite support situation, our test functions are non-zero everywhere, and now we have to worry about convergence of the integral and we have to think about some restrictions on the growth of f(x) to stay convergent.

If f(x) is "of slow growth", it is bounded by some power x^k. But since any S_1 φ vanishes in the tails faster than any power, our integral converges. So now the integral form for $\langle f,\varphi\rangle_{S1}$ is a viable distribution.

Notice that we have not yet dealt with the issue of "what is an S_1 distribution related to a Fourier Transform". We are still in the preliminary stages of things.

On page 31 Stak wants to actually prove that such an $\langle f,\varphi\rangle_{S1}$ is in fact a "distribution". It is a functional, it is linear, and the problem is always to show continuity. If you install a null sequence of test functions φ_m as defined bottom of page 30, you want $\langle t,\varphi_m\rangle \rightarrow 0$, and this is what he proves page 31 middle. This new thing is called a "**distribution of slow growth**" which really means S_1 is the φ space.

Now suppose we have some φ(x) in S_1. The Theorem D p 31 claims that φ^(u) is also in S_1. He proves this down through 1/3 of page 32. This same argument applies in the reverse direction.

Properties of the new S_1 space distributions

At this point Stak provides a list of "properties" of our new S_1 space test functions, and this is where Fourier Transforms along with e^{iux} are going to appear (as noted above, e^{iux} is never thought of here as a test function).

(5.53): $\int dx\ \varphi^{(k)}(x)\, e^{iux} = (-iu)^k \int dx\ \varphi(x)\, e^{iux}$ = doing parts k times

The parts vanish because our S_1 test function φ (and its derivatives!) vanishes faster than any power at ±∞, and the power appearing here is $x^0 = 1$ in effect. So we never have to "fudge" the value of e^{iux} at infinity. We can by inspection rewrite the above as:

(5.53) $[\varphi^{(k)}]\hat{}(u) = (-iu)^k\, \varphi\hat{}(u)$

so we now know how to relate the FT of a derivative of any test function to the FT of the test function.

(5.54) $\partial_u{}^k \{ \int dx\ \varphi(x)\, e^{iux} \} = \int dx\ (ix)^k\, \varphi(x)\, e^{iux}$

This is NOT parts integration, we are just moving the derivative inside the integral and letting it act. Both integrals shown here converge because φ vanishes "faster than any power", so nothing is fuzzy and the change of order is justified. Again by inspection the above can be written as:

(5.54) $\partial_u{}^k [\varphi\hat{}(u)] = [(ix)^k \varphi(x)]\hat{}(u)$

Next we have

(5.55) $\int du\ \varphi\hat{}(u)\, e^{-iux} = 2\pi\, \varphi(x)$ // the "inversion formula" p 23 5.34

$\int du\ \varphi\hat{}(u)\, e^{+iux} = 2\pi\, \varphi(-x)$

which we can rewrite by inspection as

(5.55) $(\varphi\hat{})\hat{}(x) = 2\pi\, \varphi(-x)$

which is a little strange and new to me. The double FT does not take you back to where you started, but takes you to 2π times where you started but with negated argument. Seems pretty clear. Maybe the 2π would go away if things were normalized differently. I have never seen this result before.

(5.56) $\int dx\ \varphi(x-a)\, e^{iux} = \int dx'\ \varphi(x')\, e^{iu(a+x')} = e^{iua} \int dx'\ \varphi(x')\, e^{iux'}$

(5.56) $[\varphi(x-a)]\hat{}(u) = e^{iua}\, \varphi\hat{}(u)$

and this is a rule for the FT of a function of translated argument. So now I gather up all our little results:

(5.53) $[\varphi^{(k)}]\hat{}(u) = (-iu)^k\, \varphi\hat{}(u)$
(5.54) $\partial_u{}^k [\varphi\hat{}(u)] = [(ix)^k \varphi(x)]\hat{}(u) = [\varphi\hat{}(u)]^{(k)}$

(5.55) $(\varphi\hat{})\hat{}(x) = 2\pi\,\varphi(-x)$

(5.56) $[\varphi(x-a)]\hat{}(u) = e^{iua}\,\varphi\hat{}(u)$

These are "properties of test functions φ in the space S_1". Because φ are in this space S_1, the various integrals implied by these properties all converge!

The Fourier Transform of a distribution is also a distribution

Recall our results above

$(f, \varphi)_{S1} = \int dx\, f(x)\,\varphi(x)$ associates distribution f with function f(x)

$(f\hat{}, \varphi)_{S1} = (f, \varphi\hat{})_{S1}$ associates a distribution f^ with function f^(x)

or

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

Again, this is just one of the Parseval things 5.35. On page 33 upper half he writes this out using t in place of f. The thing on the left is "a functional or distribution associated with the function f^(x) ". He now wants to extend this idea to the world of symbolic functions (singular distributions), so we have

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

or

$$\int dx\, t\hat{}(x)\,\varphi(x) = \int dx\, t(x)\,\varphi\hat{}(x)$$

I am pretty happy with all this. Now comes a rather amazing claim. Above I list four properties of test functions φ in S_1. He claims these properties are true if you replace φ with an arbitrary t ! I need to do these one at a time.

A. Show that: $[t^{(k)}]\hat{} = (-iu)^k\, t\hat{}$ // like 5.53

Proof:

$([t^{(k)}]\hat{}, \varphi)_{S1} = ([t^{(k)}], \varphi\hat{})_{S1}$ // move ^ to the other side

$= (-1)^k\,(t, [\varphi\hat{}]^{(k)})$ // parts k times (moving derivative rule)

$= (-1)^k\,(t, [(ix)^k\varphi(x)]\hat{})$ // using 5.54

$= (t, [(-ix)^k\varphi(x)]\hat{})$ // move sign in

$= (t\hat{}, (-ix)^k\,\varphi(x))$ // move ^ to other side

$= ((-ix)^k\, t\hat{}, \varphi(x))$ // move function $(-ix)^k$ to other side

This we have shown that $([t^{(k)}]\hat{}, \varphi)_{S1} = ((-ix)^k\, t\hat{}, \varphi)_{S1}$ for any φ. Thus $[t^{(k)}]\hat{} = (-ix)^k\, t\hat{}$ QED.

So, in order to prove an "operator relation" like this (well, a distributional equality, each side of the equation is a distribution which we could write in older notation as " t ".), we have to show the operator thing is true for all test functions. If so, then the relation is true for the distributions in the abstract. The ingredients of this proof were (1) the four properties of the φ shown above; (2) the rule for moving ^ from

one side to the other; (3) the rule of old for moving a derivative from one side to the other (parts); (4) the rule of old for moving a regular function from one side to the other.

None of this would make sense in the K_1 test function space because most of the implied integrals would not converge even for regular distributions. That is the big deal here!

B. Show that: $[(ix)^k t(x)]\hat{}(u) = [t\hat{}(u)]^{(k)}$ // like 5.54

Proof:

$([(ix)^k t(x)]\hat{},\varphi) = ([(ix)^k t(x)],\varphi\hat{})$ // move ^ to other side

$= (t, (ix)^k \varphi\hat{})$ // move function to other side

$= (-1)^k (t, (-ix)^k \varphi\hat{})$ // fiddle signs

$= (-1)^k (t, [\varphi^{(k)}]\hat{})$ // from 5.53 for φ

$= (-1)^k (t\hat{}, [\varphi^{(k)}])$ // move ^ to other side

$= ([t\hat{}]^{(k)}, \varphi)$ // derivatives move rule k times

Thus we conclude that $[(ix)^k t(x)]\hat{} = [t\hat{}]^{(k)}(x)$ QED. Again, one key step is that we can apply our various φ properties like 5.53 (in the last proof it was 5.54).

C. Show that: $(t\hat{})\hat{}(x) = 2\pi\, t(-x)$

Proof: $((t\hat{})\hat{}(x), \varphi) = (t\hat{}, \varphi\hat{}) = (t, \varphi\hat{}\hat{}) = (t, 2\pi\, \varphi(-x))$ // using 5.55

$= \int dx\, t(x)\, 2\pi\varphi(-x) = \int dx'\, t(-x')\, 2\pi\varphi(x') = 2\pi\, (t(-x), \varphi)$ QED

D. Show that: $[t(x-a)]\hat{}(u) = e^{iua}\, t\hat{}(u)$

Proof: $([t(x-a)]\hat{}(u), \varphi(u)) = ([t(u-a)], \varphi\hat{}(u)) = \int du\, t(u-a)\, \varphi\hat{}(u) = \int dx\, t(x)\, \varphi\hat{}(x+a)$ // x = u-a

But we know that

$$\varphi\hat{}(u) = \int dy\, \varphi(y)\, e^{iuy}$$

$$\varphi\hat{}(x+a) = \int dy\, \varphi(y)\, e^{iy(x+a)}$$

so the above becomes

$$= \int dx\, t(x) \int dy\, \varphi(y)\, e^{iy(x+a)} = \int dy\, \varphi(y)\, e^{iya} \int dx\, t(x)\, e^{iyx}$$

$$= \int dy\, \varphi(y)\, e^{iya}\, t\hat{}(y) = \int dy\, e^{iya}\, t\hat{}(y)\, \varphi(y) = (e^{iya}\, t\hat{}(y), \varphi(y)) = (e^{iua}\, t\hat{}(u), \varphi(u))$$

and therefore we have

$[t(x-a)]\hat{}(u) = e^{iua}\, t\hat{}(u)$ // which is 5.61

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

Proof: $([t(-x)]\hat{}(u), \varphi(u)) = (t(-u), \varphi\hat{}(u)) = (t(u), \varphi\hat{}(-u)) = (t\hat{}(u), \varphi(-u)) = (t\hat{}(-u), \varphi(u))$

Therefore $[t(-x)]\hat{}(u) = t\hat{}(-u)$ QED for 5.62.

Properties of Distributions

So lets gather up all these "properties of distributions":

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

Examples of Fourier Transforms of distributions: (34)

Again, in order to find equalities with distributions, you need to bracket the "operator" with φ.

1. Here we go:

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

and therefore we get this "operator equation": $\delta\hat{} = 1$

2. Here we go with another one: (where we use rule 5.55 above)

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

and therefore we get this "distributional equality" that: $1\hat{} = 2\pi\delta$.

Applying the previous example, then we have also that

$1\hat{}\hat{} = 2\pi\delta\hat{} = 2\pi\, 1$ and $\delta\hat{}\hat{} = 1\hat{} = 2\pi\delta$

3. From 5.56 above applied to $\varphi = \delta$ we have : $[\delta(x-a)]\hat{}(u) = e^{iua}\, \delta\hat{}(u) = e^{iua}\, 1$

So at this point, the reader gets the idea that doing a Fourier Transform of a distribution is quite easy, and everything is completely well defined. That is the whole point!

4. The big question here is this: what is the Fourier Transform of the Heaviside function H(x)? We need to think of H as a distribution, because H^ turns out to be a singular distribution. He pursues this three different ways.

(a) Doing simple manipulations and calling upon two limits from Chapter 1, Stak shows result p 35 A, and from this we conclude that the correct distributional or operator result is this:

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

which is a very strange looking result! You have to show arguments here due to the argument of the pf function. As a reminder, the pf(1/x) distribution is defined like this:

$$\langle pf(1/x), \varphi \rangle = \fint dx\,(1/x)\varphi(x)$$

Recall from Chapter 1 that $1/(x \pm i\varepsilon) \rightarrow \mp\, i\pi\delta(x) + pf(1/x)$ so that that above result is this:

$$-i\hat{H}(y) = 1/(y+i\varepsilon)$$

where however we are now mixing K_1 and S_1 results in this last step which may not be valid (but we see later it is in fact valid).

(a') In a naive function sense, suppose we just try to compute H^(u) as a function. We get

$$\hat{H}(u) = \int dx\, e^{iux} H(x) = \int_0^\infty dx\, e^{iux}$$

We could evaluate

$$\hat{H}(1) = \int dx\, e^{ix} H(x) = \int_0^\infty dx\, e^{ix} = (1/i)[\, e^{ix}\,|_0^\infty = e^{i\infty} - e^0/i \approx i$$

According to our fancy result above, we would say that $\hat{H}(1) = i\,pf(1) = i\,(1/1) = i$. So at points away from the origin, our naive method agrees with the distribution thing as long as we set $e^{i\infty} = 0$ in our naive method, which is an ugly thing to do! The same confusion arises if you do parts. So the bottom line is that this only makes sense at all if you do it in the Stakgold distribution sense. Stak puts in a little regular in part (c) below and that justifies $e^{i\infty} = 0$.

(b) Here we start over with Heaviside and observe that H'(x) = δ(x) and we do some easy fiddling using previously derived simple results to get

$1 = -ix\,\hat{H}(x)$ which suggests that $-i\hat{H}(x) \approx 1/x$ as shown both above and below

Stak notes that the above is a "distributional equation". He then claims that this equation has the following particular solution:

$H\hat{}(x) = i\ pf(1/x)$

The proof is simple: $(-ix\ H\hat{}(x),\varphi) = (-ix\ i\ pf(1/x),\varphi) = (x\ pf(1/x),\varphi) = \fint dx\ x(1/x)\varphi(x)$

$= \fint dx\ \varphi(x) = \int dx\varphi(x) = (1,\varphi)$

We can see that adding $C\delta(x)$ to our solution does not change it, since $x\delta(x) = 0$. He then does a little trick to show that $C = 1$. Somehow, the general constant C does not do the right thing.

(c) The idea here is to put in a QED like "regulator" factor with an ε, and then take a limit. This gives our little result

$-i\ [e^{-\varepsilon x}H]\hat{}(x) = 1/(u+i\varepsilon)$ // which is our result noted above as $\varepsilon \to 0$.

He then uses this $1/(u+i\varepsilon)$ result to rederive our original result again using some limits which he says are "easily calculated". So I think I have had enough of this $H\hat{}(x)$ example! I maybe ought to be collecting all these interesting little limit things.

F. Transforms in more than one variable (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 $i\mathbf{x}\bullet\mathbf{u}$ 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.

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

Exercises (37)

5.15 When $e^{-\varepsilon|x|}$ is added as shown with $\varepsilon>0$, we just enhance the convergence of f(x) with this multiplicative factor. That is to say, $g(x) = e^{-\varepsilon|x|}f(x)$ vanishes faster than a fixed power. So nothing really happens here. But if you looked at $\varepsilon\to 0-$, you would have a problem because then g(x) does not meet the requirements. So this thing is only true for $\varepsilon\to 0$ from the positive side.

5.16. Show that $t_n\hat{} \to t\hat{}$ if $t_n \to t$. Another result like the derivative. I think this would be easy to show.

5.17 Now we throw in that same $e^{-\varepsilon|x|}$ and split into the + and - parts. We can now regard ε as a sequence label and use the previous exercise to prove result 5.51 which we previously only conjectured to be true.

5.18. I would show this result by first invoking 5.53 applied to f instead of φ and then I recognize the thing on the right as f^(u) from Exercise 5.17 right above. He has a typo "n" which should be a p. Our 5.53 rule applies for any show growth world "t" as we showed above, item A in the list.

5.19. Here we realize that the traditional **Laplace Transform** is just our f_+ thing, that is to say,

$$f_+\hat{}(\omega) = \int_0^\infty dx\ f_+(x)e^{i\omega x} \qquad 5.42$$

The Laplace transform applies to a "right sided function" f(x) only, so identify $f(x) = f_+(x)$. So we have

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

Again: we realized that the Laplace Transform $f\tilde{}(s) = f_+\hat{}(is)$.

Now look at 5.43 page 28. The recovery must have $\nu > \alpha$ where α is how fast $f_+(x)$ grows. Notice that on the contour, the expo of the integrand is $\exp(-i[i\nu+u]x) = \exp(+\nu x - iux)$ with u = real. The number ν is fixed, so this $e^{+\nu x}$ should not worry us. Similarly, in the Laplace recovery we have $e^{sx} \sim e^{as}$ with a>0, but again this is a constant a. I am just trying to explain a confusion I had at first here. You don't have an expo inside the integral with a positive exponent with variable running off to infinity.

5.20. We think we know now how to compute the FT of any "distribution". So here are three functions to test this belief on. The answers are not obvious to me right now. But let's do them here:

5.20 A Compute the FT of pf(1/x).

I am having a lot of trouble doing this. Here is one of my methods:

$$\langle pf(1/x)\hat{}\ ,\ \varphi\rangle = \langle pf(1/x),\ \varphi\hat{}(x)\rangle = \fint dx\ (1/x)\ \varphi\hat{}(x) = \fint dx\ (1/x)\ \int dy\ e^{ixy}\varphi(y)$$

$$= \int dy\ \varphi(y) \fint dx\ (e^{ixy}/x)$$

So basically, the answer is that

$$\text{pf}(1/x)\hat{} \;=\; \fint dx\, (e^{\mathbf{ixy}}/x)$$

which I could have gotten just by writing it out. But with φ in there, things converge. But in this last equality, both sides are symbolic functions. So, what function is this exactly? In doc "expo over power..." I show this fact:

$$Il = \int_a^\infty dx \frac{e^{ikx}}{x} = \int_l^\infty dy \frac{e^{ikay}}{y} \quad = -[\, ci(ka) + i\, si(ka)\,]$$

which I will rewrite here as

$$\int_{\mathbf{a}}^{\infty} dx\, (e^{\mathbf{iyx}}/x) \;=\; -\,[ci(ay) + i\, si(ay)]$$

where the ci and si are the sine and cosine integral functions of GR page 928. We can probably simplify a bit by claiming that our tick integral is still symmetric, so only the even part of $(e^{\mathbf{ixy}}/x)$ can contribute, so we then claim that

$$\text{pf}(1/x)\hat{} \;=\; \fint dx\, i \sin(yx)/x \;=\; \left(\int_{-\infty}^{\mathbf{-\varepsilon}} + \int_{\boldsymbol{\varepsilon}}^{\infty} \right) dx\, i \sin(yx)/x \;=\; 2i \int_{\boldsymbol{\varepsilon}}^{\infty} dx \sin(yx)/x$$

Then we would use GR page 928 (8.231) so make this claim

$$\text{pf}(1/x)\hat{} \;=\; 2i \int_{\boldsymbol{\varepsilon}}^{\infty} dx \sin(yx)/x \;= -\,2i\, si(\varepsilon y)$$

It seems pretty clear that si(-u) = -si(u) just as for the sin function itself. But cutting to the chase:

$$\int_{\mathbf{0}}^{\infty} dx \sin(yx)/x \;=\; \text{sign}(y)\, \pi/2 \qquad \text{Schaum p 96} \qquad y \neq 0$$

so then we get

$$[\text{pf}(1/x)\hat{}](y) \;=\; 2i \int_{\mathbf{0}}^{\infty} dx \sin(yx)/x \;=\; 2i * \text{sign}(y)\, \pi/2 \;=\; +\, i\pi\, \text{sign}(y)$$

Is this right? I did find something in a Google book searching on "Fourier transforms of distributions" :

http://books.google.com/books?id=sp08AAAAIAAJ&pg=PA96&lpg=PA96&dq=%22fourier+transforms+of+distributions%22&source=bl&ots=zYvCS8SCEm&sig=VCKd_uHa4LoIQ9NaVv2dnBDNcko&hl=en&ei=Zng6Ss7hOY3QsQPW0Yz-Bg&sa=X&oi=book_result&ct=result&resnum=4

EXAMPLE 1.

$$\mathcal{F}\left[\text{v.p.}\,\frac{1}{x}\right]=\begin{cases}-\pi i & (\xi>0)\\ \pi i & (\xi<0),\end{cases} \qquad (2.76)$$

EXAMPLE 2.

$$(\log|x|)'=\text{v.p.}\,\frac{1}{x} \quad (n=1).$$

$$\hat{f}(\xi)=\int_{-\infty}^{+\infty} e^{-2\pi i x\xi} f(x)\,dx, \qquad (1.27)$$

The second shows that this author's v.p function is the same as Stakgold's pf function. He gets a different sign in his answer because he defines the FT with a negative exponential. So, I think I got the right answer!

