Stakgold Chapter 1 Raw Notes PhL 3.30.09

Paren numbers are page numbers in the book.

Appendix A.1: The strings (and membrane in part (d)). (323) **2**

1.1 The string under pressure f(x) (1) **3**

The Green's Function for the String with Fixed end points (5) 4

Uses of the Green's Function for the String (10) 5

Side topic: some calculus definitions: C^k etc. 5

Order interchange theorem for differentiation and integration. 7

Theorem about differentiating an integral 8

Proof by "careful differentiation" that $u(x) = \int dy\ g(x,y)f(y)$ solves $Lu = f$. 10

Notion that L and G are inverses. 11

Exercises: (14) 12

1.2 The Dirac Delta Function (18) **12**

δ sequences (19) 12

Exercises (21) 12

The sifting property (22) 12

Exercises: (24) 12

1.3 The Theory of Distributions (28) **13**

History 13

My own introductory comments. 13

Functions Viewed as Linear Functionals (28) 14

Test functions (29) 14

Convergence in the Space of Test Functions (30) 15

Distributions (31) 15

Locally Integrable (LI) functions (31) 15

Examples of Distributions (34) 17

Operations with Distributions (35) 18

1. Translation. (35) 18

2. Scaling (36) 19

3. Multiply by a LI function (36) 19

Equality of two symbolic functions (37) 21

4. Differentiation (37) 21

Examples of derivatives of distributions (38) 21

Exercises (40) 22

Convergence of Distributions (42) 22

Examples of the Convergence of Distributions (43-47) 24

Exercises (47-51) 26

Differential equations in Distributions (51) 32

Examples of non-classical solution situations (53) 34

Fundamental Solutions and Green's Functions (54) 36

Examples (56) 38

Exercises (57) 39

1.4 Preliminary Results on 2nd order linear ODEs (58) **40**
Theorem 1 (existence and uniqueness for initial value BCs): 40
Dependence and Independence and Wronskians (59) 40
Abel's Formula for the Wronskian (60) 41
Exercises (61) 41
The Homogeneous Differential Equation (61) 41
1.5 Boundary Value Problems (64) **44**
Big Theorem: (p66 top) 45
1. The Case of Unmixed BC's. 45
2. The Case of General BC's (67) 45
(a) initial value 45
(b) general case 46
Summary of Solutions of various ODE systems: 47
Adjoint, Symmetric, and Self-Adjoint Systems (69) 48
1.6 Alternative Theorems and Modified Green's Functions (79) **60**
System of n equations in n unknowns: the matrix problem of order n. (79) 60
Linear Integral Equations. (81) 60
Second Order ODEs (82) 61

Appendix A.1: The strings (and membrane in part (d)). (323)

I agree that it is always good to have a simple underlying basic example to think about, and this is it.

String at rest lies along the x axis on a table, tied at both ends with **tension T**. The direction perp to the string on the table is called u (think y). So our description of the string at rest is u(x) = 0 on the interval of the string. So u(x) is going to be the *deflection* of the string if we can make it deflect. The table is frictionless, gravity has no effect on the string of any interest. We are going to consider now several "problems".

(a) We apply now a **distributed force** (he calls it a **pressure**, force per length) **f(x)** in the y direction. In response to this force, the string assumes some shape u(x). Picture on page 323. Stak considers a finite piece of string from x1 to x2 as shown on page 324 bottom. The tension at the ends might not be the same, so call it S1 and S2 as shown. And angles at the ends are different θ1 and θ2. Since our piece of string (segment) is *at rest,* the x and y forces each sum to 0 and this gives the two equations shown as page 324 A. We see that the x direction of the tension is the same at both ends and must equal the tension T in the rest string, equation B. Notice that we don't really have a model here for how stretching the string might affect the tension, and we don't need one. There is just some tension function S(x). If we combine all our pieces of information, we get A.1 which describes the difference in the string slope at the two ends of our segment x1 to x2. If f(x) is at worst piecewise continuous or just "integrable", then in the limit x1 → x2, RHS = 0 of A.1 and we find that the string's slope is continuous. If f(x) is a delta function force, then there will be a change in the slope u' at the point the force is applied. But we assume f(x) is continuous, and we use the mean value theorem. [I have excellent notes on various MVT's in my Math binder] to then show p 325 A, which in the limit says -Tu" = f(x), which is our ODE of interest! Think u(a) = u(b) = 0 at the ends of the string, and this is the ODE you want to solve to find the shape of the string u(x) in response to the pressure f(x).

(b) I mentioned above that the string probably has some "**restoring force**" parallel to its direction that we did not have to know. The mechanism that causes this "elasticity" of the string is also going to create a restoring force in the y direction. We don't care about the details, we just model this restoring force by the expression F = -ku with some "spring constant k" and this gives us A.3b. It says -Tu" = $f_1(x)$ where we have just added restoring force to the pressure force so that $f_1(x) = f(x) - ku(x)$.

(c) Now let f(x) vary in time so it is f(x,t) and it causes some u(x,t). We are going to ignore any possible longitudinal modes that might somehow get activated on our elastic string and think transverse action which is in the y direction. So the first equation of page 325 B shows the x force balance equation which is the same as it was on page 325 since we assume no acceleration in the x direction of any chunk of the string. Then p 325 C shows the $F_y = ma_y$ equation with **mass density ρ** and we end up with A.4 which is the amazingly famous **"wave equation" with a driving force**. This is now a *partial* differential equation with variables t and x, so the derivatives are written using ∂. We now thrown in a "**viscous medium**" which makes a force – γ u'(x) in the y direction [linear in velocity, directed opposite v] So this thing is added into the mix in A.5 page 326 and he calls this the **"dissipative wave equation"**.

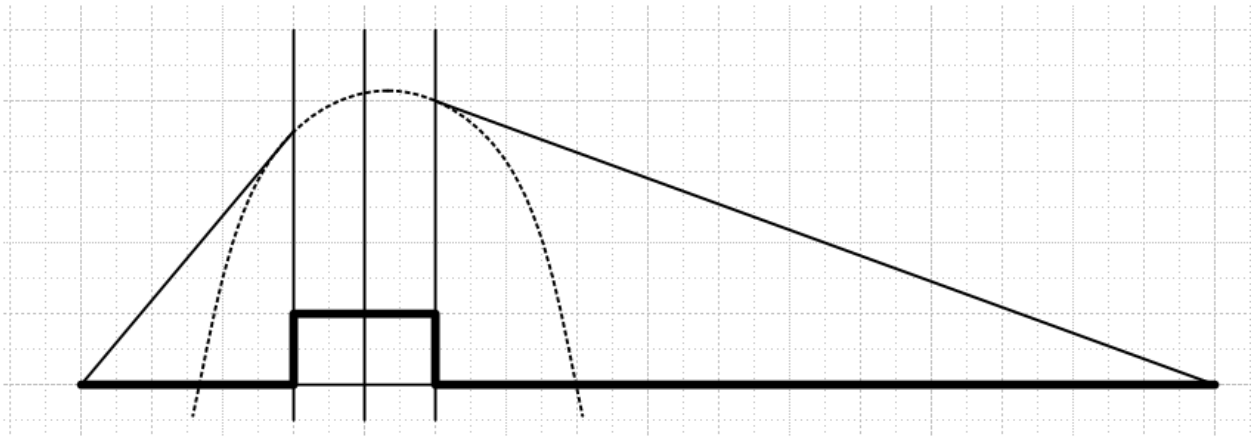
Note: I think you can imagine $T\partial_x^2 u$ as a "tension curvature force" and imagine it on the RHS of A.5. If the string is curving up, it must be deflected down, and we know the curvature force will be up and > 0. That is why we have + $T\partial_x^2 u$ if we put this on the RHS which is the "force side" of the equation. Again, the left most term of A.5 is "ma" and we just have forces.

(d) Suddenly he changes from the string to a **membrane** surface which lies in the xy plane at rest and now the pressure function f(x,y;t) is in the z direction. The claim here is that you replace $\partial_x^2 u$ in A.2 and all the other equations with $\partial_x^2 + \partial_y^2 = \nabla^2_{2D}$. He does not prove this extension of the idea, but I do prove it in "rubber sheet.doc". His point here is just to be able to write down the 2D wave equation with a driving force shown in A.6.

1.1 The string under pressure f(x) (1)

Stak quotes the ODE from Appendix A, then on p 2 considers f(x) = constant q, certainly the simplest possible problem. We integrate -Tu" = q twice and apply our BC's u(0) = u(ℓ) = 0 and get p2 A. The result is interesting in itself: under a constant pressure, the string is a parabola u = $(q/2T)[\ (\ell/2)^2 - \xi^2\ ]$ where ξ is the offset coordinate x - ℓ/2. You could think of this as the center part of the page 4 figure with the linear parts sort of phased out by gradually widening the region of constant pressure (which gives that picture)

So yes, the next problem is f(x) = p (constant) on just a finite interval centered at some value x_0. The function f(x) is then given by (1.1). To solve this problem, we "integrate" our ODE twice in each of the three regions of interest and gives names A,B,C,D to the free constants as shown page 3 A (I did all this detail). We showed in Appendix A that if f(x) is at least piecewise continuous, the ODE tells us that u'(x) is continuous. The integral of a continuous function is of course continuous so u(x) is also continuous (we are thinking of at our problematic points x ± ε/2), but we knew that already since the string does not break. So we match u and u' at the two locations of interest, giving us 4 equations in our 4 unknowns A,B,C,D and the result is given in (1.1a) -- I did not do the details here. What this solution looks like is shown in Figure 1.3, but unfortunately he has put x_0 right smack in the middle so you don't get the feel of the general solution which must look a little like this:



The peak of the parabola is now to the right of the center of the force region. And it is pretty obvious that the u"(x) is discontinuous, actually having a box shape similar to the force shown here (no surprise since that is what the ODE says!). But still u"(x) would be piecewise continuous in this example.

On page 4 he defines the "system" associated with our problem, saying u and u' are C, and u" is PC, and we have the condition that f is PC as well.

He then shows that the solution is unique.

He shows that you can superpose two solutions in the usual way since L is linear.

The Green's Function for the String with Fixed end points (5)

Then we start into our first Green's Function on page 5. The force is just $\delta(x)$ which he then describes somewhat. The sifting property is only valid if the function is continuous at the point of interest. He points out that you can consider the above square force profile as approaching $\delta(x-x_0)$ if you shrink ε but keep εp = constant, so this is our first "sequence" for a delta function.