I see that some people refer to pf(1/x) as PV(1/x) where PV means Principle Value, which is the tick integral idea. So maybe that is what "v.p." means in the above author's work.

5.20 B Compute the FT of sgn(x).

$\langle[\text{sgn}(x)]\hat{}(u), \varphi(u)\rangle = \langle \text{sgn}(x), \hat{\varphi}(x)\rangle = \int dx\ \text{sgn}(x)\ \hat{\varphi}(x)$

$= \int dx\ \text{sgn}(x) \int dy\ e^{iyx}\varphi(x) = \int dx\ \varphi(x)\ \{\ \text{sgn}(x) \int dy\ e^{iyx}\ \}$

$= \int dx\ \varphi(x)\ \{\ \text{sgn}(x)\ 2\pi\delta(x)\ \}$

But this is a bit problematical since sgn(0) is undefined. Let's try another method

$\text{sgn}(x) = -1 + 2H(x)$

We know from page 24 that $H\hat{}(x) = i\ \text{pf}(1/x) + \pi\ \delta(x)$.
We know that $1\hat{}(x) = 2\pi\delta(x)$. Therefore

$\text{sgn}\hat{}(x) = -\ 2\pi\delta(x) + 2\ \{\ i\ \text{pf}(1/x) + \pi\ \delta(x)\} = 2i\ \text{pf}(1/x)$

If I just try to compute this naively, get (keep odd part of expo)

$\text{sgn}\hat{}(x) = \int dy\ e^{ixy}\ \text{sgn}(y) = 2i \int_0^{\infty} dy\ \sin(xy) = -\ 2i\ \cos(xy)/x\ |^{y=\infty}_0 = +2i/x$ in some sense

so this naive result at least bears a resemblance to the pf result shown above. Here is some confirmation of my result above:

EXAMPLE 9.6
Spectrum of sgn t

The distribution sgn t is odd (see exercise 8.27a). Applying (9.13) to the result of example 9.4, we obtain that the spectrum of $\pi i \operatorname{sgn} t$ is the distribution $2\pi \operatorname{pv}(1/\omega)$, that is,

$$\operatorname{sgn} t \leftrightarrow -2i\operatorname{pv}(1/\omega). \tag{9.14}$$

Spectrum of $\epsilon(t)$

Since $\epsilon(t) = (1 + \operatorname{sgn} t)/2$, where $\epsilon(t)$ is the unit step function, it follows that

$$\epsilon(t) \leftrightarrow \pi\delta(\omega) - i\operatorname{pv}(1/\omega), \tag{9.15}$$

where we again have the same sign issue due to the definition of the FT. This reference is

http://books.google.com/books?id=frT5_rfyO4IC&pg=PA209&lpg=PA209&dq=%22fourier+transforms+of+distributions%22&source=bl&ots=hlTQg_bBpy&sig=CA0clnYvRR-Nz4wlBnQcVX3EBnA&hl=en&ei=nIU6SuinNI_KsQOFkpH-Bg&sa=X&oi=book_result&ct=result&resnum=7

5.20 C Compute the FT of log |x|

Rule A above says that : $[t^{(k)}]\hat{} = (-iu)^k\, t\hat{}$ or, with k = 1,

$[d/dx\ t(x)]\hat{}(u) = (-iu)\ [t(x)]\hat{}(u)$

If we choose t(x) = log|x| we then have

$[d/dx\ \log|x|]\hat{}(u) = (-iu)\ [\log|x|]\hat{}(u)$

But $[d/dx\ \log|x|]\hat{}(u)$ = $[pf(1/x)]\hat{}(u)$ so we have

$[pf(1/x)]\hat{}(u) = (-iu)\ [\log|x|]\hat{}(u)$

But from our first exercise we know that

$[pf(1/x)]\hat{}(u) = +\ i\pi\ \operatorname{sign}(u)$

So that

$i\pi\ \operatorname{sign}(u)\ =\ (-iu)\ [\log|x|]\hat{}(u)$

and

$[\log|x|]\hat{}(u) = i\pi\ \operatorname{sign}(u)/\ (-iu)\ = -\ \pi\ \operatorname{sign}(u)/u$ // I have no web confirmation.

5.21 Compute the FT of x^k

Quote first from our list or properties above,

$[(ix)^k t(x)]\hat{}(u)\ =\ [t\hat{}(u)]^{(k)}$

Then set t = 1 so we have

$$i^k [\, x^k]\hat{}(u) = [1\hat{}(u)]^{(k)} = [2\pi\delta(u)]^{(k)} = 2\pi\, \delta^{(k)}(u)$$

Therefore

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

5.22. In this example, our distribution t(x) is in n dimensions and is supposed to be an arbitrary simple layer distribution described by source density a(x) on a surface σ of n-1 dimensions. The various steps shown are all checked off and I agree with the general result 5.68 where t^ is given as a surface integral over σ of the density times the multidimensional expo.

Now we are supposed to consider a sphere in an n-dimensional space. Such a sphere we know has an n-1 dimensional surface. A uniform surface charge on this hypersphere is given by a = δ(r-R) where r is the radial coordinate in n-dimensional spherical coordinates.

This problem requires information beyond the scope of this book. We can crib the notion of spherical coordinates from wiki as follows:

****

Large digression on Spherical Coordinates in n dimensions

In n dimensions, we get one radial coordinate r and n-1 angle coordinates $\varphi_1 \ldots \varphi_{n-1}$ as follows:

$$\begin{aligned}
x_1 &= r\cos(\phi_1) \\
x_2 &= r\sin(\phi_1)\cos(\phi_2) \\
x_3 &= r\sin(\phi_1)\sin(\phi_2)\cos(\phi_3) \\
&\vdots \\
x_{n-1} &= r\sin(\phi_1)\cdots\sin(\phi_{n-2})\cos(\phi_{n-1}) \\
x_n &= r\sin(\phi_1)\cdots\sin(\phi_{n-2})\sin(\phi_{n-1})
\end{aligned}$$

The first angle φ_1 is the "polar angle", while the very last one is the "azimuth". Only this last angle ranges 0 to 2π, all the others go 0 to π.

For n = 2 we have only $\varphi_1 = \theta$, angle on a circle
For n = 3 we have $\varphi_1 = \theta$ and $\varphi_2 = \varphi$. The inverses are obtained as follows:

$$\tan(\phi_{n-1}) = \frac{x_n}{x_{n-1}}$$
$$\tan(\phi_{n-2}) = \frac{\sqrt{{x_n}^2 + {x_{n-1}}^2}}{x_{n-2}}$$
$$\vdots$$
$$\tan(\phi_1) = \frac{\sqrt{{x_n}^2 + {x_{n-1}}^2 + \cdots + {x_2}^2}}{x_1}$$

Then the volume element requires the Jacobian which I could compute from the first set above, but here is the answer:

$$d_{\mathbb{R}^n}V = \left|\det \frac{\partial(x_i)}{\partial(r,\phi_j)}\right| dr\, d\phi_1\, d\phi_2 \ldots d\phi_{n-1}$$
$$= r^{n-1}\sin^{n-2}(\phi_1)\sin^{n-3}(\phi_2)\cdots\sin(\phi_{n-2})\, dr\, d\phi_1\, d\phi_2 \cdots d\phi_{n-1}$$

Now, if we were to integrate over all angles except the polar angle φ_1 , we would get

$$dS_n = r^{n-1}\sin^{n-2}(\varphi_1)\, d\varphi_1\, C_n$$

where C_n is some constant. We could integrate this to get the area of our hypersphere:

$$S_n = r^{n-1}C_n \int_0^{\pi} d\varphi_1 \sin^{n-2}(\varphi_1) = r^{n-1}C_n\, 2 \int_0^{\pi/2} dx \sin^{n-2}(x)$$

This is a strange integral, and we consult GR page 369 which says

$$\int_0^{\pi/2} dx \sin^{n-2}(x) = (n-3)!!\,/\,(n-2)!!\ (\pi/2) \qquad \text{if n is even}$$

$$\int_0^{\pi/2} dx \sin^{n-2}(x) = (n-3)!!\,/\,(n-2)!! \qquad \text{if n is odd}$$

Example inserted. Suppose n = 3. Then $dS_3 = r^2\sin\varphi_1 d\varphi_1 C_3$ and we know from experience with this case that $C_3 = 2\pi$. The area of our 3D sphere would then be $S_3 = r^2 * 2\pi * 2 * \int_0^{\pi/2} dx \sin(x)$. This integral is

$$\int_0^{\pi/2} dx \sin(x) = -\cos(x)\big|^{\pi/2}_0 = -[\,0 - 1\,] = 1$$

so we would conclude that $S_3 = r^2 * 2\pi * 2 * 1 = 4\pi r^2$ which is in fact the area of our sphere.

But GR page 369 tells us that (where we use $\mu = n-1$ so that $n = \mu+1$)

$$\int_0^{\pi/2} dx \sin^{\mu-1}(x) = 2^{\mu-2}\ B(\mu/2,\mu/2) = 2^{\mu-2}\ [\Gamma(\mu/2)]^2 / \Gamma(\mu)$$

$$\int_0^{\pi/2} dx \sin^{n-2}(x) = 2^{n-3}\ B(n/2 - 1/2, n/2 - 1/2) = 2^{n-3}\ [\Gamma(n/2 - 1/2)]^2 / \Gamma(n-1)$$

As a check, if n=3 we get $2^0\ \Gamma(1)^2/\Gamma(2) = 1 * 1^2 / 1 = 1$. Then we have that

$$S_n = r^{n-1}\ C_n\ 2\ \ 2^{n-3}\ \{\Gamma[n/2 - 1/2])\}^2 / \Gamma[n-1]$$

But from page 15 equation 5.18 we know that

$$S_n = n\pi^{n/2}\ r^{n-1} / (n/2)! = n\pi^{n/2}\ r^{n-1} / \Gamma[\ n/2 + 1\ ]$$

As a check, $S_3 = 3*\pi^{3/2}\ r^2/\ \Gamma(5/2) = 3*\pi^{3/2}\ r^2/\ [\ 3\ \pi^{1/2}/4] = \pi^2\ r^2\ 4 = 4\pi r^2$.

The comparison is pretty ugly but tells us that

$$r^{n-1}\ C_n\ 2\ \ 2^{n-3}\ \{\Gamma[n/2 - 1/2]\}^2 / \Gamma[n-1] = n\pi^{n/2}\ r^{n-1} / \Gamma[\ n/2 + 1\ ]$$

$$C_n\ \ 2^{n-2}\ \{\Gamma[n/2 - 1/2]\}^2 / \Gamma[n-1] = n\pi^{n/2} / \Gamma[\ n/2 + 1\ ]$$

$$C_n = n\pi^{n/2}\ 2^{2-n}\ \Gamma[n-1] / \{\ \Gamma[n/2 - 1/2]^2\ \Gamma[\ n/2 + 1\ ]\ \}$$

where we have a messy combination of four gamma functions on the right. But now we have lots of ugly algebra to try to simplify this. Here we go with a complete track record:

$\Gamma[\ n/2 + 1\ ] = (n/2)\ \Gamma[\ n/2]$ Schaum p 101 16.2

so

$$C_n = n\pi^{n/2}\ 2^{2-n}\ \Gamma[n-1] / \{\ \Gamma[n/2 - 1/2]^2\ (n/2)\ \Gamma[\ n/2]\ \}$$

$$= (2/n)\ n\pi^{n/2}\ 2^{2-n}\ \Gamma[n-1] / \{\ \Gamma[n/2 - 1/2]^2\ \Gamma[\ n/2]\ \}$$

$$= \pi^{n/2}\ 2^{3-n}\ \Gamma[n-1] / \{\ \Gamma[n/2 - 1/2]^2\ \Gamma[\ n/2]\ \}$$

The next step is to write , where n/2 - 1/2 = y, 2y = n-1

$\Gamma[n/2 - 1/2]\ \Gamma[\ n/2] = \Gamma[y]\ \Gamma[y+1/2] = \pi^{1/2}\Gamma(2y)\ 2^{1-2y}$ Schaum p 102

$= \pi^{1/2}\ \Gamma(n-1)\ 2^{1-(n-1)} = \pi^{1/2}\ \Gamma(n-1)\ 2^{-n+2}$

Then we have

$$C_n = \pi^{n/2}\ 2^{3-n}\ \Gamma[n-1] / \{\ \Gamma[n/2 - 1/2]\ \pi^{1/2}\ \Gamma(n-1)\ 2^{-n+2}\ \}$$

$= \pi^{n/2}\, 2 \;/\; \{\, \Gamma[n/2 - 1/2]\, \pi^{1/2} \,\}$

$= \pi^{(n-1)/2}\, 2 \;/\; \Gamma[n/2 - 1/2]$

As a check, $C_3 = \pi\, 2 / \Gamma(1) = 2\pi$ which is correct.

Therefore, we have this as our surface area ring

$$dS_n = r^{n-1} \sin^{n-2}(\varphi_1)\, d\varphi_1\, C_n \qquad \text{where } C_n = 2\pi^{(n-1)/2} / \Gamma[n/2 - 1/2]$$

Let's check this result for the case n = 3. Then

$$dS_3 = r^2 \sin(\varphi_1)\, d\varphi_1\, C_3 \qquad \text{where } C_3 = 2\pi / \Gamma[3/2 - 1/2] = 2\pi$$

$$\Rightarrow \quad dS_3 = 2\pi\, r^2 \sin\theta\, d\theta \qquad d\Omega_3 = 2\pi \sin\theta\, d\theta$$

which is correct

Let's ignore this error for the moment and consider 5.69 where he makes this claim:

$$dV = \sin^{n-2}(\theta)\, d\theta\, r^{n-1}\, S_{n-1}(1)$$

and I compare this with my result above

$$dS_n = r^{n-1} \sin^{n-2}(\varphi_1)\, d\varphi_1\, C_n = \sin^{n-2}\theta d\theta\, r^{n-1}\, C_n$$

so his annoyingly simple result is just this:

$$C_n = S_{n-1}(1) = (n-1)\, \pi^{(n-1)/2} / [\, n/2 - 1/2]\,! = (n-1)\, \pi^{(n-1)/2} / \{(n/2 - 1/2)\, [\, n/2 - 3/2]\,! \,\}$$

$$= (n-1)\, \pi^{(n-1)/2} / \{(n - 1)(1/2)\, [\, n/2 - 3/2]\,! \,\} = 2\, \pi^{(n-1)/2} / [\, n/2 - 3/2]\,!$$

whereas I got

$$C_n = 2\pi^{(n-1)/2} / \Gamma[n/2 - 1/2] = 2\pi^{(n-1)/2} / [n/2 - 3/2]\,!$$

so again we agree.

Comment: I started off above saying "Now, if we were to integrate over all angles except the polar angle φ_1, we would get $dS_n = r^{n-1} \sin^{n-2}(\varphi_1)\, d\varphi_1\, C_n$ where C_n is some constant." Yes, I can see in the case of n = 3 that this area is a ring on a sphere of radius $r\sin\varphi_1$ and width $rd\varphi_1$. The area of this ring thing is then $2\pi\, r\sin\varphi_1$ times $rd\varphi_1$ and yes, 2π is the total angle integral for the n=2 problem, $S_2(1)$. So I admit that it is correct that $dS_3 = r^2 \sin(\varphi_1)\, d\varphi_1\, S_2(1)$ and the general case must be this:

$$dS_n = r^{n-1} \sin^{n-2}(\varphi_1)\, d\varphi_1\, S_{n-1}$$

****

Resume Exercise 5.22

So, I now (finally) agree with Stak's equation 5.69 on page 39, and I can now resume this exercise. The exponent in 5.68 contains dot product of $\boldsymbol{\xi} \bullet \mathbf{x}$ where each are vectors in R_n. Regardless of n, I guess we can go into the above spherical coordinates and align $\boldsymbol{\xi}$ with our "z axis" so $\boldsymbol{\xi} \bullet \mathbf{x} = \xi r \cos(\varphi_1)$. This is a little wobbly for me, but I think it is right, and it is what he shows in 5.69. Honesty is my policy.

Now remember what we are doing here: $t(x) = 1\ \delta(r-R)$ is our "simple layer" on a sphere, where r is the radial coordinate of the vector x. We are trying to compute the FT $t\hat{}(x)$. Our result 5.69 shows that this thing is just a number which is a function of $|\xi|$ which is the radial component of ξ. I would be inclined now to write 5.69 as follows:

$$t\hat{}(x) = R^{n-1} S_{n-1}(1) \int_0^{\pi} d\theta'\ e^{iRr\cos\theta'} \sin^{n-2}(\theta')$$

where "r" is the radius of the n-dimensional vector **x** in R_n and θ' is a dummy integration variable. So, we start with a spherically symmetric distribution $t(x) = t(r) = \delta(r-R)$, and we find that the Fourier transform of this is also spherically symmetric and can be written $t\hat{}(x) = F(r)$ as shown above.

All that remains is to set n = 2 and n = 3 and validate his last two results.

n = 2: $t\hat{}(x) = R^1 S_1(1) \int_0^{\pi} d\theta'\ e^{iRr\cos\theta'}$

Use GR page 482 5 with $\nu = 0$ and $\beta = Rr$ to get

$$\pi^{1/2} * 1 * \Gamma(1/2)\ J_0(Rr) = \pi^{1/2} * 1 * \pi^{1/2}\ J_0(Rr) = \pi\ J_0(Rr)$$

so then

$$t\hat{}(x) = R\ S_1(1)\ \pi\ J_0(Rr)$$

But $S_1(1)$ is the volume of a 1D sphere, probably just 1.

$$S_1(1) = 1\ \pi^{1/2} / (1/2)! = \pi^{1/2} / \Gamma(3/2) = \pi^{1/2} / [\pi^{1/2}/2] = 2$$

Maybe you have a 1D sphere be a line segment and the two ends of the segment represent the "area" of this 1D sphere and that is why you get 2. So the final answer is

$t\hat{}(x) = 2\pi R\ J_0(Rr)$ n = 2 agrees with 5.70

n = 3: $t\hat{}(x) = R^2 S_2(1) \int_0^{\pi} d\theta'\ e^{iRr\cos\theta'} \sin^1(\theta')$ $S_2(1) = 2\pi 1$

We can use the same GR integral p 482 5 with $\nu = 1/2$ this time, so

integral = $\pi^{1/2}(2/Rr)^{1/2}\ \Gamma(1)\ J_{1/2}(rR)$

Schaum p 138 shows that $J_{1/2}(rR) = (2/\pi rR)^{1/2} \sin(rR)$ so then our total answer is:

$R^2 * 2\pi * \pi^{1/2}(2/Rr)^{1/2} * (2/\pi rR)^{1/2} \sin(rR)$

$= R * 2\pi * (2/r)\ \sin(rR)\ = 4\pi\ (R/r) \sin(rR)$ // agrees with 5.71

Note that "r" is conjugate to a spatial "r" and he calls it ξ. The combination rR is dimensionless.