So we regard $Lu = \delta$ as an "accessory problem" to our problem $Lu = f$. His notation at this point is $g(x|x_o)$ where x_o is the location of the delta function force. We require the BC's on g. He then just solves the problem and whittles the number of unknown constants from 4 to 2 to 1. Two are removed from the end BC's on g. One is removed from requirement that g be continuous. But we know that g' will not be continuous (saw that in Appendix A) and in fact Appendix A shows page 324 A1 that page 8 B is true (later -1/T will be called $1/a_o$). But he rederives it anyway on page 8 from scratch which is OK. This jump condition on the first derivative then determines the last constant A and our first computed Green's is shown in 1.6 and again in 1.6a with the usual $x_>$ type notation (meaning $x_> = x$ when $x > x_o$, else $x_> = x_0$).

He next claims that $g(x|x_o)$ is symmetric, and I always have a problem with this. Look at 1.6 and just by brute force make the swap:

$(\ell\text{-}x)x_o \qquad 0 \le x_o \le x \le \ell$

$(\ell\text{-}x_0)x \qquad 0 \le x \le x_0 \le \ell$

If you include both ends of the range specification, then I agree, it is patently symmetric. But the sentence at the bottom of page 9 about the reciprocal nature is a little hazy to me.

He wraps up this section on page 10 by giving two equivalent specifications of our accessory problem, where the second makes no use of the δ symbol.

The solution g is unique, just as he showed for the solution to the regular problem.

Uses of the Green's Function for the String (10)

This section opens with a heuristic proof of the fact that u = Gf is the solution to Lu=f. The idea is to of course write $u(x) = \int dy\ g(x,y)f(y) = \Sigma_i\ g(x,y_i)\ f(y_i)$ = superposition of delta function point pressures at points y_i. For each such point, you know that g is the solution since Lg = δ. So just add them up. I guess I never thought of it in quite this way. My usual approach is

$$Lu = L\int dy\ g(x,y)f(y) = \int dy\ [L\ g(x,y)]f(y) = \int dy\ \delta(x-y)f(y) = f(x)$$

but OK, so $\int dy\ g(x,y)f(y)$ is a continuous superposition of point pressure sources.

Now Stak points out that my casual application above of L is not obviously justified.

Side topic: some calculus definitions: C^k etc.

When, you ask, are you allowed to say this:

$$\partial_x\ [\ \int dy\ F(x,y)] = \int dy\ [\partial_x\ F(x,y)]$$

I think this would be true at any point x where F(x,y) is "differentiable" which means it has to be smooth with no cusp at the point x. If F(x,y) has what I will call a cusp, then dF/dx will only be piecewise continuous at the cusp location -- there will be a "jump" in dF/dx at the point. Here are some standard facts and definitions:

CONTINUOUS

(0) A function is **continuous** if it has no breaks in the interval of interest. A break is a point where the function has different values on the two sides of the point. If you plot things at such a point, your graph has a vertical segment at such a point.

Question: Can a continuous function have an infinity in the interval? Like f(x) = 1/(x-0.5). Clearly the function is discontinuous at x = 0.5. But what about $f(x) = 1/(x-0.5)^2$? It might be "continuous" when x = 0.5, who knows what happens "way up there" in terms of connectedness of the string.

This points out that the I don't have the real meaning of continuous in my definition. The real definition is that small Δ input => small Δ output, we usually say

For any ε, you can find finite δ such that $|f(x) - f(a)| \leq \varepsilon$ when $|x-a| \leq \delta$.

Consider $f(x) = 1/(x-a)^2$. The condition above becomes

For any ε, you can find δ such that $|1/(x-a)^2 - \infty| \le \varepsilon$ when $|x-a| \le \delta$ (ie, $\forall x \ni |x-a| \le \delta$)

I pick ε = 1. No matter how small (but finite) you pick δ, you have $|1/(x-a)^2 - \infty| = \infty$, so you can never get $|f(x) - f(a)| \le 1$.

Conclusion: f(x) is not continuous at any point x where $f'(x) = \infty$. If you have a closed interval, then if f(x) is not continuous at any point in that interval (which includes the end points), then f(x) is not continuous on the interval.

Theorem: If a function f(x) has an infinity within a closed interval, f(x) is not continuous on that closed interval.

Proof: if f(x) has an infinity somewhere, then f '(x) will be infinite at that same point, QED.

(1) If a function is continuous, then the integral of that function is also continuous (this seems obvious).

(1a) What about the derivative of a continuous function? I think it would be continuous for reasonable functions, but I think there are cases where the derivative is not continuous.

(2) wiki on smooth function: "Consider an open set on the real line and a function f defined on that set with real values. Let k be a non-negative integer. The function f is said to be of **class C^k** if the derivatives f', f'', ..., $f^{(k)}$ exist and are continuous (the continuity is automatic for all the derivatives except the last one, $f^{(k)}$ -- *that is to say, if $f^{(k)}$ is continuous, so are all the lesser derivatives by fact (1) above*). The function f is said to be of class **C^∞, or smooth**, if it has derivatives of all orders. f is said to be of class **C^ω, or analytic**, if f is smooth and if it equals its Taylor series expansion around any point in its domain."

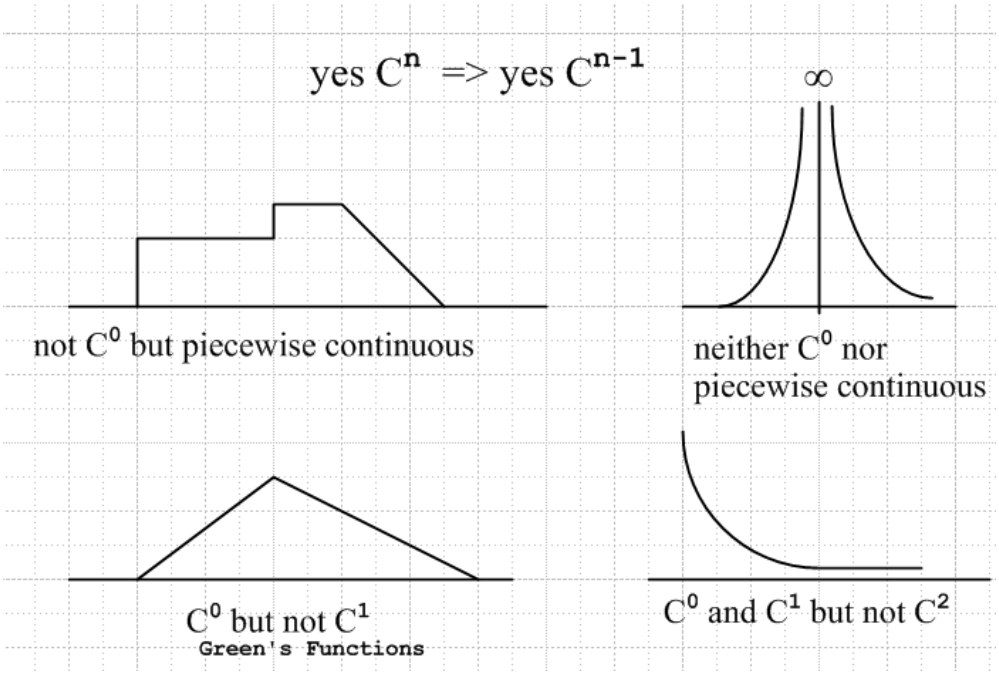
A function is **" not C^k "** if $f^{(k)}$ does not exist or is not continuous at some point in the interval.

C^0 = class of continuous functions on the interval. [to assist search for this section: C1 and C2]

C^1 = class of functions whose first derivative exists and is continuous everywhere in the interval. This is also the class of **"differentiable functions"**. If a function is not C^1, then there is at least one point in the interval where the derivative either does not exist or is not continuous. If f is a C^1 function, then f ' is continuous, and by (1) above, f itself is continuous. So, a function which is C^1 is also C^0. A function which is C^k is also C^{k-1} .... C^2, C^1 and C^0 . So this is a nested set of classes. Functions which are C^1 are sometimes said to be "**continuously differentiable**". If f is C^1, then f ' is C^0. If f is C^2, then f ' is C^1 and f " is C^0. And so on.

If a function f fails to be C^0 "at a few points" in the interval, then it is **"piecewise continuous"**. The web says there is another condition for piecewise: f ' must be finite as you approach a break from either side.

So, with all these words, what would you say about **Green's g(x,x$_0$)** shown as a triangle in Stak? It is continuous so it is C^0. The derivative is not continuous at x_0 so it is **not** C^1. I seems to be piecewise continuous.



Now resume on our question:

The following section between horizontal bars is very special purpose, not a useful general theorem, so don't pay too much attention to it! See also an external doc about interchange on the same subject, which also has no great conclusion.

Order interchange theorem for differentiation and integration.

Question: when can you say $\partial_x [\int dy\, F(x,y)] = \int dy\, [\partial_x F(x,y)]$?

Answer: Well, just write out the LHS:

$$\partial_x [\int dy\, F(x,y)] = [\int dy\, F(x+dx,y) - \int dy\, F(x,y)]/dx = \int dy\, [F(x+dx,y) - F(x,y)] / dx$$

If the limit exists, you are OK, so the answer is that F(x,y) must be C^1 in variable x. In this case

$$\partial_x [\int dy\, F(x,y)] = \int dy\, \partial_x F(x,y)$$

and order interchange is allowed.

The Green's function is not C_1, so the interchange is not allowed ! Theorem: yes C^n => yes C^{n-1}

Theorem about differentiating an integral.

Problem: Suppose F(x,y) is piecewise continuous in x and has a problem only at x = y. Let's just say it is C^1 in *both* x and y except at the point x = y. So F(x,y) is discontinuous at x = y with a finite jump. Break the integral into two pieces separated at the problem point: each side is "open" at the point, by the way.

$$[\int dy\, F(x,y)] = \int^{x-} dy\, F_-(x,y) + \int_{x+} dy\, F_+(x,y)$$

Here F_- is the C^1 function to the left of x, and F_+ is the C^1 function to the right of x. Now recall this fact:

$$d/dx [\int^{x} dy\, h(y)] = h(x) \qquad \text{// on the RHS we set the y argument of h to x.}$$

$$d/dx [\int^{x} dy\, h(x,y)] = + h(x,x) + \int^{x} dy\, (\partial h(x,y)/\partial x) \qquad (*)$$

$$d/dx [\int_{x} dy\, h(x,y)] = - h(x,x) + \int^{x} dy\, (\partial h(x,y)/\partial x)$$