Comments: There are various books on the Fourier Transforms of Distributions. Here is one:



where we recognize the name of the author as a Bateman Boy.

5.7 PDE's for Distributions (39)

This section opens with a very fancy claim. I have pondered it for a few hours, and I am at a complete loss as to how I would show it was true. I have hunted on the web, but Green's theorems get into differential forms and exotica I know nothing about, as in my Buck calculus book. I can't even get a search handle on the theorem that is claimed here. I don't know what to search for! [But now I have two docs on this subject and I have verified the claim, see "parts and greens.doc".]

Claimed Theorem 5.72: For L a general differential operator, there exists a "current" **J** with this property:

$vLu - uL^*v = \nabla \bullet \mathbf{J}$

where current **J** is an n-vector which is bilinear in u and v (with various derivatives). L* is of course the formal adjoint of L. Why on earth should this difference on the left be the divergence of some vector function? (that was my big mystery)

Exercise: figure this out! I think you just do multiple parts with a multi-index $a_k(x)D^k_x$ sample term to prove that it is true.

Given this "Green's theorem", then 5.73 follows at once from the n-dimensional "divergence theorem". This is one of the obscurest things Stak has rolled out so far! I was able to follow this pretty easily in Vol I where we had just a single variable, because we could do "parts", but I don't really know how to "do parts" in a more general sense. My only weapons are the short list of "integral theorems" on my little sheet. In the special case of the Laplace operator $L = \nabla^2$, yes I am able to show that there exists a current **J**, because this follows from what I call Green #2. I first show that

RHS Green #2 $= \int dS\, \hat{n} \bullet \{ \psi \nabla \varphi - \varphi \nabla \psi \}$ then $= \int dV\, \nabla \bullet \{ \psi \nabla \varphi - \varphi \nabla \psi \}$

so in this tiny case, we have $\mathbf{J} = \psi \nabla \varphi - \varphi \nabla \psi$. But this is a very special case where I have some information on the theorem already.

I have to let this go because it could take me 5 years to learn enough to prove it. So let's grit our teeth and accept that 5.72 and 5.73 are true, and we can wonder why Stak gives no reference. Perhaps these facts won't play a major role in what we are about to do in this volume, or perhaps we will get more hints later as to why they are true. [Well it took about 3 days, not 5 years.]

Examples: (40)

Example 1: $L = \nabla^2$. This is the one I can actually do. I agree that the current J is as given in 5.74 in n-dimensions, and I agree with the Green's Theorem of 5.75, also in n dimensions. But let's just do this again:

$$\int dV\, (v \nabla^2 u - u \nabla^2 v) = \int dS\, \hat{\mathbf{n}} \bullet \{ u \nabla v - v \nabla u \} \qquad \text{// this is Green \#2}$$

where we think of $\hat{\mathbf{n}} \bullet \nabla = \partial / \partial n$, the derivative in a direction normal to the surface σ. If we call the thing in curly brackets **J**, then we can claim from the divergence theorem that the surface integral shown above is equal to a volume integral of $\nabla \bullet \mathbf{J}$ and then we can claim that $v \nabla^2 u - u \nabla^2 v = \nabla \bullet \mathbf{J}$. So this whole thing hinges on these two integral theorems: Green's Second Identity, and The Divergence Theorem. I have derived Green #2 on my sheet using the Divergence Theorem applied to $\mathbf{A} = u \nabla v$ and I have used a fancy vector identity along the way.

Example 2: $L = \nabla^2 \nabla^2$ which he calls ∇^4 and the "biharmonic operator". I have never heard of such a thing, but it is obviously self-adjoint (formally) and he quotes the current J in this case and then in 5.76 we get the full bore Green's theorem with all the terms of the current on the RHS.

Example 3: $L = \partial_t - \nabla^2$ which he calls the **heat conduction operator**. In this case L* has a sign change for the time derivative. Again he states the current J and then the Green's Theorem.

Exercise: confirm that the current **J** here is correct. That is to say, prove that this is true:

$$v(\partial_t - \nabla^2)u - u(-\partial_t - \nabla^2)v = \nabla \bullet [\, \mathbf{e}_t\, uv - (v \nabla u - u \nabla v) \,]$$

The non-time terms on the two sides balance according to Example 1, so we only have to show this:

$$v(\partial_t u) + u(\partial_t v) = \nabla \bullet (\mathbf{e}_t\, uv)$$

where $\nabla \bullet$ is the usual 3D divergence operator. The RHS of the above, according to an identity p 8, is

$$\nabla \bullet (\mathbf{e}_t\, uv) = \mathbf{e}_t \bullet \nabla(uv) + (uv)\, \nabla \bullet \mathbf{e}_t = \mathbf{e}_t \bullet \nabla(uv) = \partial_t(uv)$$

Note that e_t is just a fixed unit vector $\hat{t}$ which has no spatial divergence. But we are now done, because it is obvious that $\partial_t(uv) = v(\partial_t u) + u(\partial_t v)$.

So here then is a statement of our Example 3 result:

$$\int_R dxdt \{v [\partial_t - \nabla^2]u - u [-\partial_t - \nabla^2]v \} = \int_R dxdt \ \nabla \bullet [\mathbf{e}_t uv - (v \nabla u - u \nabla v)]$$

$$= \int_\sigma d\mathbf{S} \bullet [\mathbf{e}_t uv - (v \nabla u - u \nabla v)]$$

$$= \int_\sigma (d\mathbf{S} \bullet \mathbf{e}_t) uv - \int_\sigma d\mathbf{S} \bullet (v \nabla u - u \nabla v)]$$

Note: The volume integral in (5.79) is over the n+1 dimensional volume and is written dxdt. This volume integral is over some region R_{n+1} in n+1 dimensional spacetime. The boundary of R_{n+1}, called σ_n, is an n dimensional surface. A surface element on this boundary would be written dS_n. In general this dS_n is going to involve some mixture of the t coordinate with some of the n x coordinates. So let's then rewrite the above Green's Theorem with more clarity on things:

$$\int_{R_{n+1}} d^n x \, dt \{v [\partial_t - \nabla^2]u - u [-\partial_t - \nabla^2]v \}$$

$$= \int_{\sigma_n} (d\mathbf{S}_n \bullet \mathbf{e}_t \ uv - \int_\sigma d\mathbf{S}_n \bullet (v \nabla u - u \nabla v)] \qquad (*)$$

Later on page 195 we will get an important example of an n+1 dimensional region R_{n+1} which is a certain "cylinder". One dimension of the cylinder is coordinate t (vertical). At time t, we have region $R = R_n$ which is a region of actual space, so in some sense we have $R_{n+1} = t \times R_n$ for our cylinder. This region R_n itself has a spatial boundary we might call σ_{n-1} with purely spatial surface patch dS_{n-1}. Thus we have $dS_n = dt \, dS_{n-1}$. When we try to compare page 195 to page 41, Stak has changed notation so we have to pay attention. For the moment I will put primes on the page 195 quantities

page 41	page 195	
$dx = dx^n$	$dx' = dx^n$	// so dx = dx', the same notation
$R = R_{n+1}$	$R' = R_n$ (space only)	
$dS = dS_n$	$dS' = dS_{n-1}$ (on boundary of space only region R_n)	
$\sigma = \sigma_n$	$\sigma' = \sigma_{n-1}$ (boundary of space only region R_n)	

Thus we could say dS = dt dS'. Let's now take (*) above and convert it to this page 195 notation. First, here is just the LHS of (*).

$$\int_0^T dt \int_{R_n} d^n x \ \{ v\,[\partial_t - \nabla^2]u \ - \ u \ [-\partial_t - \nabla^2]v \}$$

Now recall that we might write $d\mathbf{S_z} = dxdy\ \hat{\mathbf{z}}$ for a path on the x-y plane in regular 3D stuff. The direction of this patch is the "normal" here being $\hat{\mathbf{z}}$ and we could say $d\mathbf{S_z} = d\mathbf{x} \times d\mathbf{y}$ since we know that x and y are perpendicular, and of course we would have $d\mathbf{S_z} \bullet \hat{\mathbf{x}} = 0$ and $d\mathbf{S_z} \bullet \hat{\mathbf{y}} = 0$.

Thinking now about $d\mathbf{S_n}$, we can write

$$d\mathbf{S_n} = d\mathbf{S_{n-1}} \times d\mathbf{t} \qquad \text{and} \quad d\mathbf{S_n} \bullet \hat{\mathbf{t}} = 0 \text{ and } d\mathbf{S_n} \bullet d\mathbf{S_{n-1}} = 0$$

After all, we know that "t" is orthogonal to all n spatial coordinates, so we can break it out in this way. The point here is that we now have "directional information" on $d\mathbf{S_n}$.

Example 4: $L = \partial_t^2 - \nabla^2$ which he calls the **wave operator**. Again he states J and Green's. Unit vectors are called e_i and e_t for time. He rewrites the Green's several different ways using a mysterious new vector q called **the "transversal"** to the surface σ. This gives 5.83 which has a more standard form and probably something like this would be involved in the general proof of the claim made above. This example is totally incomprehensible.

Exercises: We are going to need this stuff later, I now know, so figure it all out please.

(a) Prove p 41 A which claims this:

$$v\Box^2 u - u\Box^2 v = \nabla \bullet \{ e_t(v\partial_t u - u\partial_t v) - (v\,\nabla u - u\,\nabla v) \}$$

where $\Box^2 = \partial_t^2 - \nabla^2$. As before, we can balance the ∇^2 part of things as per Example 1, so we have only to show this:

$$v\partial_t^2 u - u\partial_t^2 v = \nabla \bullet \{ \mathbf{e_t}(v\partial_t u - u\partial_t v) \}$$

As in Example 3, the RHS becomes $\mathbf{e_t} \bullet \nabla(v\partial_t u - u\partial_t v) = \partial_t (v\partial_t u - u\partial_t v)$. The cross terms cancel and we then just find this to be equal to $v\partial_t^2 u - u\partial_t^2 v$, QED!

(b) Question: what is this "transversal vector q" thing? We can write out his little equations like so:

$$(\mathbf{q} - \hat{\mathbf{n}}) \bullet \hat{\mathbf{t}} = 0 \qquad \Rightarrow \qquad (q_x, q_y, q_z, q_t) = (-n_x, -n_y, -n_z, n_t)$$
$$(\mathbf{q} + \hat{\mathbf{n}}) \bullet \hat{\mathbf{x}} = 0$$
$$(\mathbf{q} + \hat{\mathbf{n}}) \bullet \hat{\mathbf{y}} = 0$$

$(\mathbf{q}+\hat{\mathbf{n}}) \bullet \hat{\mathbf{z}} = 0$

You have to think of a surface σ in a 4D space, so that the normal to the surface has 4 components and is written (n_x, n_y, n_z, n_t). Our space is R_4 really. This vector is a 4D unit vector. Then the transversal vector is what you get by negating the spatial components only of the normal vector. Obviously **q** is also a unit vector in R_4 so we could write it as $\hat{\mathbf{q}}$. Notice that

$$\hat{\mathbf{n}} \bullet \hat{\mathbf{q}} = (n_x, n_y, n_z, n_t) \bullet (-n_x, -n_y, -n_z, n_t) = -n_x^2 - n_y^2 - n_z^2 + n_t^2$$

$$= -n_x^2 - n_y^2 - n_z^2 - n_t^2 + 2n_t^2 = -1 + n_t^2$$

so in general this dot product is not zero so $\hat{\mathbf{q}}$ is not in general perp to $\hat{\mathbf{n}}$, and therefore $\hat{\mathbf{q}}$ is not in general tangent to the surface at the point of $\hat{\mathbf{n}}$. You could think of t as "just another coordinate" in this problem. You can draw pictures with only x, and then with only x and y. Have to be careful with notation here because $\hat{\mathbf{n}}$ usually means a spatial unit vector.

(c) Prove 5.83. We have to show that

$$\hat{\mathbf{n}} \bullet \{ \mathbf{e}_t(v\partial_t u - u\partial_t v) - (v \nabla u - u \nabla v) \} = (v\, \partial u/\partial q - u\, \partial v/\partial q)$$

where I presume that $\partial u/\partial q = \hat{\mathbf{q}} \bullet \nabla_4 u$ where I imply the 4D gradient here. Now

$\nabla_4 u = (\nabla u + \mathbf{e}_t \partial_t)$ where I just add the 4th coordinate's contribution

Therefore:

$$(v\, \partial u/\partial q - u\, \partial v/\partial q) = v\, \hat{\mathbf{q}} \bullet \nabla_4 u - u\, \hat{\mathbf{q}} \bullet \nabla_4 v$$

$$= v\,(-\hat{\mathbf{n}} \bullet \nabla u) + v\, \hat{\mathbf{n}} \bullet \mathbf{e}_t\, \partial_t u - (u \leftrightarrow v)$$

Note: In this notation, ∇ is the spatial gradient only, while $\hat{\mathbf{n}}$ is the full 4D unit vector. In each dot product, some components give 0, ie, when we write $\nabla \mathbf{u}$ we imply that the time component is 0. Now we continue

$$= \hat{\mathbf{n}} \bullet \{ (- v \nabla u + u \nabla v) + \mathbf{e}_t (v\partial_t u - u\partial_t v) \}$$

But this is in fact the LHS of our equation to be proved, so we have shown RHS = LHS, QED. So if we now combine 5.81 and what I just proved here, we end up with 5.83. Very good.

Now I agree that if we have only spatial x, then our 2D space is x,t and then our volume is the enclosed area of some closed curve C we draw in the x,t plane. In this case, the surface integral dS is in fact just a line integral and we get 5.83A.

Example 4A: $L = \partial_t^2 - c^2\nabla^2$ where I have now added speed c. How are things now different?

(1) equation 41A last two terms have extra factor c^2, and same for 5.81

Now let's improve notation, then see what the new expression for q must be:

$n = (n_x, n_y, n_z, n_t) = (\mathbf{n}, n_t)$

$q = (q_x, q_y, q_z, q_t) = (\mathbf{q}, q_t)$

(c version A) **Prove 5.83.** We have to show that

$$n \cdot \{ \mathbf{e}_t(v\partial_t u - u\partial_t v) - c^2(v \nabla u - u \nabla v) \} = (v\, \partial u/\partial q - u\, \partial v/\partial q) \qquad (*)$$

Start by writing: (where both q and ∇ are 4-vectors)

$\partial u/\partial q = q.(\nabla u) \qquad \nabla u = (\nabla u, \partial_t u)$

$\partial u/\partial q = \mathbf{q} \bullet \nabla u + q_t \partial_t u$

Then we can write for the RHS of (*),

$(v\, \partial u/\partial q - u\, \partial v/\partial q) = v(\mathbf{q} \bullet \nabla u + q_t\partial_t u) - u(\mathbf{q} \bullet \nabla v + q_t\partial_t v)$

$= (v\, \mathbf{q} \bullet \nabla u - u\, \mathbf{q} \bullet \nabla v) + (v\, q_t\partial_t u - u\, q_t\partial_t v)$

$= \mathbf{q} \bullet (v \nabla u - u \nabla v) + (v\, q_t\partial_t u - u\, q_t\partial_t v)$

Now let's work on the LHS of (*)

$n \cdot \{ \mathbf{e}_t(v\partial_t u - u\partial_t v) - c^2(v \nabla u - u \nabla v) \}$

$= n_t (v\partial_t u - u\partial_t v) - c^2\, \mathbf{n} \bullet (v \nabla u - u \nabla v)$

$= - c^2\, \mathbf{n} \bullet (v \nabla u - u \nabla v) + n_t (v\partial_t u - u\partial_t v)$

From the second term we see that we need $q_t = n_t$. From the first term, $\mathbf{q} = - c^2\, \mathbf{n}$. Therefore

$n = (n_x, n_y, n_z, n_t) = (\mathbf{n}, n_t)$
$q = (q_x, q_y, q_z, q_t) = (-c^2\mathbf{n}, n_t)$

Strict and Generalized Solutions of Differential Equations (42)

Our ODE is now Lu=s, pretty compact notation for this complex subject! Regardless of the nature of "s", the solution "u" is called "a **strict solution** in an open region R of R_n" if Lu=s at every point in R and if u is C^p where L is order p. Probably this will be the case if s is a regular function (I am guessing).

If s is a distribution, like a delta function, then provided <Lu,φ> = <s,φ> for all K_n test functions, a solution u is called a **generalized solution**. Remember that this is not a local "point" equality Lu=s, it is a global equality since φ might have any support in R_n . So our generalized solution must be true for all φ you can make in R_n . You can restrict this to some region R and then have a generalized solution in R where you then only allow test functions with support in R.

As usual, we can always swing things to say <Lu,φ> = <u,L*φ>, so then <u, L*φ> = <s,φ> must be true for u to be a generalized solution.

Homogeneous Equation (42)

Of course one of our favorite equations is Lu=0 and now we want <u, L*φ>= 0 for all test functions.

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

The proof is given and we are not surprised at this result. Notice the use of Green's theorem 5.73 where one of the two functions u and v is a test function φ and the region R is the support region for φ. [We are back to the K_n test function space: we use Sn only when worrying about FT's.] We know that φ (and all derivatives) must vanish on the boundary since that is the stated limit of support, and that means the current J which is bilinear in u and φ (say) must vanish on the boundary. So basically we are proving in these words that <Lu,φ> = <u,L*φ> on region R where φ has support only in R. When we "swing" L over and it becomes L*, our "parts" vanish due to the nature of φ's support! Otherwise we would have non-zero parts due to the boundary current J. I think this is a very major point that I have not given enough attention to. This is one of the main reasons for defining K_n as we do! It makes those parts in the form of the surface term go away.

Let's review this proof again. First we always have this Green's thing being true for some current:

$$\langle u,L^*\varphi\rangle = \langle Lu,\varphi\rangle - \int d\mathbf{S}\bullet\mathbf{J}$$

If we have a homogeneous solution u of Lu=0, the first term on the RHS = 0. If we can show the second term is also zero, then we have shown <u,L*φ> = 0. This fact, I now realize, is the *definition* of L being a "generalized solution" of Lu=0. I missed that point above.

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

Again a proof is given. This is a little harder, because we start with the softer fact <Lu,φ> = 0 and we have to show the harder fact that Lu=0 at every point in R. The extra fact needed her about C^p causes Lu to be a continuous function which he then writes as q, and then he shows q = 0 using the fact that it is a smooth object.

[p 44] He claims that there exist generalized solutions which are NOT strict solutions in R. This must occur when the solution u is not C^p. We recall that things got singular in Chapter 1 when $a_0(x)$ vanished in the interval (a,b). In n dimensions, there is some matrix analog of this situation that we will study later.

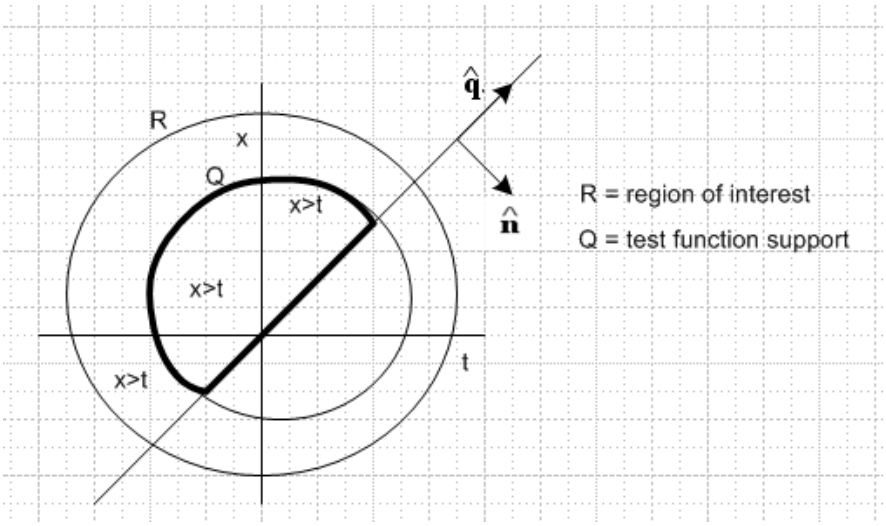
If we have this situation, then things are singular, and then a generalized solution u might not be strict. I suspect that this just means you have a solution which -- itself or its p derivatives -- is a distribution.

Example 1 (p 44) Wave with 1D space.

Wave equation in 1D space Lu=0. It is obvious that u(x,t) = f(x-t) is a solution, and this is a wave going to the right with unit velocity. First we think of f as C^2 and we are talking about a strict solution. But how about a solution u(x,t) = H(x-t) Heaviside, which is not C^2. He now shows that in fact this is a generalized solution of Lu=0. He starts with 5.89 where notice the integration is restricted to x > t and u does not appear inside since it is H.

The following steps are each hard for me to follow, so detailed notes are needed.

(a) Derive p 44 A. This equation is a statement of 5.83a, but we need a picture



The integral in 5.89 is only over the intersection of Q and R to the left of the line x = t since this is where x > t. This arises from the Heaviside function for u. So we can think of this semidisk (in my drawing) as the volume of the Green's integral, and then the "surface" for that integral is a line integral around the semidisk. Here is that line integral from 5.83a: (note sign is changed overall)

$$\int_{\texttt{semidisk}} d\ell\, (u\, \partial\varphi/\partial q - \varphi\, \partial u/\partial q) \qquad\qquad u = H(x-t)$$

Now he says we can throw out the contribution of the circular part because, this circular part being a boundary of Q, we know that φ and all derivatives vanish on that boundary. We are left with

$$\int_{\texttt{x=t segment}} d\ell\, (u\, \partial\varphi/\partial q - \varphi\, \partial u/\partial q)$$

Now I think we should regard this line segment as the limit of one lying just inside the semidisk where u = 1. Then the first term of the above agrees with what he shows on the RHS of p 44A. But what about this

second term, and what is q here? I have drawn $\hat{\mathbf{q}}$ by negating the spatial component of $\hat{\mathbf{n}}$ and we see that it lies along the line segment. Since u = H is a constant (call it 1) along the entire x=t line, the derivative $\partial u/\partial q = 0$. So the second term vanishes, even though φ has support along the line segment. So we have now shown that

RHS second term in p 44A $= \int_{\texttt{x=t segment}} d\boldsymbol{\ell}\, (\partial\varphi/\partial q)$

(a) Derive p 45 A. Now we can write this integral as $\int d\boldsymbol{\ell}\, (\partial\varphi/\partial\boldsymbol{\ell})$ where $\boldsymbol{\ell}$ is a coordinate along the line segment. But this integral is then φ(upper right end) - φ(lower left end) of the line segment (this is my integral theorem #9 on sheet 4). But these endpoints are on the Q boundary, so we get each number being 0, and thus

$$\int_{\texttt{x=t segment}} d\boldsymbol{\ell}\, (\partial\varphi/\partial q) = \int d\boldsymbol{\ell}\, (\partial\varphi/\partial\boldsymbol{\ell}) = \varphi(\text{upper right end}) - \varphi(\text{lower left end}) = 0 - 0 = 0$$

which is what p 45 A says. But for x>t, we know that Lu = 0 because this is a continuous region and we have 5.87 as our homo equation Lu = 0.

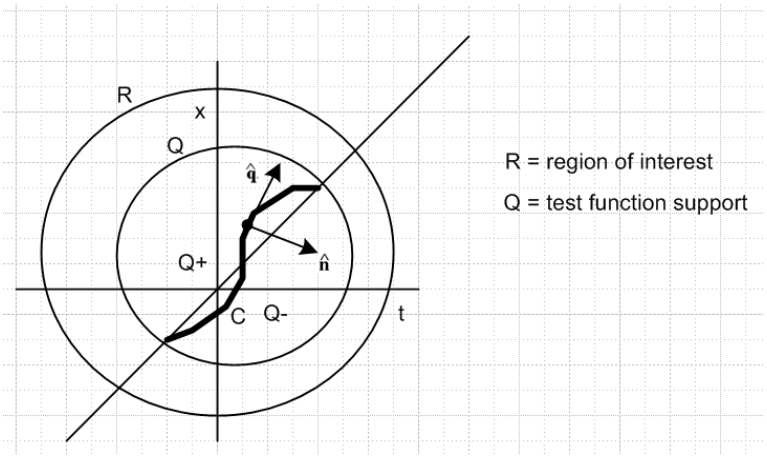
Therefore, everything on the RHS of 44A is 0, and we conclude that

$\langle u, L^*\varphi \rangle = 0$ // as claimed on page 45 B.

Since this is true for every test function φ, this shows that h = H(x-t) is in fact a *generalized* solution of our ODE Lu = 0 (again, since this is the definition of u being a generalized solution).

Comments: Notice the general flow of things here. We want to change from L*φ to Lu, flipping the derivatives from one side to the other inside an integral. We have an integral because we are discussing distributions, and they always involve integral-like forms which diffuse things out over some region of your space R_n and bring in the test function φ. In the very general case, we pick up "parts" which have the form shown in 5.73 where we have to figure out what J is for each L we want to think about. Once we have L in the form Lu, we can set Lu=0 or Lu=s based on our ODE. Then we have to deal with the "parts term". In our last example, since Lu = 0, that switched-over integral gave 0, and he showed that the parts gave zero as well.

Page 45 general discussion: Stak claims that our example above with Heaviside is an example of a more general situation which we can represent by this picture



I think the curve C (now not just a straight line) represents the most general discontinuity you might have in your solution function u(x,t), at least for the wave equation in 1D. If you stay away from this curve, things are smooth, so we have now smooth support regions Q+ and Q-. In our example, it happened that u was 0 in the Q- region, but in general that is not the case.

Now, if u is some generalized solution, it must satisfy page 45 C (still doing 1D wave L), then we break the support integral over Q into the two parts shown in D. Then just as we did in our example, each of these area integrals can be written in the form p 44A, and the first integral will be zero and we have only the second integral to worry about. That is to say, we have the (negated) right side of 5.83a to worry about, which now appears as (showing parts from both integrals)

$$\int_{C+} d\ell\, (u\, \partial\varphi/\partial q - \varphi\, \partial u/\partial q) \; - \int_{C-} d\ell\, (u\, \partial\varphi/\partial q - \varphi\, \partial u/\partial q)$$

Here I have defined C+ and C- to be a line integral in the same direction, but just inside its respective side. I guess I have chosen the vectors n and q for the Q+ side here in both integrals. Thus, we can combine these to get

$$\int_{C} d\ell\, (\Delta u\, \partial\varphi/\partial q - \varphi\, \Delta(\partial u/\partial q)) \qquad \text{// which is 5.90 p 45}$$

and we pick up two discontinuities across the contour C.

For the most general case that the q vector is in some arbitrary direction, this integral vanishes only if both discontinuities vanish at all points along C, and this is equation p 45 E. On page 46 top, these seem to lead to a contradiction, since we assumed I think that C was a line of discontinuity of u. [So I think the point here is that for general q directions, you conclude that $\Delta u = 0$ AND that $\Delta(\partial u/\partial q) = 0$. As discussed soon, he claims that this means there is in fact no discontinuity at all on your line C, so in fact you have in fact a strict solution. But we are seeking a generalized solution that DOES have some discontinuity across C, and this will then NOT be a strict solution. That is the case we now consider:]

Suppose that q is tangent so that q.n = 0. This would imply p 46C be true at every point on C, where as usual e_t and $e_1 = \hat{\mathbf{x}}$ are the fixed unit vectors. This means C is a straight line either x = t + C or x = -t + B where B is some constant, and these lines are called **characteristics**. If we have this case for C, namely that C is a characteristic of the problem, then variable q is along the characteristic (since n.q=0) and then we can replace variable q with the line integral variable ℓ as shown in D, so we now require D to be true if

u is to be a generalized solution. We then trivially rewrite D as 5.91 and the first integral is of a perfect line differential and equals the difference in end point values which are both 0 as in our example above, so we are then left with only the second integral. Thus, for u to be a generalized solution, we must have E be true (and our curve C must be a characteristic). This E integral is zero when Δu = constant. [I think this means that there is no tangential discontinuity but says nothing about a perpendicular discontinuity across C.]

So this ends our example of the 1D wave equation situation, and I have underlined in red on page 46 a precise statement of what we have learned here: (1) for the 1D wave, the only curves C which can "carry" discontinuities of u are the characteristics, and (2) the discontinuity of u across a characteristic must be a constant.

Example 2 (p 46 bottom) Laplace in n dimensions.

Now we want to think about the Laplace equation $L = \nabla^2$ in n dimensions.
The equation is Lu=0. We imagine a test function supported on some region Q in R_n, and we imagine that the solution has some singularity on a surface σ which divides Q+ and Q-. [σ generalizes C]

He injects a caveat at this point saying that we assume a "simple discontinuity" across the surface σ. I think by this he means that u is not infinite at any point along the surface, on either side, and that we just have some finite discontinuity of finite values on the two sides. I think this may be what justifies writing the integral as Q+ and Q- pieces and ignoring a possible delta function type contribution from the integration contribution of the surface σ itself, despite its having zero volume.

Now, assuming the above, we break our volume test function integral as in E into the two parts, and for each part, we use our Green's theorem. This was 5.83a for the wave L in, but for the Laplace L it is 5.75 which is simpler that the wave case in that we only have the n vector to think about and no q vector. In fact, everything is exactly as it was before with q replaced by n, and we arrive at our requirement p 47 A which must be true for u to be a generalized Laplace solution! In this case, there are no "characteristics" with n.q=0 or any of that, we must have to have Δu and $\Delta(\partial u/\partial n)$ separately be 0 (always these things are for arbitrary φ). Saying $\Delta u = 0$ he says implies that all derivatives tangential to σ are continuous across σ, but the second condition says the normal derivative must also be 0. So waving the wand, he says that everything is smooth across our hypothetical simple singularity boundary σ. Thus, it is not really a singularity. In other words, if we assume we have a generalized solution with such a singularity, we find that there is no singularity and everything is smooth in Q. Said another way: for Laplace Lu=0, every generalized solution is a strict solution, and you cannot have "strange things" happen out in the middle of nowhere the way you can have in the wave L case! Very good, this was a long tough section for me.

Comments: 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!

Exercises (47)

Exercise 5.23. OK, in the above Example 2 we treated Laplace in R_n, but in Example 1 we treated wave only in R_1 for the space part. Here we want to look at the more general wave L with n spatial dimensions! The first question: What does the Green's Theorem say in this new case? We have already computed this thing and it is shown in 5.83. However, because we now have $c \neq 1$, the q appearing in 5.83 is this:

$$n = (n_x, n_y, n_z, n_t) = (\mathbf{n}, n_t)$$
$$q = (q_x, q_y, q_z, q_t) = (-c^2\mathbf{n}, n_t)$$

What are the characteristics of this problem? Well, first off, if we imagine Q+ and Q- divided by surface σ, our parts are

$$\int_\sigma dS\ (u\ \partial\varphi/\partial q - \varphi\ \partial u/\partial q)$$

where this replaces our curve C segment integral in Example 1. We think of the starting volume integral as two volumes, Q+ and Q-. For each integral, we use Green's, so each integral results in a σ surface integral. When we add these and pick the q vector say on the Q+ side, we get,

$$\text{parts} = \int_\sigma dS\ (\Delta u\ \partial\varphi/\partial q - \varphi\ \Delta(\partial u/\partial q))$$

In order to find possible "characteristics", we imagine a case where q is tangential to the surface σ. In the simple Example 1 case, we were able to throw out the $\Delta(\partial u/\partial q)$ term because we had a Heaviside function which was "1" all along one side of the surface. A similar thing is going to happen here. We know that a solution to our wave equation Lu=0 is some $u(\mathbf{r},t) = f(\mathbf{r}\bullet\hat{\mathbf{k}}\text{-}ct)$ as one can trivially show. Then we consider the Heaviside version $H(\mathbf{r}\bullet\hat{\mathbf{k}}\text{-}ct)$. This thing is "1" along the surface, so $\Delta(\partial u/\partial q) = 0$ if q is tangential. This leaves us with the other term $\int_\sigma dS\ (\Delta u\ \partial\varphi/\partial q)$. But here $\Delta u = 1$, so we have

$\int_\sigma dS(\partial\varphi/\partial q)$, which generalizes our previous result $\int_C d\ell\ (d\varphi/d\ell)$. This last we wrote as $\varphi() - \varphi() = 0 - 0 = 0$. So what does our integral theorem #9 look like in this more general case? Try this:

$$\int_\sigma dS(\partial\varphi/\partial q) = \int_\sigma dS\ \hat{\mathbf{q}} \bullet \nabla\varphi = \int_\sigma dS\ \{\ \nabla\bullet(\hat{\mathbf{q}}\varphi) - (\nabla\bullet\hat{\mathbf{q}})\varphi\ \}$$

Now I am going to just assume that σ is a planar surface which means $\hat{\mathbf{n}}$ is fixed, which then means $\hat{\mathbf{q}}$ is fixed, which means $\nabla\bullet\hat{\mathbf{q}} = 0$ so the second term above is 0. For the first term, I use the divergence theorem in this form, where I now treat dS as the "volume" which has boundary Dσ

$$\int_\sigma dS\ \nabla\bullet(\hat{\mathbf{q}}\varphi) = \int_{D\sigma} dS' \bullet (\hat{\mathbf{q}}\varphi)$$

Then we argue that the boundary Dσ aligns with the boundary of the Q support region at on such boundary we now that $\varphi = 0$, so this term goes away.

Our first surface that comes to mind is where the argument of H($\mathbf{r}\bullet\hat{\mathbf{k}}$-ct) is 0, or the surface $\mathbf{r}\bullet\hat{\mathbf{k}} = ct$. This surface is a plane with normal vector $\hat{\mathbf{k}}$ which is perp distance ct from the origin.

So here is the upshot: if we assume a planar surface σ, we have a surface which supports $\Delta u = 1$ and so we have a true generalized solution of the wave equation. So such a planar surface is the characteristic of this problem. Our assumption of a "tangential q" means n.q = 0. Write this out,

$$n = (n_x, n_y, n_z, n_t) = (\mathbf{n}, n_t)$$
$$q = (q_x, q_y, q_z, q_t) = (-c^2\mathbf{n}, n_t)$$

$$n.q = -c^2\,\mathbf{n}\bullet\mathbf{n} + n_t^2 = 0 \quad \Rightarrow \quad c^2\,\mathbf{n}\bullet\mathbf{n} = n_t^2$$

which we can rewrite as

$$c^2 \Sigma_i\, n_i^2 = n_t^2 \qquad \text{or} \qquad c^2 \Sigma_i\, (n.e_i)^2 = (n.e_t)^2$$

which is the desired exercise result (5.94). The characteristics are $\mathbf{r}\bullet\hat{\mathbf{k}} \pm ct$ = constant, which generalizes the result $x \pm t$ = constant after C on page 46, where now $\hat{\mathbf{k}}$ is any spatial unit vector you want.

Exercise 5.24. $L = \partial_{x1} + \partial_{x2}$. This is not a case formerly considered. What is the Green's for this L?

$$vLu + uLv = \text{div } J = \partial_{x1}J_1 + \partial_{x2}J_2$$

$$LHS = vLu + uLv = v\,(\partial_{x1} + \partial_{x2})u + u\,(\partial_{x1} + \partial_{x2})v$$

$$= v\,(\partial_{x1}u) + u\,(\partial_{x1}v) + \text{same for } x_2$$

$$= \partial_{x1}(vu) + \partial_{x2}(vu)$$

So I would then claim that $J_1 = J_2 = vu$ so that

$$\mathbf{J} = (vu)\,(\mathbf{e}_1 + \mathbf{e}_2)$$

Green's then says

$$\int dx_1dx_2\,(vLu - uL^*v) = \int_C d\ell\ \mathbf{n}\bullet\mathbf{J} = \int_C d\ell\,(n_1 + n_2)\,uv$$

Now, we want to find a generalized solution u such that $\langle Lu,\varphi\rangle = 0$ which means $\langle u,L^*\varphi\rangle = 0$. Our Green's above then says:

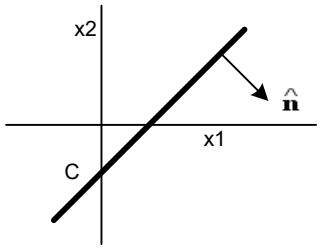
$$\int dx_1dx_2\,(\varphi Lu - uL^*\varphi) = \int_C d\ell\ \mathbf{n}\bullet\mathbf{J} = \int_C d\ell\,(n_1 + n_2)\,u\varphi$$

$$\langle u,L^*\varphi\rangle = \langle\varphi,Lu\rangle - \int_C d\ell\,(n_1 + n_2)\,u\varphi$$

As usual, we wrote our volume (area) as two pieces separated by a potential discontinuity. We try for this solution $u = H(x_1 - x_2)$ so our candidate line is $x_1 - x_2$ = constant. We have a picture similar to the one we had above with the Q+ and Q– regions. The Q- region gives 0 since H, the Q+ region gives 0 since Lu=0, so we are left with

$$\langle u, L^*\varphi \rangle = \int_C d\ell\, (n_1 + n_2)\, (\Delta u)\varphi = \int_C d\ell\, (n_1 + n_2)\, \varphi$$

Here is our picture



and we can see that $n_1 > 0$, and $n_2 < 0$ and $n_1 + n_2 = 0$ so our parts = 0, and we then conclude that

$$\langle Lu, \varphi \rangle = \langle u, L^*\varphi \rangle = \int_C d\ell\, (n_1 + n_2)\, \varphi = 0$$

for any φ with support Q, etc etc. So this shows that $u = H(x_1 - x_2)$ is a *generalized* solution. We know that anything of the form $u = f(x_1 - x_2)$ formally solves the $Lu = 0$. Notice that $x_1 = x_2$ + constant is a characteristic, but that $x_1 = -x_2$ + constant is NOT a characteristic because in that case, $n_1 = n_2 > 0$ and we cannot kill off the parts term.

Exercise 5.25. Here we look at $L = \nabla^4$ which is the fancy Laplace thing for which we know J from page 40. The RHS of the Green's theorem involves only n, not q, as with regular Laplace. So I think this fact will yield the same conclusion which we got for regular Laplace. That is, we will get something like that shown ion 47 A and we conclude that there can be no discontinuity.

Exercise 5.26. On page 44 we showed that H(x-t) and H(x+t) were generalized solutions of the 1D wave equation. Here we want to show that δ(x-t) and δ(x+t) are also generalized solutions. They do of course have the right form f(x-t).

(a) Try if first myself.