Therefore we get (the idea is that on the left side, really x = x– etc so d/dx = d/dx- ; also, it is the second term on the RHS that requires C^1 of F_- for x < y, etc etc.)

$$d/dx [\int^{x-} dy\, F_-(x,y)] = + F_-(x,x-) + \int^{x-} dy\, (\partial F_-(x,y)/\partial x)$$

$$d/dx [\int_{x+} dy\, F_+(x,y) = - F_+(x,x+) + \int_{x+} dy\, (\partial F_+(x,y)/\partial x)$$

Now, we said F(x,y) was assumed C^1 in variable x away from x=y. That means dF/dx is C^0 (continuous) away from x=y. That means that we have $(\partial F_-(x,y)/\partial x) = (\partial F_+(x,y)/\partial x) = (\partial F(x,y)/\partial x)$ because this thing is continuous Thus, we can combine the two integrals to get

$$d/dx [\int dy\, F(x,y)] = \int dy\, (\partial F(x,y)/\partial x) + \{ F_-(x,x) - F_+(x,x) \} \qquad (**)$$

Side question: go back to (*) above which said this:

$$H(x) = d/dx [\int^{x} dy\, h(x,y)] = + h(x,x) + \int^{x} dy\, (\partial h(x,y)/\partial x)$$

The side question is this: what is dH/dx ? In particular, how do we compute d/dx [h(x,x)] ??? The only logical thing I can see is this:

$$d/dx\ [\ h(x,x)] = \partial_{\mathbf{x}}\ h(x,y)\ |^{\mathbf{y=x}}\ + \partial_{\mathbf{y}}\ h(x,y)\ |^{\mathbf{y=x}}$$

If this is correct, then if we do another d/dx on (**) we get

$$d/dx\ \{d/dx\ [\ \int dy\ F(x,y)]\}\ = d/dx\ \{\ \int dy\ (\partial F_{-}\ (x,y)/\partial x)\ \}\ + d/dx\ \{\ F_{-}(x,x)\ -\ F_{+}(x,x)\ \}$$

and we now want to look at the right terms. They would be

$$d/dx\ \{\ F_{-}(x,x)\ -\ F_{+}(x,x)\ \}$$

$$=\ \partial_{\mathbf{x}}\ F_{-}\ (x,y)\ |^{\mathbf{y=x}}\ + \partial_{\mathbf{y}}\ F_{-}\ (x,y)\ |^{\mathbf{y=x}}\ - \partial_{\mathbf{x}}\ F_{+}\ (x,y)\ |^{\mathbf{y=x}}\ - \partial_{\mathbf{y}}\ F_{+}\ (x,y)\ |^{\mathbf{y=x}}$$

$$=\ \{\partial_{\mathbf{x}}\ F_{-}\ (x,y)\ |^{\mathbf{y=x}}\ - \partial_{\mathbf{x}}\ F_{+}\ (x,y)\ |^{\mathbf{y=x}}\ \}\ + \{\partial_{\mathbf{y}}\ F_{-}\ (x,y)\ |^{\mathbf{y=x}}\ - \partial_{\mathbf{y}}\ F_{+}\ (x,y)\ |^{\mathbf{y=x}}\}$$

If we know that F(x,y) is continuous at x = y, then we know that $F_{-}(x,x)\ -\ F_{+}(x,x)\ = 0$, and then it will turn out that the sum of the 4 terms shown will be exactly 0. But if F is not continuous at x=y, we really have to maintain these terms. So from now on, we shall assume F(x,y) is continuous at x = y. So we conclude in this case that:

$$d/dx\ \{d/dx\ [\ \int dy\ F(x,y)]\}\ = d/dx\ \{\ \int dy\ (\partial_{\mathbf{x}}F\ (x,y)\ \}\ + 0 \qquad (***)$$

Now here is (**) from above with F replaced by Q.

$$d/dx\ [\ \int dy\ Q(x,y)]\ =\ \int dy\ (\partial_{\mathbf{x}}Q\ (x,y))\ +\ \{\ Q_{-}(x,x)\ -\ Q_{+}(x,x)\ \} \qquad (**)'$$

If we set $Q = \partial_{\mathbf{x}}F$ this becomes

$$d/dx\ [\ \int dy\ \partial_{\mathbf{x}}F\ (x,y)]\ =\ \int dy\ (\partial_{\mathbf{x}}\partial_{\mathbf{x}}F\ (x,y))\ +\ \{\ \partial_{\mathbf{x}}F_{-}(x,x)\ -\ \partial_{\mathbf{x}}F_{+}(x,x)\ \}$$

Combining this last line with (***) we conclude that

$$d^{2}/dx^{2}\ [\ \int dy\ F(x,y)]\}\ = \int dy\ (\partial^{2}{}_{\mathbf{x}}F\ (x,y))\ +\ \{\ \partial_{\mathbf{x}}F_{-}(x,x)\ -\ \partial_{\mathbf{x}}F_{+}(x,x)\ \}$$

Let's state these results as a theorem:

Theorem:

(a) Consider F(x,y) which has some kind of "problem" at x = y, something is perhaps not continuous or not differentiable there. Let $F_-(x,y) \equiv F(x,y)$ when $y < x$ and $F_+(x,y) \equiv F(x,y)$ when $y > x$. If we assume that $F_\pm$ are both C^1 in variable x ($\partial_x F_\pm$ = continuous, meaning $F_\pm$ = differentiable), then we find that

$$d/dx\ [\ \int dy\ F(x,y)] \quad = \int dy\ (\partial F\ (x,y)/\partial x) \ + \quad \{\ F_-(x,x)\ -\ F_+(x,x)\ \}$$

Notice that we have not assumed that F is C^1 at the point y=x, nor that F is C^0 there. That is to say, F might be discontinuous and/or non-differentiable at y = x.