Before looking at the hints, what would I do? How does this differ from the Heaviside example #1 above? We start this way:

$\langle Lu, \varphi \rangle = \langle u, L^*\varphi \rangle = 5.89$

= something like p 44A by using Green's theorem

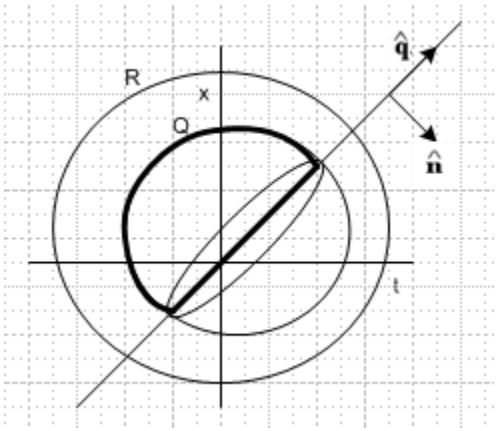
The first term in our new version of p 44A will have contributions from the Q+ and Q- areas. However, in each of these areas, we will set Lu=0 even for u being a δ. So we are left only with the second term, which in general looks like this:

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

Now we can again imagine that our line C is the line x = t, so that q = $\boldsymbol{\ell}$ and this becomes

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

In Example 1 we had u = Heaviside so that Δu = 1. But what do we do now when u = δ? It is a δ that peaks along our characteristic line. Maybe I should *back up* to the "volume" (ie, area) integral for a region surrounding our line C. That would be an area integral over the thin oval shown below



In other words, we don't use Green's theorem at all, we just do the "volume" integral directly. As already noted, the Q+ and Q- regions give 0 since Lu=0 even for u = δ, so we have only the oval to worry about. then we have

$$\langle Lu,\varphi\rangle = \langle u,L^*\varphi\rangle = \int_{\mathtt{oval}} dxdt\ \delta(x-t)\ (L\varphi)$$

Suppose I convert the oval to a thin rectangle and change variables so that

t' = x+t = distance along the line C t'+x' = 2x t'-x' = 2t
x' = x-t = distance perp to the line C

Now let's say that Φ(t',x') = φ(x,t). What then is L'? If we ignore scaling factors for L' and for the Jacobian, we are going to end up with this

$$= \text{constant} * \int dx'dt'\ \delta(x')\ [L'\ \Phi(t',x')]$$

$$= \text{constant} * \int dt' \; [L' \, \Phi(t',x')]|^{x'=0}$$

Since Φ is an arbitrary test function, why should this come out being 0?

Aside for some details. Well, go back and check things a bit

$$dt' = dx + dt$$
$$dx' = dx - dt$$

$$\partial/\partial t' = (\partial t/\partial t') \, \partial/\partial t \; + (\partial x/\partial t') \, \partial/\partial x = 1/2(\partial/\partial t \; + \; \partial/\partial x)$$
$$\partial/\partial x' = (\partial t/\partial x') \, \partial/\partial t \; + (\partial x/\partial x') \, \partial/\partial x = 1/2(-\partial/\partial t \; + \; \partial/\partial x)$$

$$J' \equiv (\partial_{t'}^2 \; - \partial_{x'}^2) \; = 1/4 \, *(\, [\, \partial_t + \partial_x]^2 \, - \; [\, -\partial_t + \partial_x]^2 \,)$$

$$= \; 1/4 \, * \, (2\partial_t^2 + 4\partial_t\partial_x + 2\partial_x^2) \; = 1/2 \, (\partial_t^2 + 2\partial_t\partial_x + \partial_x^2) \; = 1/2 \, (\partial_t + \partial_x)^2$$

Hmm, I was wrong about that. So let's go the other way

$$\partial/\partial t = (\partial t'/\partial t) \, \partial/\partial t' \; + \; (\partial x'/\partial t) \, \partial/\partial x' \; = \; \partial/\partial t' \; - \partial/\partial x' \qquad \partial_t = \partial_{t'} - \partial_{x'}$$
$$\partial/\partial x = (\partial t'/\partial x) \, \partial/\partial t' \; + (\partial x'/\partial x) \, \partial/\partial x' \; = \; \partial/\partial t' \; + \partial/\partial x' \qquad \partial_x = \partial_{t'} + \partial_{x'}$$

$$L = (\, \partial_t^2 \; - \partial_x^2) \; = (\, [\, \partial_{t'} - \partial_{x'}]^2 \, - \; [\, \partial_{t'} + \partial_{x'}]^2 \,)$$

$$= \; 2 \, (\partial_{t'}^2 \, - 2\partial_{t'}\partial_{x'} + \partial_{x'}^2) \quad = 2 \, (\partial_{t'} - \partial_{x'})^2 \; \equiv \; L'$$

OK, so let's go back and redo things a bit.

$$\langle Lu,\varphi \rangle \; = \langle u,L^*\varphi \rangle \; = \int_{oval} dxdt \; \delta(x-t) \; (L\varphi)$$

$$= \text{constant} * \int dx'dt' \; \delta(x') \; [L' \, \Phi(t',x')]$$

$$= \text{constant} * \; \{ \int dt' \; [(\partial_{t'} - \partial_{x'})^2 \, \Phi(t',x')] \; \} \; |^{x'=0}$$

I would like to somehow cast this as the integral of a perfect differential and then claim that we get the difference between two endpoint values both of which are 0 due to the support of φ being 0 there. But I am unable to achieve this perfect differential form, so I guess I will now look at the "hints".

(b) Look at the Hints

$$\langle u,L^*\varphi \rangle \; = \langle u,L\varphi \rangle = \langle \, \delta(x-t),L\varphi \rangle = \int_{plane} dxdt \; \delta(x-t) \; [L\varphi](x,t) = \int_{-\infty}^{+\infty} dt \; \; [L\varphi](t,t)$$

He claims this integral is 0 for every φ. He says to do exactly my change of variables above. So he claims that you can show this:

$$\int_{-\infty}^{+\infty} dt \ [(\partial_t^2 - \partial_x^2)\varphi(x,t)]\,|^{x=t} = 0 \qquad \text{for all } \varphi$$

His method is to introduce

$$\psi(x,t) = [(\partial_t^2 - \partial_x^2)\varphi(x,t)]$$
$$\psi(t,t) = [(\partial_t^2 - \partial_x^2)\varphi(x,t)]\,|^{x=t}$$

Then show that

$$\int_{-\infty}^{+\infty} dt\ \psi(t,t) = 0$$

So I agree with his statement of the problem we have to solve, showing this is 0 for all φ. So basically he is saying that I was on the right track. His contribution is to remind me that we can have those infinite endpoints and not just end on the support of φ. So back to what I was doing

$$= \text{constant} * \{ \int_{-\infty}^{\infty} dt' \ [(\partial_{t'} - \partial_{x'})^2 \Phi(t',x')] \ \}\,|^{x'=0}$$

t' = x+t = distance along the line C t'+x' = 2x t'-x' = 2t
x' = x-t = distance perp to the line C

Plan A: I can regard the integrand as three terms. I think each term is separately 0.

$$\{ \int_{-\infty}^{\infty} dt' \ \partial_{t'}^2 \Phi(t',x')] \ \}\,|^{x'=0}$$

Here, do parts twice, in each case the parts vanish since φ is a test function, and in each case the other term is 0 because the function is just 1.

$$-\{ \partial_{x'} [\int_{-\infty}^{\infty} dt' \ \partial_{t'} \Phi(t',x')] \}\,|^{x'=0}$$

For the inner integral, do parts once and get 0, so this term is 0.

$$-\{ \partial_{x'}^2 [\int_{-\infty}^{\infty} dt' \ \Phi(t',x')] \}\,|^{x'=0}$$

This is the term I have trouble with! The inner integral is some f(x') that does not vanish. So my little three terms method is not working.

Plan B: So back up and try what I was doing before

$$= \text{constant} * \{ \int_{-\infty}^{\infty} dt' \ [(\partial_{t'} - \partial_{x'})^2 \Phi(t',x')] \ \} \ |^{x'=0}$$

$t' = x+t$ = distance along the line C $\qquad t'+x' = 2x \qquad t'-x' = 2t$
$x' = x-t$ = distance perp to the line C

$\partial/\partial t = (\partial t'/\partial t)\, \partial/\partial t' + (\partial x'/\partial t)\, \partial/\partial x' = \partial/\partial t' - \partial/\partial x' \qquad \partial_t = \partial_{t'} - \partial_{x'}$
$\partial/\partial x = (\partial t'/\partial x)\, \partial/\partial t' + (\partial x'/\partial x)\, \partial/\partial x' = \partial/\partial t' + \partial/\partial x' \qquad \partial_x = \partial_{t'} + \partial_{x'}$

Change variables BACK to t and x. But then let dt' = dt somehow, and then we have

$$= \text{constant} * \int_{-\infty}^{\infty} dt \ \{ \partial_t^2 \varphi(t,x)] \ \} \ |^{x=t}$$

$$= \text{constant} * \int_{-\infty}^{\infty} dt \ \{ \partial_t^2 \varphi(t,x=t) \}$$

Now in this form we can do parts twice and everything vanishes. Or we could regard it as a perfect differential and do our difference of finite endpoints, each being 0.

OK, I am not very happy about this exercise. Go back to the very beginning:

Show: $\int_{-\infty}^{+\infty} dt \ [(\partial_t^2 - \partial_x^2)\varphi(x,t)] \ |^{x=t} = 0$ for all φ

Write integral as

$$\int_{-\infty}^{+\infty} dt \ [(\partial_t - \partial_x) (\partial_t + \partial_x)\varphi(x,t)] \ |^{x=t}$$

Change variables and get something like this

$$\int_{-\infty}^{+\infty} dt' \ [(\partial_{t'}, \partial_{x'})\Phi(x',t')] \ |^{x'=0}$$

$$= \int_{-\infty}^{+\infty} dt' \ \partial_{t'} \{ \partial_{x'}\Phi(x',t')] \ |^{x'=0}\} \qquad (*)$$

Now we have our in the curly brackets

$$\{ \partial_{x'}\Phi(x',t')] \ |^{x'=0}\} = \{ \partial_{x'}\varphi(x(x',t'), t(x',t')] \ |^{x'=0}\}$$

$$= \{ \partial\varphi/\partial x * \partial x/\partial x' + \partial\varphi/\partial t * \partial t/\partial x' \} \ |^{x'=0}$$

But since support for φ and all derivatives is finite, this function vanishes at ± ∞, and therefore the parts vanish in the above (*) if we do parts there, and thus we have shown that

$\langle u = \delta(x-t), L\varphi \rangle = 0$ for all test functions φ

and therefore we have shown that

$\langle L\,\delta(x-t), \varphi \rangle = 0$

and that therefore δ(x-t) is a *generalized* solution of Lu=0. If we start with δ(x+t) instead, we get:

Show: $\int_{-\infty}^{+\infty} dt\ [(\partial_t^2 - \partial_x^2)\varphi(x,t)]\,|^{x=-t} = 0$ for all φ

Write integral as

$$\int_{-\infty}^{+\infty} dt\ [(\partial_t - \partial_x)(\partial_t + \partial_x)\varphi(x,t)]\,|^{x=-t}$$

Change variables but define x' and t' the reverse of before and you get same result as before

$$\int_{-\infty}^{+\infty} dt'\ [(\partial_{t'}\,\partial_{x'})\Phi(x',t')]\,|^{x'=0}$$

so the same conclusion obtains. Enough!!!

Exercise 5.27. Now we have $L = \partial_{x_1}\partial_{x_2}$ and we want to show that $\delta(x_1 - a)$ is a generalized solution.

$$\langle Lu,\varphi \rangle = \langle u, L\varphi \rangle = \int dx_1 dx_2\ \delta(x_1-a)\,[\partial_{x_1}\partial_{x_2}\varphi\,]$$

$$= \int dx_2\ [\partial_{x_1}\partial_{x_2}\varphi(x_1, x_2)\,]|^{x_1=a} = \int dx_2\ \partial_{x_2}\{\partial_{x_1}\varphi(x_1, x_2)\}\,|^{x_1=a}$$

This is like the last problem and we do parts and claim the parts are 0 since any combination derivative of φ vanishes at the ∞ endpoints. And of course then $\delta(x_2 - b)$ by symmetry is also a generalized solution.

In a naive sense, Lu=0 has strict solution $u(x_1, x_2) = f(x_1) + g(x_2)$ where f and g are differentiable. We can think of $f(x_1) = \delta(x_1 - a)$ as a special case in the symbolic function sense.

Exercise 5.28. We showed that for the 1D wave equation (Example 1), there are characteristics x±t = K so that we could have a generalized solution with a line of discontinuity. In Exercise 5.23 I generalized this to n spatial dimensions and the characteristics were planes or something like that. On the other hand, we showed for the general n-dimensional Laplace equation (Example 2) that you cannot have such a generalized solution with a discontinuity, only a strict solution. The main difference was that in Example 1 (wave equation) we had "q" appearing as in 5.83 in the parts part, whereas in Example 2 (Laplace), we

had only "n" appearing as in 5.75. In the Heat equation (page 41) we have 5.79 as a starting point and it is not obvious to me which way it will go, so let's copy paste and modify something we did above:

Somehow the answer must be this

$$\text{parts} = \int (v\,\partial_{\mathbf{n}}u - u\,\partial_{\mathbf{n}}v + n_t uv)dS$$

where $\partial_{\mathbf{n}}u = \mathbf{n} \bullet \nabla u$ in a spatial sense only, as in the Laplace case. There is no need for a q vector I suspect in the heat case. So the heat case is like the Laplace case but we have an extra term uv which we have to show cannot support a discontinuity.

I guess I will put off this exercise for yet another rainy day. Certainly if the result is important, we are going to be told what it is soon enough. Remember the structure of this volume:

theory chapter, Laplace operator, heat and wave operator, variational

so heat will get lots of mention. Note that heat and wave both have time, so "evolution".

5.8 Fundamental Solutions (48)

For some reason, the n-dimensional Green's function g(x|ξ) is called E(x|ξ), maybe just to remind us we are not necessarily in the 1D case, or maybe just because we don't yet have BC's imposed. The distributional PDE of interest is $LE = \delta_\xi$ as in 5.95, and if we think we have a solution E_ξ, we should in theory be able to show that $\langle E_\xi, L^*\varphi\rangle = \langle LE_\xi, \varphi\rangle = \langle \delta_\xi, \varphi\rangle = \varphi(\xi)$ for all test functions. This solution E(x|ξ) is called the "fundamental solution for L with pole at ξ". We are in n-dimensions total here, so $E_n(x|\xi)$.

Page 49 then has a section for the constant coefficients case (which is really the main interest of this entire volume!) and he uses the convolution notation to obtain results that are familiar to us from volume 1, in the form of 5.101 in general, and 5.100 for the constant coefficient case. In this case, we can always write $L_x = L_{x+b}$ for any constant b and this is what gives the simplification in this case. [Note that neither 5.100 or 5.101 is used in this long section 5.8.]

Most of what we do below will use ξ = 0, the reason being that we will be using constant coefficient L's, and we always know that $E_\xi(x) = E_0(x-\xi)$ so we might as well just solve $LE = \delta_0 = \delta$ for $E_0(x)$ and then make the replacement $x \to x-\xi$ for $LE = \delta_\xi$.

Now we are going to study certain operators L one at a time and compute E. Here is an outline of what is to come:

(1) Laplace Equation (p49) . He will do this in 3, then 2, then 1, then n dimensions.

(2) Helmholtz equation (53). This is Laplace with an extra λ term. This is really the wave equation where we have replaced ∂_t^2 with $-k^2$ which is the λ term. But here, we are just "doing the math".

(3) Heat Equation (58) one time and n spatial dimensions

(4) Wave Equation (61) same

(5) Wave equation with damping (65) same

(1) Laplace Equation (49). We take our LE = $(-\nabla^2)E = \delta$ equation and integrate over a little sphere and we get 1 due to the unit source. Then we apply the divergence theorem to get

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

where in the last forms we apply to the sphere in n dimensions. We have assumed that E is spherically symmetric, see various notes below for why this assumption is justified.

3D: (n=3) Here we know that our Laplace is $1/r^2\partial_r(r^2\partial_r E) = 0$ for r>0 since we read this off our general math notes. The general solution is E = C/r + D and we set D = 0 since we want E(r=∞) = 0. The condition above then tells us that $-1 = S_3(1)\,(-C/r^2)\, r^2 = -S_3(1)\, C$ and $S_3(1) = 4\pi$, so C = $1/4\pi$ and we get the famous result that E(r) = $1/(4\pi r)$ as our solution -- the potential of a unit point charge. [We are certainly not surprised to learn that the only solution of Poisson with a δ charge is spherically symmetric.]

Comment: I would say that this solution to our distributional equation is in fact a *function*. True, it diverges at r = 0, but that does not make it not a function. Remember that Laplace only has "function" solutions, no generalized solutions!

2D: This time we look at our cylindrical coords (or the new polar doc I just wrote) and we see that our Laplace is $1/r\partial_r(r\partial_r E) = 0$. The solution is Cln(r) + D so $(\partial E/\partial r) = C/r$ and our condition is

$$-1 = S_2(1)\,(\partial E/\partial r)\, r^{2-1} = 2\pi\,(C/r)\, r = 2\pi C \qquad C = -1/2\pi \qquad E = -\ln(r)/4\pi$$

and again we have a function as our solution. (I think this is the potential of a unit/length line charge)

1D: Here he shows that E(r) = |r| is the symmetrical solution, page 51.

Comment: in all these cases, we are *looking for* solutions of the form E(**r**) = E(r). Just because the equation LE = δ looks the same in any rotated coordinates does not tell me that there are no solutions that depend on angles. For example, the hydrogen SE is "rotationally invariant" in this way, but we know there are all kinds of angular solutions which are the famous orbitals. For some reason Stak does not discuss, we are restricting our interest to E(r) solutions.

Explanation: The SE says $\nabla^2\psi + (V(r)-E)\psi = 0$ and is *homogeneous* so the usual separation of variables works and you get spherical harmonics for angular dependence of solutions. As shown at the end of this section below, if you have $L\psi = f(r)$ where f(r) is some *inhomo* driving term (like δ) which is spherically symmetric, then you must have $\boldsymbol{\ell} = 0$ for any spherical harmonic and solutions are indeed only radially symmetric.

On pages 51-52 Stak then *proves* that <E,L*φ> = φ(0), special case of 5.96. He does this in the 3D case where we know that E = 1/4πr. He uses Green's theorem applied to a ball with a spherical hole in it and this gives 5.107. He then does some ε limit work on the central ball and out pops the desired result.

nD: We consider finally the LE=δ equation for Laplace in n-dimensions and again *looks for* E(r) type solutions. The result for n > 2 is $E_n(r) = C_n / r^{n-2}$ and we then just carry through our program to evaluate our constant C_n and get the final result for $E_n(r)$ as shown in 5.111.

Comment: In 1D, we had Lg = δ and when we integrated over the x=0 point we got $\int Lg\, dx = 1$. Only the second derivative in L contributed here and we got $a_2(0)\, g'(0)|^{+}_{-} = 1$ which was our jump condition on the Green's g. In > 1 dimensions, we integrate over a small spherical region around x = 0 and use the divergence theorem to set the constant. This theorem does not make much sense for n = 1 ? :

Comments on the Divergence Theorem

Divergence theorem in n dimensions: $\int dV\, \nabla\bullet\mathbf{A} = \int d\mathbf{S}\bullet\mathbf{A}$

Divergence theorem in 1 dimension where $A_x = A$ is the only component of **A** ,

$$\int dV\, \nabla\bullet\mathbf{A} = \int d\mathbf{S}\bullet\mathbf{A}$$

$$\int_a^b dx\, dA/dx = A(x)|^b_a = A(b) - A(a)$$

But we can make this interpretation where p are the points which make the boundary

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

and in this way retain the validity of the divergence theorem even for n = 1. In this n=1 case, we can regard the "volume" as a line segment (a,b), and we can regard the bounding surface as "the points a and b at the ends of the segment", where the right endpoint has normal pointing to the right, and the left end point has the negative of this normal, so you get the difference as shown.

Notice that this has NOTHING to do with "parts integration". I will clarify this later. Notice on the "integral theorems" page that Divergence Theorem, Parts Integration, and Green's Theorem for Laplace L are three distinct theorems!

(2) Helmholtz Equation (53).

As usual, we define $\sqrt{\lambda}$ to have a positive imaginary part and we set $\sqrt{\lambda} = i\mu$, etc. He puts the cut to the right I think. We then assume a radial solution to the Helmholtz in R_n, and we use the radial ∇^2 form we found in 5.110. This gives 5.114. He converts this to Bessel's with a simple substitution, and writes the

solution as Hankels. If $\sqrt{\lambda}$ has positive imaginary part, we throw out H_2 to get just H_1 which has expo decay and we arrive at 5.116 for our solution $E_n(r; \lambda)$. We apply our usual trick to find the constant, in this case making use of the small r behavior of the $H^{(1)}$ function (ie, we use a tiny sphere in our little divergence theorem application). Now we have the constant, as in 5.118. We can replace the $H^{(1)}$ with a modified Bessel K function if we want as in 5.119.

OK, so we have our general solution for $n \geq 2$ given by this 5.119. He then writes out the n=2 case which is 5.120 which is a K_0 function. For n=3 we get a simple Yukawa potential 5.121. And finally for n = 1 we get just the Yukawa exponential part.

He comments (but thank goodness does not show it) that you can prove as we did in the Laplace case that the $E_n(r;\lambda)$ we have found here is indeed the distribution which solves the distribution ODE in the distribution sense. [ie, it is a generalized solution.]

When λ is on the plus real axis, we no longer have $\text{Im}\sqrt{\lambda} > 0$ so we have to ponder more what the solution is based on our problem, this will be considered later.

So at this point we have completely solved our little Green's problem, the result being 5.118 for the quantity $E_n(r; \lambda)$ as a Hankel-1 which is valid for all values of n and for λ not on the real axis.

Now he is going to re-solve the same problem using another method. He sets $\lambda = -1$ arbitrarily and rewrites out Helmholtz in 5.123. Apply an expo and Fourier-Transform (spatially) the whole equation 5.123 and it becomes the trivial p 56 A. I am used to seeing an ODE become a polynomial situation in k-space (here α-space). The solution is trivially 5.126.

Doing things the FT way, we are now faced with using the inversion formula to find our actual E_n function back in coordinate space. We are faced with a familiar-to-me looking integral 5.127. This integral in n=3 would be this one:

$$(2\pi)^{-3} \int_0^\infty dr\ r^2/(r^2+1)\,[2\pi \int_{-1}^{1} dz\,]\ e^{-irz} \qquad \text{where } z = \cos\theta \text{ and } \rho = r\sin\theta$$

But he does this in the general case, picking "z" as the last coordinate (not x_3 but α_n) and having "ρ" be the length of the remaining coordinates (not $\rho^2 = x^2+y^2$ but $\Sigma\alpha_i^2$). He ends up with an integral representation 5.128 which still has $\lambda = -1$. I have not done any details here! He then does a <u>scaling trick</u> to get this argument to be $-\mu^2$ instead of -1 and we get 5.129 for our "general" answer -- ie, we have computed the inversion integral. However, what we have is still an integral, albeit a 1D one. The main use of this mess is that it lets us derive a little <u>recursion formula</u> 5.130 for the E_n (which steps by two). Although we know E_2 as a K_0 function in 5.120, he prefers to keep it as the single integral representation from 5.129 (because this makes it easy to differentiate wrt r when we use that recursion relation). So we now have E_2 and E_3 as in p 57 A and B, and we can then <u>use the recursion rule to get ALL the E_n as shown top of page 58.</u> Fine. Actually, he uses the E_1 solution as the base for the odds in the first line of 5.131.

We can now compare the page 58 solutions obtained by the spatial FT method to the solution 5.119 obtained by the Bessel-equation method. The connection is that $\lambda = -\mu^2$ and $\mu = -i\sqrt{\lambda}$. For example, you can compare the first method's result for n=3 shown in 5.121 to the second method's result p 57 B.

So the upshot is that we know the Helmholtz radial solutions for any n (ie, in R_n) and for any λ (and here $\lambda = -\mu^2$ as on p 53).

I am still unclear why we are only looking at "spherically symmetric" solutions of these distributional ODE's. Certainly if we have spherically symmetric BC's that would be useful. But there has been no mention so far of any BC's whatsoever, that is certainly coming soon.

Note added in retrospect: in the development of the second method above, at one point we arrive at equation p 57 C which says $-\nabla^2 E_n + \mu^2 E_n = \delta(x)$ in R_n. It is this equation that we then solve to get the second-method odd and even solutions shown at the top of page 58. In the following sections, we are going to encounter this same equation several times with μ^2 being replaced by some other "constant". We can then just "steal" the solutions on page 58, replacing μ^2 appropriately.

(3) The Heat Equation (58).

In 5.132 we have the heat conduction equation where u is temperature and a is the thermal diffusion constant. This equation was derived in Appendix A of volume 1 pages 327-328. There, the general result is A.10 where c is specific heat and k is thermal conductivity, both functions of everything. If you assume c and k are both constants, then define a = k/c and you get A.11 which is our 5.321. Somehow the ∇^2 spreads out the heat, but I don't have a toy model for this process, but could come up with one I think.

So our Green's problem is then 5.133, but coefficients are constants so we can use 5.134 instead with p 58 A to get the full result (as we do many times).

Now comes the new element. We impose the condition that u = 0 for negative time. I suppose we could have used any constant for this, but u=0 is fine by me for simplicity. Not sure if this would be absolute zero or what in the physics model. Energy in a volume is c u. so I guess u = 0 means "no energy stored". No thermal energy that would be. Fine.

So we look for a solution of 5.135 which is H(t) in time. We now get the following interesting theorem proven before our very eyes.

Theorem: The solution of 5.135 (known as the "causal" fundamental solution) can be thought of as the solution of a slightly different problem, namely 5.136. This problem is the homo Lu(x,t)=0 with u(x,0) = δ(x). So here we imagine that we have a heat injection pulse δ(x), sort of a big bang of unity heat at the spatial origin, then things just progress from there. On the lower half of page 59, he proves that these two problems are related as he shows, so they have the same solutions. The main point is that you can identify δ(x) as u(x,0+).

Let's review this little theorem. u solves 5.136 and we define C by equation A. Fine. Then

$$\nabla^2 C = \nabla^2[\, H(t)\, u(x,t)] = H(t)\, \nabla^2 u \qquad \text{verifies B}$$

$$\partial_t C = \partial_t[\, H(t)\, u(x,t)] = H(t)\, \partial_t u(x,t) + u(x,t)\, \partial_t H(t) = H(t)\, \partial_t u(x,t) + u(x,t)\delta(t) \qquad // \text{ C}$$

Now why do we write the last term as u(x,0+)δ(t) ? I don't like it, I will put u(x,0)δ(t). Now assemble our two pieces to construct the LHS of 5.135,

$$\partial_t C - a\nabla^2 C = H(t)\, \partial_t u(x,t) + u(x,0)\delta(t) - a\, H(t)\, \nabla^2 u$$

$$= H(t)[\, \partial_t u(x,t) - a\nabla^2 u] + u(x,0)\delta(t)$$

But since u satisfies 5.136 for t > 0, the first term vanishes and we get

$$\partial_t C - a\nabla^2 C = u(x,0)\delta(t) \qquad \text{for } t > 0$$

But according to 5.136, we assume $u(x,0) = \delta(x)$ so we then have

$$\partial_t C - a\nabla^2 C = \delta(x)\delta(t)$$

Therefore, the solution to problem 5.136 generates, via equation A, the solution C to 5.135. Since the whole program here exists for $t > 0$, that is why we think of 0+ as the limit from the t>0 side, fine.

[p 60] So, what is the solution of 5.135? He does a spatial FT on our 5.135 problem, leaving time as it was, and we get a very simple result that $C_n\hat{}(\alpha,t) = \exp(-a\,\alpha^2\, t)$ as in 5.139, we are in k-space here, $\mathbf{k}=\boldsymbol{\alpha}$. Now of course we want to use the "inversion formula" (my "recovery" formula) to compute $C_n(x,t)$. As usual, n is the number of spatial dimensions. The inversion formula involves the standard shifted gaussian integral and we get as our final answer 5.140 which is this: $C_n(r,t) = (4\pi at)^{-n/2} \exp(-r^2/4at)$ which is spherically symmetric. You can see that as $t \to 0$, the expo localizes stuff at $r = 0$ and of course the limit of this thing as $t \to 0$ is going to be $\delta(x)$.[We could show this using our δ limit technology, but it has to be true according to 5.136.] Note in 5.137 the distinction between functions u and C. Function u(x,t) is not even talked about for $t < 0$, while $C = 0$ in this range. That is, $C = Hu$ as in p 59 A.

So we have solved the problem and found that basic Green's solution, analogous to $1/4\pi r$ for the Laplace equation for $n = 3$. But here, the result has this simple form for any n! Well, for Laplace, we had 5.111 for any n>2 which is also pretty simple. The Helmholtz solutions E_n were a little messier top of page 56.

Now let's ponder a method Stak *could* have used to solve this problem but does not use until the next section below. He could have started with 5.135 and done a time (only) Laplace Transform with variable s. I think we would end up then with $s\,\widetilde{C}_n(x,s) - a\,\nabla^2\widetilde{C}_n(x,s) = \delta(x)$ which we could write as

$$-\nabla^2\widetilde{C}_n(x,s) + (s/a)\,\widetilde{C}_n(x,s) = \delta(x)/a$$

This looks very similar to p 57 C and I am sure I could follow this through and come up with a solution to the heat equation which uses the page 58 top formulas. Of course those formulas would be for $\widetilde{C}_n(x,s)$ and then we would have to invert the Laplace transform to get $C_n(r,t)$. But we already have such a simple and general answer for $C_n(r,t)$, this is hardly worth doing!

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

Our Green's equation is stated in 5.142 with the driven wave equation in 5.141. Notice that 5.142 is exactly like the heat equation 5.135 except now we have a second time derivative instead of a first. The velocity parameter is now "c".

The claim is made and then proven that we can instead solve the "initial value problem" p 61A and get C indirectly that way, where $C = Hu$. The proof of this "theorem" fills the lower half of page 61. We now have two "initial conditions" instead of one as we had in the heat problem.

Last time we did a spatial FT. *This time* we are going to instead do a *temporal* Laplace Transform. Our initial value problem is restated in 5.145. (We set $c = 1$ and then restore it at the end.) I think of

$C_n(t=0)$ as the initial "position" of the wave, which is zero deflection everywhere, and $\partial C_n/\partial t(t=0)$ as the initial "velocity" of the wave, which is a velocity pulse concentrated at the origin of space. We are dropping a rock into the origin of a pond and we will watch the waves that result. Just after the rock hits the water, position = 0 still, but velocity has that $\delta(x)$ appearance.

Now the temporal LT gives us the simple equation 5.146 which is the same as equation p 57C which we already solved, but now $\mu = s$!! So we now know $\widetilde{C}_n(x,s)$ by cribbing the results from top of page 58, where over ~ means Laplace Transform in Stak notation. The cribbed results are shown in 5.147. All we have to do now is the inverse LT. Luckily, appearances of s are very localized in these odd and even solutions In fact, these functions are so simple we can just look up the inverse transforms in a table. For example, Schaum p 170 last two items are the ones we want! We then have results for $C_n(x,t)$ as shown in 5.148 and 5.149. He then quotes the specific results for n = 1,2,3.

You can see from 5.148 (n = odd = 3,5,...) that the entire wave is concentrated at the spherical surface r = t because you have nothing but derivatives of delta functions at this surface. However, 5.149 (even) shows a solution that has a wavefront at r = t, but has a "wake" behind it, just meaning that the inside of the time sphere has wave in it, whereas for n = 3,5... there is no wave inside the sphere. In n = 3, our 3D world, we get just an expanding sphere, but our rock in pond should be n = 2 with a finite wake.

Notice that none of these wave solutions are oscillatory! This is all impulse response stuff. And all this time we have c = 1, so the "wavefront sphere" moves out at velocity 1 meter/sec, say.

Now Stak wants to show that the solutions we have found really do solve the problem we started with. He should have pointed out this little fact: in the previous examples of Laplace, Helmholtz and Heat, although our Green's ODE had a distribution as a driving term (delta), the (spherically symmetric) solutions were always true functions, except perhaps at r = 0, and this was true for any n. True, in the heat case, we had Heaviside in time so we had a step function that was discontinuous, but that is only a mild distribution. In our wave case (no damping), the odd solutions are all pure distributional being derivatives of delta functions. So things are perhaps more interesting here.

So how do we show that the solutions really are "distributional solutions" ? His reading is that we should show all three results in 5.153. Here t is a parameter, and we are dealing with R_n *space* for our functional definitions. The first item is the homo equation of 5.145 for t > 0 (where he has already swung the ∇^2 to the RHS). And then the two t=0 BC's are written in distributionally sensible forms below that.

He does not attempt a complete verification for arbitrary n. Instead, he just does the cases n = 1 and then n = 3. I did not study each of these verifications but I don't think there is any rocket science here. I get the point.

His final "act" is to replace c = 1 with c = c, the velocity of the wave, in the right places, and we then have our odd and even solutions as in 5.155 and 5.156.

So far, so good!

(4) The Wave Equation (65) -- damped

So now we have γ linear term as shown in 5.157. We first transform from u to w with $u = w\, e^{-\gamma t}$ which throws our equation into the form 5.158. Notice that the driving term also got altered as in p 65 A. The point of doing this is to <u>remove the linear time derivative</u> in favor of a multiple of the function. We end up with the Green's problem shown in 5.159. As we did in the undamped case, we do a time Laplace transform so the second time derivative becomes s^2 and we have 5.160. Once again, we end up with our

famous page 57 C equation form $-\nabla^2 E_n + \mu^2 E_n = \delta(x)$ where now $\mu^2 = (s^2 - \gamma^2)$. So we crib the solutions from the top of page 58. The n = 1 solution is then shown in 5.161 where we set $\mu(s) = \sqrt{s^2 - \gamma^2}$.

At this point we have a long song and dance about the analytic structure of this function of s. He sets $\gamma = 1$ and will restore it later, so we then have $\mu(s) = \sqrt{s^2 - 1}$ which we know has branch points at s = -1 and s = 1. All he is saying on most of page 66 is that we want to put both branch cuts to the left, as we normally do with a square root. Then for s < -1 the two branch cuts cancel and there is no discontinuity, and the only discontinuity is in the range (-1,1). He shows this in *great* detail probably because the reader is not as familiar with this general idea as I am due to my history and the current date.

So now we want to do the Laplace inversion on our n=1 cribbed solution 5.161 and this is the vertical contour integral shown in 5.165. The real part a is > 1 in this case, since that is part of the transform rule in this case (go back to page 38 and ponder this if you want). If we close to the right in the upper page 67 drawing, we get $C_1(x,t) = 0$ and this is as expected for r > vt (v=1). The wavefront has not yet arrived. W then close to the left for r < vt and pick up the discontinuity across the cut. Thus, we arrive at result p 68 A and I have no complaints to this point. Now here is the discontinuity of the function shown. Let b be the absolute value of the square root $\sqrt{s^2 - 1}$ which is $b = \sqrt{1 - s^2}$. Then we have

below = $e^{st}\, e^{-(-ib)r}/2(-ib)$ above = $e^{st}\, e^{-(+ib)r}/2(+ib)$

below - above $= e^{st}\, e^{-(-ib)r}/2(-ib) - e^{st}\, e^{-(+ib)r}/2(+ib)$

$= e^{st}/(2ib)\{ -e^{ibr} - e^{-ibr} \} = -e^{st}/(2ib)\{ e^{ibr} + e^{-ibr} \}$

$= -e^{st}/(2ib)\ \{ C + iS + C - iS\} = -e^{st}/(2ib)\ \{ 2\cos(br)\} = -e^{st}\cos(br)/(ib)\ \}$

$= + i\, e^{st}\cos(br)/b$

This is the discontinuity shown in 68 A (below - above), so we then have

$C_1 = (1/2\pi) \int_{-1}^{1} ds\, e^{st}\cos(br)/b$ // agrees with 5.166

$C_1 = (1/2\pi) \int_{-1}^{1} ds\, e^{st}\cos(r\sqrt{1 - s^2})/\sqrt{1 - s^2}$

$= (1/2\pi)\{ \int_{-1}^{0} ds\, e^{st}\cos(br)/b + \int_{0}^{1} ds\, e^{st}\cos(br)/b \}$

$= (1/2\pi)\{ \int_{0}^{1} ds\, e^{-st}\cos(br)/b + \int_{0}^{1} ds\, e^{st}\cos(br)/b \}$

$= (1/\pi) \int_{0}^{1} ds\, ch(st)\cos(r\sqrt{1 - s^2})/\sqrt{1 - s^2}$

Now try $s = \cos\theta$ so that $\sqrt{1 - s^2} = \sin\theta$ then we have

$= (1/\pi) \int_0^{\pi/2} d\theta \, ch(t\cos\theta) \cos(r \sin\theta)$

Digression on an Integral.