(b) If we now ALSO assume that F is in fact continuous at x = y, AND that $\partial_x F_\pm$ are both C^1 in variable x ($\partial^2_x F_\pm$ = continuous, meaning $\partial_x F_\pm$ = differentiable) then we get

$$d^2/dx^2\ [\ \int dy\ F(x,y)] \quad = \int dy\ (\partial^2_x F\ (x,y)) \ + \ \{\ \partial_x F_-(x,x)\ -\ \partial_x F_+(x,x)\ \}$$

Proof by "careful differentiation" that $u(x) = \int dy\ g(x,y)f(y)$ solves Lu = f.

Application of Theorem 1: Let F(x,y) = g(x,y)f(y) as on Stak page 11 1.10, and g is linear thing for our string problem shown in 1.6. We know that g is C^1 except at x = y. We assume f is C^0, just continuous.

The second line above then becomes:

$$d^2/dx^2\ [\ \int dy\ g(x,y)f(y)] \ = \int dy\ [\partial^2_x\ g(x,y)]\ f(y)) \ + \ \{\ \partial_x\ g_-(x,x)\ -\ \partial_x\ g_-(x,x)\ \}\ f(x)$$

But since g of 1.6 is linear, we have $\partial^2_x\ g(x,y) = 0$ everywhere, so the above becomes

$$d^2/dx^2\ [\ \int dy\ g(x,y)f(y)] \ = \{\ \partial_x\ g_-(x,x)\ -\ \partial_x\ g_-(x,x)\ \}\ f(x) \ \ = (-1/T)\ f(x) \qquad \text{// p 8 B}$$

But we started with 1.10 saying $u(x) = \int dy\ g(x,y)f(y)$ so we have now shown that

$$d^2/dx^2\ u(x) \ = (-1/T)\ f(x) \qquad \text{or} \qquad Lu = f$$

We have shown this by what is usually called **"careful differentiation".**

Now compare this with "my" method, which is perhaps "careless" :

$$L\{\ \int dy\ g(x,y)f(y)\ \} = \quad \int dy\ \{Lg(x,y)\}f(y) \ = \int dy\ \{\delta(x-y)\}f(y) = f(x) \qquad \text{or} \qquad Lu = f.$$

Something sure is fishy here! I am not allowed to do the order interchange as I do at the first equal sign because g is not C_1 at x = y. Above I said that ∂^2_x g(x,y) = 0 everywhere, but now I am saying ∂^2_x g(x,y) ~ δ(x-y). So we need somehow to reconcile these two worlds! Stak has not mentioned "my method". Somehow the δ method knows about the careful differentiation and gets the right answer! So at least we have an element of suspense here as we "go forward" in this chapter.

Notice that we did not use any δ functions in our proof above that u(x) = $\int$ dy g(x,y)f(y) is in fact a solution of Lu= F.

Notion that L and G are inverses.

Start with u(x) = $\int$ dy g(x,y)f(y) ≡ Gf. We know that Lu = f, so we get

Lu = L(Gf) = (LG)f = f => LG = 1 when acting on f-class functions

But we can also say

(GL)u = G(Lu) = Gf = u => GL = 1 when acting on u-class functions

What are these "classes"? Going back to page 4, I think they are these:

u-class =functions where u" is piecewise continuous. Then u' = C^0 and u = C^0 so both u' and u are continuous. This agrees with the words there on page 4. Thus, u-class is C^0 and C^1 class and is in fact C^2 except at a few points.

f-class = piecewise continuous so f = weak C^0 (not continuous and not differentiable at some points) So f-class is C^0 except at a few points.

So u-class is a stricter class by two levels! The u-class is smoother than the f-class. Any function in u-class is also in f-class! Thus we have (LG)u = u and (GL)u = u and for u-class functions it seems that we can say LG = GL = 1 so that G = L^{-1} and L = G^{-1}.

Advantages of the Green's Function method.

1. The BC's are built into g, so as you change f, no need to recompute constants.

2. We have the interpretation of u(x) = $\int$ dy g(x,y)f(y) as a superposition of responses to point loads.

3. We can combine a continuous load (pressure) with point pressures as p 13 A.

4. If we shift the BC's to be more general, u(0)=a and u(ℓ) = b instead of 0 and 0, the same g for the 0,0 problem can be used but we get a simple extra term as in p 14B. This result is derived more simply I think in Exercise 1.1 using superposition. [and appears later in the chapter in the n=2 stuff]

5. The g function allows us to relate an ODE to an integral equation. Here, he talks only about the EV equation situation. The ODE is $Lu = \lambda u$ and the IE is $u = \lambda Gu$. Use $L = G^{-1}$ to move back and forth.

$Lu = \lambda u \Rightarrow GLu = \lambda Gu \Rightarrow u = \lambda Gu$ so ODE $\Rightarrow$ IE

$u = \lambda Gu \Rightarrow Lu = \lambda LGu \Rightarrow Lu = \lambda u$ so IE => ODE

Exercises: (14)

All are good, some on string, some on heat flow. I really should do these, but will let them go by the wayside for now.

1.2 The Dirac Delta Function (18)

Math can use a theory of concentrated sources, so here we do it.

δ sequences (19)

Our first δ sequence is s_k of p 19 A, and it corresponds to the H sequence r_k 19B which. The second thing is the integral of the first thing from -∞ to x. Pictures on page 19 and 20. These things become the δ function and the θ function in the limit $k \to \infty$, but Stak uses H for Heaviside function. The exercises point out various useful things. The myth that $\delta(0) = +\infty$ is thrown out by counterexamples where $\delta(0) = -\infty$ and $\delta(0) = 0$. The value must somehow be undefined for δ(x) at x=0. Similarly, the value of H(0) is undefined, though it happens to be 1/2 in the page 20 example.

Exercises (21)

The sifting property (22)

The usual sifting rule with subscript k still present is called "the action of s_k on φ(x)" and $A_k[\varphi]$ is his notation for this. He shows that in the limit, $A_k[\varphi] = \varphi(0) = \int dx\, \varphi(x)\, \delta(x)$. For the particular δ sequence we had in p 19A, he carefully proves the sifting rule, consuming all of page 23, of interest I suppose. Now we finally have our definition:

a sequence is a δ sequence if $A_k[\varphi] \to \varphi(0)$

He then shifts us out to x_o and we can redo things for $\delta(x-x_o)$ as in p 24 A. He will use the → notation in two different ways, as I always do, shown p 24 B, there the left notation is more compact if you know what is being said.

By the way, the functions φ(x) used here must be continuous and bounded for use in the sifter thing.

Exercises: (24)

1.13. Show various given δ sequences have the sifting property and thus really are δ sequences.

1.14 He ponders the dipole moment, but does not yet mention δ'(x). He shows that the - δ'(x) sequence he uses does in fact have action φ'(0). Later we will do parts to get this result.

1.15 If s(x) is any normalized function, then k s(kx) → δ(x). So this of course let's you construct sequence from any function you want that is integrable -∞ to ∞.

1.16 A lemma thing needed for the next item. It says that if $\lim_{k\to\infty}\{ \int_a^b dx\, s_k(x)\}$ is always 1 when a and b straddle x=0 (no matter how close!) and is always 0 otherwise, then s_k must be a delta sequence.

1.17 He shows here that $f_k(x) = (1-x^2)^k$ is a δ sequence. We are then told to take φ(x) as one which vanishes beyond ±1/2. The claim is that the action here will be a polynomial on our half interval which approximates φ(x). He shows that as k→∞ the approximation for φ improves to perfection. The error is in the L_1 norm (uniform convergence). This is the **Weierstrass Approximation Theorem** which is mentioned in Chapter 2. These polynomials are dense in L_1 or L_2 for example.

1.18 We get an example of an H sequence and the famous sinc function as a δ sequence.

1.3 The Theory of Distributions (28)

History

"Generalized functions" were introduced by Russian Sergei **Sobolev** (m1989) in **1935**.

"They were independently introduced in the late 1940s by Laurent **Schwartz** (m2002), who developed a comprehensive theory of distributions. In 1950, Laurent Schwartz was awarded the Fields medal (equivalent to the Nobel prize for mathematics) for his work on distributions. He was the first French mathematician to receive the Fields medal. Schwartz encountered serious problems to enter the United states to receive his medal, because of his sympathy for Trotskyism. Jewish, he had the usual WWII problems as well. The theory of distributions clarifies the (then) mysteries of the Dirac Delta function and Heaviside step function. It helps to extend the theory of Fourier transforms and is now of capital importance to the theory of partial differential equations."

My own introductory comments.

Since I have already done Chapter 2, I know about functionals like F[φ] which map F:H→R where H is some Hilbert Space, probably a function space of some sort. I learned the Riesz Rep Theorem:

" Any continuous and linear functional which acts on any subspace D of any Hilbert space H can be "represented" as T[x] = <x,f> where f is some unique element of H.

So there is a homomorphism between elements f of H and continuous functionals F. You could say that the space of continuous functionals ↔ H. He is going to refer to the space of functionals defined on D as D'.

Of course when represented in this way, the continuity in the sense $\varphi_n \to \varphi$ means that $F[\varphi_n] \to F[\varphi]$ is obvious since the scalar product is always continuous. Recall that continuous $\leftrightarrow$ bounded, by the way. Stak is not really saying that <,> is a inner product except in passing on page 28 because the whole subject of inner product spaces does not "come up" until the next chapter.

I think this is what Stak is now saying:

A **distribution** is a continuous linear functional on $D = C^{\infty}_{test}$, the space of "test functions" and I suppose the larger HS is just $H = L_2$.

I guess we would write a distribution as $T[\varphi] = \langle t,\varphi \rangle$ where now we use the letter t instead of f.