Update: Jim found a relevant integral on GR p 510 which makes the delicate $\sqrt{t^2 - r^2}$ connection to the separated variables r and t. This GR integral says:

$$J_0(\sqrt{z^2 - Z^2}) = (2/\pi) \int_0^{\pi/2} d\theta \cos[z\sin\theta] \, ch[Z\cos\theta] = J_0(i\sqrt{Z^2 - z^2}) = I_0(\sqrt{Z^2 - z^2})$$

So I can translate this to say, using Z = t and z = r

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

I have no idea how you would derive the GR integral, however! I suspect is has another form:

$$J_0(\sqrt{a^2 + b^2}) = (2/\pi) \int_0^{\pi/2} d\theta \cos[a\sin\theta] \cos[b\cos\theta] \qquad \text{// yes, this is true}$$

There is an addition theorem Bateman p 101 with $\varphi = \pi/2$ which says this

$$J_0(\sqrt{a^2 + b^2}) = J_o(a)J_o(b) + 2 \Sigma_{n=1}^{\infty} J_n(a)J_n(b)\cos(n\pi/2)$$

$$= J_o(a)J_o(b) + 2 \Sigma_{n=2,4..}^{\infty} J_n(a)J_n(b)(-1)^{n/2}$$

$$= J_o(a)J_o(b) + 2 \Sigma_{k=1}^{\infty} J_{2k}(a)J_{2k}(b)(-1)^k$$

so somehow that cos cos integral must equal this thing. We could use these expansions from AS p 361

$\cos[a\sin\theta] = J_0(a) + 2\Sigma_{k=1} J_{2k}(a) \cos(2k\theta)$ " associated series"
$\cos[b\cos\theta] = J_0(b) + 2\Sigma_{k=1} (-1)^k J_{2k}(b) \cos(2k\theta)$

Jam these into our integral above to get

$$\int_0^{\pi/2} d\theta \cos[a\sin\theta] \cos[b\cos\theta] = J_0(a) J_0(b) \int_0^{\pi/2} d\theta$$

$$+ 2\Sigma_{k=1} J_{2k}(a) J_0(b) \int_0^{\pi/2} d\theta \cos(2k\theta)$$

$$+ 2\Sigma_{k=1} (-1)^k J_{2k}(b) J_0(a) \int_0^{\pi/2} d\theta \cos(2k\theta)$$

$$+\ 2\Sigma_{k=1}\ J_{2k}(a)\ 2\Sigma_{k'=1}\ (-1)^{k'}J_{2k'}(b)\ \int_0^{\pi/2} d\theta\ \cos(2k\theta)\ \cos(2k'\theta)$$

The required integrals are these

$$\int_0^{\pi/2} d\theta = \pi/2$$

$$\int_0^{\pi/2} d\theta\ \cos(2k\theta) = \sin(2k\theta)/2k|^{\pi/2} = \sin(k\pi)/2k\ = 0$$

$$\int_0^{\pi/2} d\theta\ \cos(2k\theta)\ \cos(2k'\theta) = (1/2)\ \int_0^{\pi/2} d\theta\ \{\ \cos(2(k+k')\theta) + \cos(2(k-k')\theta)\ \}$$

$$= (1/2)\ \int_0^{\pi/2} d\theta\ \{\ \cos(2k_+\theta) + \cos(2k_-\theta)\ \} = (1/2)\ \int_0^{\pi/2} d\theta\ \cos(2k_-\theta) = \pi/4\ \delta_{k,k'}$$

where we keep this last term because the difference of two k's could be 0! Thus our evaluated integral becomes

$$\int_0^{\pi/2} d\theta\ \cos[a\sin\theta]\ \cos[b\cos\theta] = J_0(a)\ J_0(b)(\pi/2)\ + 0 + 0 +$$

$$+\ 2\Sigma_{k=1}\ J_{2k}(a)\ 2\Sigma_{k'=1}\ (-1)^{k'}J_{2k'}(b)\ \pi/4\ \delta_{k,k'}$$

$$= (\pi/2)\{\ J_0(a)\ J_0(b) + 2\ \Sigma_{k=1}\ (-1)^k\ J_{2k}(a)\ J_{2k}(b)\ \}\quad =\quad (\pi/2)\ J_0(\sqrt{a^2 + b^2}\ )\qquad\qquad \text{QED}$$

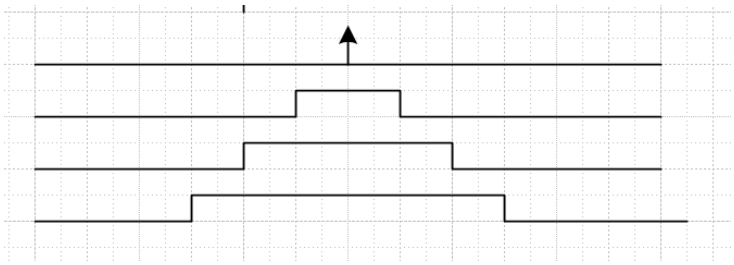
Lo and behold, this is the addition theorem which gives $J_0(\sqrt{a^2 + b^2}\ )$. So yes, Jim's first impression was exactly right, some kind of addition theorem lies behind this integral. A derivation of the addition theorem is given in Bateman Bessel p 43. The "associated series" are derived Bateman p 7.

Moving right along, I assume that 5.167 is the correct answer for C_1. To get C_3, we use the cribbed recursion formula 5.130 which is where p 68 B comes from ($1/2\pi r * 1/2 = 1/4\pi r$) . So then 5.168 is the specific solution for n = 3 and it involves a delta function and a wake. He says n = 2 is dealt with as an exercise later.

So far we had $\gamma = 1$. In the last section he shows a trick to get general γ back into the result. So the final solution to our Green's version of 5.157 is this:

$u(x,t) = e^{-\gamma t}\ C_n(r,t,\gamma)$ = as shown in p 68 C

[p 69] Here he compares our undamped versus damped solutions for a 1D string. These results do seem a bit strange to me. What would I expect a taut string to do if there is no damping? I guess the math at least says you just get this wave going down the string [strike, not pluck]



With damping, the outgoing wave has a softer shape as he shows in his p 69 picture, which is 5.169.

Question about those radial solutions.

We have some LE = δ in n-dimensions. Let's be specific and take $\nabla^2 E = \delta$. The point driving term here is definitely "at the origin". But let's replace δ = f(r) and consider $\nabla^2 E = f(r) = \nabla^2_\Omega E + \nabla^2_r E$. Our usual method is separation, so use $\nabla^2_\Omega = -L^2/r^2$ and our separated equation is $r^2(\nabla^2_r E - f(r)) = -r^2\nabla^2_\Omega E = L_\Omega^2 E$. We try $E(r,\Omega) = F_\ell(r)\,\psi_{\ell m}(\theta,\varphi)$ and obtain

$$r^2(\nabla^2_r E - f(r)) = L_\Omega^2 E$$
$$r^2(\nabla^2_r F_\ell(r)\,\psi_{\ell m}(\theta,\varphi) - f(r)) = L_\Omega^2 F(r)\,\psi_{\ell m}(\theta,\varphi)$$
$$r^2[\,\psi_{\ell m}(\theta,\varphi)\,\nabla^2_r F_\ell(r) - f(r)] = F_\ell(r)\,L_\Omega^2\psi_{\ell m}(\theta,\varphi)$$

$$r^2[\,\nabla^2_r F_\ell(r) - f(r)/\psi_{\ell m}(\theta,\varphi)\,]/F_\ell(r) = L_\Omega^2\{\,\psi_{\ell m}(\theta,\varphi)/\psi_{\ell m}(\theta,\varphi)\,\} = \ell(\ell+1)$$

But this does NOT separate due to the presence of the inhomo f(r) term! In the last equation above, no matter what function $F_\ell(r)$ is, the LHS is a function of angles and the RHS is not, so there is no solution. The only solution is if $\ell = 0$.

Theorem: The only solutions of the equation LE = δ are spherically symmetric. [But if you include the effect of boundary conditions, non-symmetric boundary conditions destroy the conclusion here.]

I guess I could do an operator proof of this for n = 3 say. We have

$$L_r E(\mathbf{r}) = f(\mathbf{r})$$

$$RL_rR^{-1}RE = Rf(\mathbf{r}) \quad \Rightarrow \quad L_{r'}E(\mathbf{r}') = f(\mathbf{r}') \quad \text{in a rotated system}$$

where $\mathbf{r}' = R(\mathbf{r})$ in the usual sense. But if f(r) is symmetric, then $f(\mathbf{r}) = f(\mathbf{r}') = g(r)$ and then

$$L_{r'}E(\mathbf{r}') = L_rE(\mathbf{r})$$

Now if in addition L_r is a rotational scalar like ∇^2, then we also have $L_{r'} = L_r$ and then

$$L_rE(\mathbf{r}') = L_rE(\mathbf{r}) \quad \Rightarrow \quad L_r\{\,E(\mathbf{r}') - E(\mathbf{r})\,\} = 0 \quad \text{suggests } E(\mathbf{r}') - E(\mathbf{r}) = G(r).$$

Exercises (69-72)

Today is July 11, 2009 and I have just finished bringing my meta notes current to this point. I reached page 69 on June 27, so a full 2 weeks have passed (including a Torrey trip) during which all I can claim to have done is meta notes for this huge Chapter 5 (up to page 69 of course). I have read through this set of exercises and they all concern finding more of these Green's function type solutions (spherical) for different operators L. I know it would be good to do these exercises, but I have lost so much momentum that I instead want to try and finish out this chapter and perhaps go into Chapter 6. At some future time (and it often works out that way) I will come back and tackle some of these problems. Stak has an interesting comment in his Exercise 5.37, inciting the reader to make an attempt. But not today.

5.9 Classification of PDEs (72)

This is a subject I have long been interested in, and finally after 30+ years maybe I will learn something about it. I am going to alter the presentation order here a bit.

The Cauchy Problem. (73)

Consider Lu = f where L is a full bore order p LPD operator with continuous a_k coefficient functions. Stak skips these words but I will add them: you are considering Lu=f in some region of R_n called R. Imagine that there is some hypersurface σ of dimension n-1 within R. Imagine that the solution u and all its derivatives (up to order p-1 but not p) could be "specified" on this surface σ (not necessarily a closed surface). This specification is called "**the initial data**", as if it were perhaps at time t = 0 in a time evolution situation. Issues that come to mind: (1) if you specified this information arbitrarily, it seems likely that it might somehow violate Lu=f in the region of the surface. Ie, not any spec of u on σ would be "legal". (2) If it were legal, would this then "force" (determine completely) the solution in all or part of region R?

Before facing these issues, Stak claims that this initial data can be specified more compactly and perhaps more consistently as "**the Cauchy data on σ**". Think of u(**x**) as a scalar field in R_n and look in the neighborhood of the surface σ. At some point P on the surface there are many derivatives of u that "exist" at point P. The idea is to do what we do in general relativity: zero in on the point P and notice that since σ is assumed to be smooth, there is a "tangent plane" which touches σ at point P. In n = 3, this is really a plane, but in n = 4, it is a plane with 3 dimensions, and so on . In n = 2, the plane is a line. Very close to P, this tangent plane aligns perfectly with the surface σ. The idea then is to form a set of local coordinates which we call $\{ \lambda, \xi_2, \xi_3 \ldots \xi_n \}$. This set of coordinates has its origin at point P on σ. There is exactly one "normal coordinate" which we call λ, and all the remaining n-1 coordinates lie in the tangent plane. I think this means that **n**•**v**= 0 if **n** is in the λ direction and v is some vector lying in the tangent plane. This set of coordinates is obviously related in some linear fashion to some standard set of coordinates $\{x_1, \ldots .x_n\}$ that you might have for your space R_n . The relation is by some rotation and translation I presume. I also presume that the ξ_k are orthonormal, but I don't think he requires that to be true.

Now, in this framework, we can classify the many terms which appear in (the transformed) Lu. Those which are of the pure form $a_{\lambda 00\ldots 0}(\mathbf{x})\, \partial_\lambda^n u$ are derivatives perpendicular to the tangent plane. If you were

to plot the function $u(\lambda,\xi_2....)$ versus λ along a short line segment which was perp to (and penetrates) the tangent plane at P, you get some curve of u versus λ which we could plot on a 2D graph. At the point P, this curve has slope $\partial_\lambda u$ and it has curvature $\partial^2_\lambda u$ and so on. All these derivatives are called **normal derivatives**, since they are normal to the surface σ.

All other derivatives -- those which involve at least one of the ξ_i coordinates -- are called **tangential derivatives**. They need not be "pure" the way the normal ones are. For example, $\partial^2_\lambda \partial_{\xi 3}$ would be called a tangential derivative.

Now here is the claim. If you know the function u AND all its normal derivatives to order p-1 on your surface σ, then you also know all the tangential derivatives as well. Why is this? Imagine you are at some point P. You know everything about function u and all $\partial_\lambda{}^n u$ in a little region of σ surrounding P, because we just said we know u and its normal derivatives on all of σ. Thus, you can compute any tangential derivative directly because we never have to leave the surface σ to do so. For example, $\partial_{\xi 3}\partial_\lambda u = [\partial_\lambda u(\xi 3+\Delta\xi 3)- \partial_\lambda u(\xi 3)]/ \Delta\xi 3$. One thing we can then say is that we can compute ALL derivatives up to order p - 1. But of course we can compute things like $\partial_{\xi 3}{}^{126}\, \partial_\lambda u$ as well. That is, as long as the part relating to λ is order p-1 or less, we can compute up to infinite order on any of the ξ parts!

So here is the point. If you know the function u AND all its normal derivatives to order p-1 on your surface σ, then you know the function u and ALL of its derivatives (normal and tangential) to order p-1, and when you transform back to your x_i coordinates, you then know all x_i space derivatives to order p-1. This limited amount of specification data (u and its normal derivatives up to order p-1) is known as **the Cauchy data on σ.** If you are given $Lu = f$ in some region R of R_n, and if you are given some Cauchy data on a surface σ within R, you are facing **The Cauchy Problem**. This, then, is how we talk about "boundary values" in the PDE world, though Stak has not used the phrase "boundary values" at all.

So this bears directly on our concern (1) mentioned above. If we specify the Cauchy data on σ, then that determines all the other derivative data so we definitely don't want to try to specify the tangential derivatives in addition to the Cauchy data. The fact is that such extra data *cannot* be specified in addition to the Cauchy data since it is completely *determined by* the Cauchy data.

We are left then with the idea that maybe we are free to specify the Cauchy data arbitrarily on σ, but we don't really know yet. It is a possibility.

So this takes us to the horizontal pencil line on page 74. Let's now back up to consider the introduction.

Introduction (72)

Consider the order p = 1 equation $Lu = a_0\partial_x u + a_1 u = f$ in the space R_1. From our experience, we know that if you specify $u(x_0)$ at some single point x_0 in your range (a,b), that completely determines the solution u everywhere in the region. In R_1 a hypersurface is a point x_0 and there is no "room" for any normal or tangential derivatives. So in this case $u(x_0)$ is the entire Cauchy data! The solution is unique.

Stak then shows exactly how a computer could construct the solution u(x). We start at x_0 where we know $u(x_0)$. The ODE lets us compute $\partial_x u(x)|^{x=x0}$ so we know that too (provided $a_0 \neq 0$!!). He shows this in 5.173. We then do a first order computation of $u(x_1) = u(x_0) + \partial_x u(x)|^{x=x0}\, \Delta x$. If the steps are small enough, you can iterate and compute u(x) in the range x_0 to b, and if you go backwards, then from x_0 back to a.

If $a_o(x_o) = 0$, the method fails because we cannot compute $\partial_x u(x)|^{x=x0}$. This derivative could be infinite (no existe) or just impossible to determine (indeterminate). I guess this would be a problem anywhere along your iteration, which is why we need $a_o(x) \neq 0$ on the entire interval (a,b).

Now Stak makes this subtle comment for our example. (1) For $Lu = a_0\partial_x u + a_1 u = f$, we can compute $Lu|^{x0}$ directly from this ODE -- it is just $f(x_o)$. (2) If $a_o(x_o) = 0$, then $Lu = a_1 u$ and we can alternatively compute Lu from $u(x_0)$, it is just $Lu = a_1(x_o)u(x_0)$. But these two computations must agree if u is really a solution to Lu = f, and therefore in the case $a_o(x_o) = 0$, you *cannot* arbitrarily specify $u(x_o)$! If you pick some arbitrary $u(x_o)$ value, then it is likely there will be no solution at all!

Now how do we generalize the above comments to ODE L of order p ? When we try to do our computer iteration in this case, we need to iterate not just u, but some of its derivatives.

$$u(x_1) = u(x_0) + \partial_x u(x_0)\,\Delta x$$
$$\partial_x u(x_1) = \partial_x u(x_0) + \partial_x^2 u(x_0)\,\Delta x$$
$$\partial^2_x u(x_1) = \partial^2_x u(x_0) + \partial_x^3 u(x_0)\,\Delta x$$
$$\ldots$$
$$\partial^{p-1}_x u(x_1) = \partial^{p-1}_x u(x_0) + \partial_x^p u(x_0)\Delta x$$

Why do we need all this information? Look at the next step of the iteration of the first line:

$$u(x_2) = u(x_1) + \partial_x u(x_1)\,\Delta x$$

So, in order to move to x_2 , we need $\partial_x u(x_1)$, and this is why we need the *second* line in our group above. Now let's do the next step:

$$u(x_3) = u(x_2) + \partial_x u(x_2)\,\Delta x$$

Now we need $\partial_x u(x_2)$ and we compute that from the iterated second line

$$\partial_x u(x_2) = \partial_x u(x_1) + \partial_x^2 u(x_1)\,\Delta x$$

but now we see that we need $\partial_x^2 u(x_1)$, and this is why we need the *third* line in our group above.
Now do the next step:

$$u(x_4) = u(x_3) + \partial_x u(x_3)\,\Delta x$$

which says we need this:

$$\partial_x u(x_3) = \partial_x u(x_2) + \partial_x^2 u(x_2)\,\Delta x \qquad \text{// from second line in group, iterated twice}$$

but now we need $\partial_x^2 u(x_2)$ which comes from the third line iterated once,

$$\partial^2_x u(x_2) = \partial^2_x u(x_1) + \partial_x^3 u(x_1)\,\Delta x$$

But to get this, we need to know $\partial_x{}^3 u(x_1)$ which comes from the *fourth* line, and this is why we need the *fourth* line in our group above.

When does this process end? If it fails to end, then we have to track an infinite number of quantities in each step, and that would take forever to do. The way it is supposed to end is this. We have

$$Lu = a_o \partial^p{}_x u + L'u = f \qquad \Rightarrow \qquad \partial^p{}_x u(x_0) = [\, f(x_0) - L'u\,(x_0)\,] \,/\, a_o(x_0)$$

As long as $a_o(x_0) \neq 0$, we have our Cauchy data giving all the derivatives at x_0 up to order p-1, then the ODE itself tells us the derivative of order p (as just shown above), sort of closing the loop. By differentiating the ODE once, we could then compute $\partial^{p+1}{}_x u(x_0)$ and so on. Thus, all the higher derivatives are completely determined, so our "vector" of iterating quantities needs only go up to p-1 and that is why the Cauchy data needs only go to this point. Notice our last group equation above

$$\partial^{p-1}{}_x\, u(x_1) = \partial^{p-1}{}_x\, u(x_0) + \partial_x{}^p u(x_0)\Delta x$$

The ODE has provided us with $\partial_x{}^p u(x_0)$ which is required here, so we can compute our vector through p-1.

In this ODE order p example we have that same conclusion: If $a_o(x_0) = 0$, we can compute Lu both from the ODE ($Lu = f(x_o)$) and from just the Cauchy data (L'u) , so you cannot specify the Cauchy data arbitrarily! Specing such data would likely cause there to be no solution to Lu=f.

We now resume the discussion on page 74 below the pencil line.

In our examples above we saw that "we have a problem Houston" when we can compute Lu directly from the Cauchy data alone, since we are likely to get conflicting results by computing it as Lu = f(P) where P is some point on our "initial data hypersurface σ" which was just point x_0 in those examples. If for some σ we have this problem at some point P on σ, then that surface σ is called "**characteristic for L at P**" and one says that "**L is an interior differential operator at P with respect to σ**". Stak gives us no comment as to the significance of the word "interior", being insensitive to the curious normal student. So fine, these are just definitions.

Example: Work in R_2 where coordinates are x_1 and x_2 so any candidate hypersurface σ is just a curve C in the x_1 x_2 plane. Suppose we select a p=1 L. Then the Cauchy data is just u on the curve C. Because p = 1, there are no normal derivatives in the Cauchy data, p - 1 = 0. Of course there is still a normal derivative at each point P on this σ we call C. Specifically now, let $L = \partial_{x1}$. Suppose the curve C has a point P on it where slope = 0. Then at that point P, we would be able to compute Lu just from the Cauchy data, since the curve C lies along the x_1 direction at P, since $L = \partial_{x1}$, and since we know u on the curve!!! So we expect this to be a problem situation. For such a curve C, C would be characteristic for L at P, we get to try out our new terminology. But now he throws in a new term:

Definition: a "characteristic curve" for $L = \partial_{x1}$ is a horizontal line. Any surface σ (here C) which matches (tangent wise) a characteristic curve at some point P is going to be a problem, will be characteristic at P for L. so I guess more generally, a "characteristic surface" is one which, if your arbitrary surface σ matches it at some point P, you will have a problem in that you can compute Lu from the Cauchy data at point P.

Fact: right now I see no connection between this use of the word characteristic, and the line $x = \pm t + K$ we encountered dealing with the 1D wave operator. There, the characteristic was a surface which could maintain a normal discontinuity of u.

We now come to **Theorem 1** on page 75. I agree that (1) and (2) are really the same by definition. The interesting new claim is (3) which says that if you are at a problem point P, then "the coefficient of $\partial_{\lambda}^{p}u$" in Lu vanishes at P. As I read the proof of this claim, I agree that 5.175 says this: "if b(P) = 0, then L is interior at P". I agree that L' is interior no matter what, by its definition. But the last sentence on this page -- stated right at the critical point in the proof -- makes no sense to me:

"If L is interior, then since $\partial_{\lambda}^{p}u$ is unknown from the Cauchy data (that is true), b(P) must vanish. "

I think this sentence does make sense:

"If L is interior at P, b(P) must vanish and vice versa."

It is true that if $\partial_{\lambda}^{p}u$ happened to be 0 at P, then L would be interior even if $b(P) \neq 0$.

So we have shown that (3a) if b(P) = 0, then L is interior; (3b) if L is interior, then b(P) = 0 unless by some coincidence $\partial_{\lambda}^{p}u$ happened to be 0 . So these are the two proof directions.

Theorem 2: If P is a problem point, then knowing the Cauchy data does *not* let you compute ∂_{λ}^{p} at P.

Well a problem point occurs when b(P) = 0, and then this theorem is obvious from what we just stated. You cannot divide through by b(P) and get $\partial_{\lambda}^{p}u$ from the DE.

Theorem 3: If P is *not* a problem point, then you *can* compute $\partial_{\lambda}^{p}u$ at P from the Cauchy data.

OK, since $b(P) \neq 0$, divide through and you have it. These theorems seem pretty empty.

Theorem 4: If P is not a problem point, then since we can compute $\partial_{\lambda}^{p}u$ at P by Theorem 3, when we transform back to the x_i coordinates, we will know all derivatives through order p at point P if we know the Cauchy data at point P. (These theorems are all specific to operator L of order p and Lu = f)

We would like to have a Theorem 5 which states: "if P is *not* a problem point on σ, then the Cauchy data does in fact yield a unique solution in the neighborhood of P, since we know all derivatives through order p". This would address my concern (2) stated right at the start of this section: Does the Cauchy data support a unique solution to Lu=f in any sense? (The way position and velocity determines a unique solution of a second order ODE.) The answer is that the Cauchy data may or may not support a solution! Stakgold apologizes for this sad situation. All our theorems are "negative" in that they tell us situations where we know there won't be a solution near P, but we don't have a positive theorem 5. We then have to "abandon our general theory" and just study specific but important and useful examples. That is what the rest of this volume is going to do!

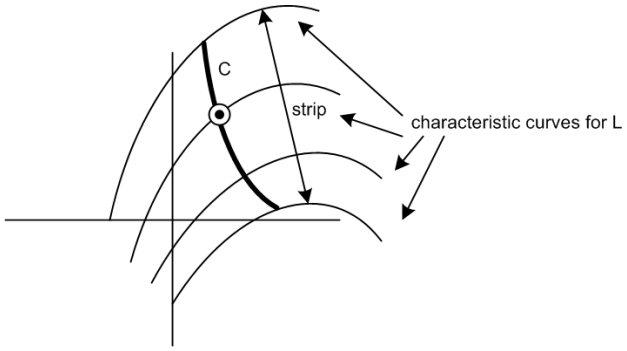
First order PDE's

We open with $L = \partial_x$ and $Lu=0$ which has solution $u(x,y) = F(y)$ for any F. Our space is R_2 which we now draw as (x,y). In Case 1, our curve C is a vertical line segment. Our Cauchy data is just u on C. We know that u(x, y) = u(x-on-C, y) which just means u is the same on any horizontal line. In both Cases 1 and 2, our curve specifies only one u for each horizontal line, so we are OK. In Case 3. the initial data makes no sense unless u = constant on the horizontal segment shown. In this case, all we know is that this is the value of u on the entire horizontal line, we know nothing else! In Case 4, for each horizontal line with two intersections on C we would have to have u(I) = u(II) or we are inconsistent, so pretty clearly we cannot arbitrarily specify u on C. These are just simple examples to nail down the general discussion above. In this Case 4, our C has just one "problem point" P, but that is enough to put the kibosh on the entire Cauchy problem with this C.

Cauchy problem for first order PDE (76)

Theorem 4 said that knowing all derivatives at P through order p (here order 1) means P is not a problem point. Now suppose P and Q are two neighboring points on C separated by dx,dy. Can we compute our two derivatives $\partial_x u$ and $\partial_y u$ at P? We write the two equations A and B on page 79 which are trivially true, and we try to solve them using Cramer's Rule for these derivatives. We find that we can solve them *unless* b dx = a dy, where a and b are coefficients in $L = a\partial_x + b\partial_y$ + rest. This means that at any point P on C, there is a certain "slope" dy/dx = a(P)/b(P) which would be a problem. As long as we avoid the "bad slope" (the characteristic) at each P on C, we will avoid the "bad P problem". Now when someone hands you dy/dx = f(x,y), you can write dy = f(x,y)dx. If we start at any point P in the x-y plane, if we move dx, this tells us how much we should move in y (dy) to stay on the curve implied by dy/dx = f(x,y) and passing through point P. If f(x,y) is separable, I can do the integration manually, but if now, a computer could do it. For each point P in the plane, this will give you a curve. So, there is some set of characteristic curves for this problem in the x-y plane, some sort of topo map thing, but no example is given. If a and b are constants, then dy = b/a dx and the characteristic curves are all straight lines with slope b/a in the x-y plane. So the characteristic curves are a set of parallel lines of slope b/a. Fine!

Now comes an unproven theorem which requires a picture:



Imagine a C which is not tangent to any of the characteristic curves. This means we don't have any "problem points" on C. One says that C is "nowhere characteristic". The unproven theorem states that the Cauchy problem wherein c is specified on this curve C (hypersurface σ) has a unique solution within the strip bounded by the two extremal characteristic curves. You can see how this is just a "warped version" of our example where $L = \partial_x$ and the characteristic curves were horizontal lines. That is to say, a warped

version of Case 1 or 2, take your pick. The theorem also says: if our specified u values are discontinuous along C, say at the point I have marked with a dot, then that discontinuity propagates into the solution along the characteristic curve which passes through the dot!

If we look back now at page 47 Exercise 5.24, we found that $L = \partial_x + \partial_y$ with Lu=0 did in fact support generalized solutions! We showed there that H(x-t) was such a solution. This has a discontinuity along the characteristic x = t. So what we are finding here is that, in general, any first order PDE is capable of supporting a discontinuity "out in the middle of nowhere" which runs along a characteristic curve. To realize such a solution, just make your Cauchy data for u(P) have a discontinuity at some point P. On the other hand, recall that $L = \partial_x^2 + \partial_y^2$ cannot support such a discontinuity.

Cauchy problem for second order PDE (79)

Stay in R_2 with x,y and write out the most general 2nd order PDE as in 5.182. It is the three coefficients of the 2nd order derivatives a,2b and c that are our main interest. We know that P is "not a problem point" (C is not characteristic at P) if we can compute all (three) second derivatives at P. We then write 3 equations mimicking what we did in the first order case, as shown A,B,C page 80. If the det shown is not zero, then we can solve for the three derivatives and P is not a problem point. The det condition can be written as in 5.184. So this is the equation for our new set of characteristic curves for this problem! There are three obvious cases at a given point P:

The square root argument is negative, so dy/dx is complex, so has no real solution, so we never have to worry about P being a characteristic point. In fact, there are no characteristic curves at all through such a point P. This is the case for Laplace, since a = c = 1 and b = 0 so b^2 -ac = -1. Since a,b,c are constants, the conclusion applies at *all* points P: there are no characteristic curves anywhere in the x-y plane. Thus, you can never propagate out a discontinuity by having a discontinuity on C. This agrees with the fact we found earlier that the Laplacian cannot support a discontinuity out in free space. This is the **elliptic** case, though we don't see any ellipses nearby.

The square root argument is positive. Then every point P has not one, but two dangerous slopes dy/dx to worry about. This suggests to me that there would then be two sets of characteristic curves you would want to plot, and you want to choose a C that avoids being tangent to either set. This is the **hyperbolic** case. The 1D wave equation has a = 1 and c = -1 so b^2 - ac = +1 so that case goes here. The space is of course not x-y but t-x in this case.

Finally, if the square root is exactly 0 at a point P, there is only one characteristic at that point, and this is the **parabolic** case.

Comment on the names. If you have a general quadratic equation in two variables:

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0$$

then the nature of the curve this equation represents is determined by $B^2 - 4AC$ which, it happens, is a rotational invariant under $R_z(\theta)$. The curves are these:

$B^2 - 4AC < 0$ => ellipse

$B^2 - 4AC = 0$ => parabola
$B^2 - 4AC > 0$ => hyperbola

Strangely, this information does not appear in any book I own!

Stak has used 2b in 5.182 just to get rid of these factors of 4. So in our R_2 PDE world we could have

$$A\partial_x^2 + B\partial_{xy} + C\partial_y^2 + D\partial_x + E\partial_y + F = 0$$

and then we have a little analogy with the conic section equation. That is where the names come from, but I don't think there is any further significance!

So everything is looking good now, we are on page 81 and cooking with gas.

Examples (81)

Laplace and Helmholtz are indeed elliptic.

The 1D wave is hyperbolic. We have in that case

$dx/dt = \pm c$ => $dx = \pm c\,dt$ $x = \pm ct + \text{constant}$ = characteristic curves

so the two families of characteristic curves are a trellis of lines of slope ±c.

The heat equation is parabolic.

We can always find the characteristic curves from 5.184.

He claims at this point that our Cauchy Problem approach works well with hyperbolic equations, and we then get two big examples which are in fact related by a coordinate transformation.

Hyperbolic Example 1: $L = \partial_{xy}$ and $Lu = 0$ (81)

(1) The Cauchy Problem.

Derive p 82A. The path going horizontally between Q and P has dy = 0, so $\partial u/\partial y$ makes no contribution, and we have $du = \partial u(x,y)/\partial x\, dx$ only. Integrate and that is what p 82A says. This has nothing to do with the "differential equation" despite the text here, there is a semicolon after 82A. The DE does tell us that $\partial u/dx$ does not vary with y, so $\partial u/\partial x = f(x)$ only. If we were to change the contour in 82 A to be "along C", from Q to R, we have exactly the same values of $\partial u/\partial x$ being integrated since they don't depend on y. (We could integrate over any path between our two x values, but we like the path on C because the Cauchy data + PDE + no problem points says we know u, $\partial u/\partial x$, and $\partial u/\partial y$ at every point on C.) Now to get 82 B he just imagines that $y = \psi(x)$ is the invertible equation for our curve C and result 82C then also follows. But 82 C gives us the answer to our problem because everything on the RHS is given by the Cauchy data (and we are given the curve C as ψ). The characteristics are vertical and horizontal lines, and we avoid with our C being tangent to any of these characteristic curves.

To summarize: we are given the PDE $\partial_{xy}u = 0$, and we are given some Cauchy data on the curve C shown. We are free to specify u and $\partial u/\partial\lambda$ along this curve (that is the Cauchy data). From this we have to go off and compute $\partial u/\partial x$ along the curve, something he has not pointed out, but I think I could do that.

Well, this last subject is addressed a bit bottom of page 82. We can talk about coordinates (s,n=λ) at a point instead of (x,y). Then ds is along C, and dn is perp to C. I agree with p 82E. This let's us use the official Cauchy data $\partial u/\partial\lambda = \partial u/\partial n$ on C, but we have to still compute $\partial u/\partial s$ along C, which we can certainly do since we know u along C.

If C is a slope 1 line segment, I agree with 5.186. Now all we need is u at the two segment endpoints and the Cauchy $\partial u/\partial n$ along the segment. I suspect we shall see this sometime in the future. This is just another way to write our known answer to this problem!

(2) The section just reminds us that if C is tangent to a vertical or horizontal line, we cannot specify arbitrary Cauchy data and the problem is inconsistent.

(3) Here Stak demonstrates a completely different kind of BC situation. Here we don't supply "the Cauchy data" on some C (meaning u and $\partial u/\partial\lambda$). Instead, we supply just some u data on 0A and 0B. So this is NOT a Cauchy problem treatment. The initial data is f(x) and g(y), general solution is F(x) + F(y). After some simple algebra, we end up with this conclusion

u(P) = u(R) + u(Q) - u(O) ie u(x,y) = g(y) + f(x) - g(0) IS the solution.

see figure page 83. The thing u(P) is out in the middle somewhere, the thing we want a formula for. The other three items are part of the specified data! We have to require that f(0) = g(0).

(4) We get exactly the same solution u(P) = u(R) + u(Q) - u(O) is we warp the vertical OB specification segment into a curve C, except the point O is now as shown in the page 84 figure.

Conclusion: the problem $\partial_x\partial_y u = 0$ is a pretty simple hyperbolic roblem, we have several different ways to deal with the boundary conditions, one of which is the Cauchy data method. We do have to be a little careful of our curve C when we want to use one as part of the BC's.

Hyperbolic Example 2: $L = \partial_t^2 - c^2\partial_x^2$ and Lu = 0 (starts p 84)

This is our famous 1D wave equation. I have confirmed that the little transform to (η,ξ) which he shows does convert this into a constant times $\partial_\eta\partial_\xi$. Under this transform, the vertical and horizontal characteristics in the η,ξ space map into the diagonal lines which appear in rest of his pictures. Of course we know this directly since $dy/dx = (B \pm \sqrt{B^2-4AC})/2A$ and now B = 0, $A = -c^2$ and C = 1 if we associate our traditional y with t. Then we have $dy/dx = \pm\sqrt{4c^2}/(-2c^2) = \pm 1/c$

(1) Method #1. Stak quotes without derivation a little geometric solution equation p 85 A which says that u(x,y) out in space is given in terms of two values of u along the t=0 horizontal boundary, AND an integral of the "velocity" $\partial u/\partial t$ between those two points. This result is highly non-obvious to me, but it is very similar to what we saw in 5.186 and there is that direct connection to $\partial_\eta\partial_\xi$, so at least I can see that equation (5.189) follows at once because we are just renaming things. So here we are specifying certain

initial information which is the wave shape u(x,t) at t = 0 between x-ct and x+ct, and also the wave velocity in that same range of x for t = 0. Notice that, for a given x, as time moves on, we need more initial data to be able to have our answer. This is because Q and R move farther apart. He calls this the d'Alembert Formula and I am sure I have never seen this before!

The wiki page http://en.wikipedia.org/wiki/D'Alembert's_formula has a pretty direct derivation and mentions this transformation connection.

d'Alembert's formula

From Wikipedia, the free encyclopedia

In mathematics, and specifically partial differential equations, **d´Alembert's formula** is the general solution to the one-dimensional wave equation:

$$u_{tt} - c^2 u_{xx} = 0,\ u(x,0) = g(x),\ u_t(x,0) = h(x),$$

for $-\infty < x < \infty,\ t > 0$. It is named after the mathematician Jean le Rond d'Alembert.

The characteristics of the PDE are $x \pm ct = \text{const}$, so use the change of variables $\mu = x + ct, \eta = x - ct$ to transform the PDE to $u_{\mu\eta} = 0$. The general solution of this PDE is $u(\mu,\eta) = F(\mu) + G(\eta)$ where F and G are C^1 functions. Back in x, t coordinates,

$$u(x,t) = F(x+ct) + G(x-ct)$$

u is C^2 if F and G are C^2.

This solution u can be interpreted as two waves with constant velocity c moving in opposite directions along the x-axis.

Now consider this solution with the Cauchy data $u(x,0) = g(x), u_t(x,0) = h(x)$.

Using $u(x,0) = g(x)$ we get $F(x) + G(x) = g(x)$.

Using $u_t(x,0) = h(x)$ we get $cF'(x) - cG'(x) = h(x)$.

Integrate the last equation to get

$$cF(x) - cG(x) = \int_{-\infty}^{x} h(\xi)d\xi + c_1$$

Now solve this system of equations to get

$$F(x) = \frac{-1}{2c}\left(-cg(x) - \left(\int_{-\infty}^{x} h(\xi)d\xi + c_1\right)\right)$$

$$G(x) = \frac{-1}{2c}\left(-cg(x) + \left(\int_{-\infty}^{x} h(\xi)d\xi + c_1\right)\right)$$

Now, using

$$u(x,t) = F(x+ct) + G(x-ct)$$

d´Alembert's formula becomes:

$$u(x,t) = \frac{1}{2}\left[g(x-ct) + g(x+ct)\right] + \frac{1}{2c}\int_{x-ct}^{x+ct} h(\xi)d\xi$$

Once again: this last formula is a statement of the solution of Lu=0 if you are given u and u_t data at t = 0. Suppose the initial velocity profile h is 0 and the initial shape is g(x) at t=0. Imagine a horizontal string with transverse coordinate y = u(x,t) and longitudinal coordinate x, and you somehow create some initial pulse shape at t = 0, but all pieces of the little pulse part of the string are vertically at rest. You have some

sort of jig form that holds the pulse shape and you just release the jig at t = 0. The wave equation does not allow a pulse to just sit there, it has to move at c. What happens is that it resolves itself into two waves of half the amplitude, and one goes off to the left and the other to the right, both waves going a velocity c.

(2) Method #2. (p 86). Here we consider a string tied at both ends, but allowed to have an arbitrary shape and velocity profile at t = 0. Here Stak illustrates a method known as "**the method of characteristics**" which is certainly a bit obscure. The x-t space is shown in the page 86 figure. Using the method of the previous page, our initial data gives us a solution anywhere in region I, a triangle.

How can we get the solution region II? Well, we know the solution along AC with I, and just on the other side in II it will be the same apart from a possible jump constant which, recall, you can have across a characteristic if it is present along curve C. In this case, to have a jump you would have to have the string have a non-zero deflection at x = 0+, which I think can only happen in some kind of fancy distributional solution of this problem, but he is good to include the jump possibility.

So relative to region II, we know u along characteristic AC, and we know that u=0 along the left vertical edge (a noncharacteristic), so we use the little method of the figure on page 84, adapted to our situation in region II.

So the same thing to get region III.

Then region IV is like the figure on page 83 where you have u on two characteristics. Again you might have jumps.

In this way, you build up your solution in all regions.

Somehow, I can sense this is related to waves bouncing back and forth as time moves forward, in some obscure way.

So this final section ties together all the previous little methods we discussed considering our two hyperbolic examples.

Stak has not really explained why Cauchy is so useful for hyperbolics and not for elliptics. It is true that the elliptics have no characteristics, so hard to use any method that uses "characteristic curves".

And so ends this incredibly painful and long Chapter 5, today is 7.13.09.