Now here is the concept he is trying to put across here: we are used to regular functions f(x) which I would write as <x |f>. He is just saying you can think of other (what Schiff would call) "representations" such as <φ|f> where you let φ range over some set of functions, $\varphi \in D$ say, as opposed to $x \in R$. You can trade one table listing of the function $\langle x_i |f \rangle$ for another table listing $\langle \varphi_i|f \rangle$ and either one can be used to define the function. [but I would think φ_i would have to be a complete set etc.]

Why is he so *strict* about D being C^{∞} with all functions having the same (a,b) boundaries of zeroness? And why so strict about the meaning of $\varphi_n \to 0$ for a test function, as on p 30 bottom? Right now, I don't yet know. So let's now return to normal notes.

Functions Viewed as Linear Functionals (28)

He mentions the notion of alternative "tables" to list a function, and he refers to the φ_i as "accessory functions". The <.> is called in inner product in passing, and $\langle f,\varphi_i \rangle$ is called a "functional". He dredges up some examples where alternatives to functions are used: (1) mass described by expansion in moments where the functions are then powers. (2) For certain BC's, $\sin(ix) = \varphi_i(x)$ are a useful set of "accessory functions" and we can do Fourier series expansion. (3) physical measurements involve a finite time, so you end up not with f(t) but with some integral of f(t) against a weight function.

Test functions (29)

These are C^{∞} and they have to vanish outside some finite interval you specify. I am in fact amazed there are such functions, and he comments that he too was amazed, but in p 29 A he produces a simple example of a test function which happens to be positive, graph page 30. It turns out that on the inside of x=1, every derivative of this function vanishes. If φ(x) is a test function in (-1,1), then the function $\varphi([x-x_o]/\varepsilon)$ is a test function in $(x_0-\varepsilon, x_0+\varepsilon)$. The test functions form a linear space D.

Why do those derivatives work at x = 1? Write

$\varphi(x) = \exp(-q(x)) = \varphi$	$q(x) = (1-x^2)^{-1}$ $\quad \varepsilon = 1-x > 0$
$\varphi'(x) = (-1)\, \varphi\, q'(x)$	$q'(x) = (-1)\,(1-x^2)^{-2}(-2x)$
$\varphi''(x) = (-1)[\, \varphi'q' + \varphi q'']$	$q''(x) = (-1)(-2)\,(1-x^2)^{-3}(-2x)$ + less singular term at x=1

In general, the terms in $\varphi^{(n)}(x)$ have the form $\varphi(1-x^2)^{-N} = e^{-q} q^N$ for various $N \geq 0$. As x approaches +1 or -1 "from the inside" at either end, we have $q \to \infty$. But write $e^{-q} q^N = e^{(-q + N \ln q)} \to 0$ for any finite N = 0,1,2,3.... Therefore each term in $\varphi^{(n)}(x) \to 0$ at the points ±1 and form a perfect match to the attached straight lines. Since these are the only points were φ might have "derivative problems", we see that in fact φ is in C^∞ everywhere on the x axis! This is a pretty amazing function.

Convergence in the Space of Test Functions (30)

A very strict meaning is applied to $\varphi_n \to 0$ if φ_n is a set of test functions: φ_n and all derivatives of φ_n must $\to 0$, and must do so uniformly on the interval, not just in the mean or other convergence. This is uniform convergence over that interval. [but I don't think this concept is used anywhere later.]

Distributions (31)

Now he says that a functional "is a rule t". For me, that rule is T[φ] = <t,φ>. (action of t on φ) [See separate notes, I would now write this as (t,φ) and reserve <f,g> to refer to an inner product.]

A **distribution** is a continuous linear functional on $D = C^\infty_{test}$, the space of "test functions". Space of all these linear functionals (space of all distributions) is called D' . He shows D' is also a linear space like D.

Locally Integrable (LI) functions (31)

Now we go off on an LI function binge. He wants to say that some distributions can be "generated by locally integrable functions", and he defines locally integrable bottom page 31. If you have a set of nested intervals $I_n \to 0$, then a locally integrable function has this property: $\lim_{I_n \to 0} \int dx|f(x)| = 0$. Since this is obviously not true for a "delta function", a delta function must not be locally integrable (it is of course integrable in the generalized function sense as we see later).

In general any function which has a non-integrable singularity like f(x) = 1/x is not going to be LI. Perhaps as the set of nested intervals in this case you could think of $I_n = [0,1/n]$ and then the integral of 1/x does not $\to 0$ as this interval $n \to \infty$.

Here is the wiki definition of a **locally integrable function**:

Formal definition

Formally, let Ω be an open set in the Euclidean space $\mathbb{R}^n$ and $f:\Omega \to \mathbb{C}$ be a Lebesgue measurable function. If the Lebesgue integral

$$\int_K |f| dx$$

is finite for all compact subsets K in Ω, then f is called **locally integrable**. The set of all such functions is denoted by

$$L^1_{loc}(\Omega)$$

Of course the details of this are beyond my understanding, since Lebesgue and compact are still on the do list. But the idea seems to be this for our application: the integral has to be finite for any bounded interval you can come up with within your starting interval. So consider the function 1/x on (-1,1). The closed set

denoted by [0,0.2] is compact and the integral has to be finite therefore on this interval, but of course we know it is not, so that is why 1/x is not locally integrable. But this argument does not explain why δ(x) is not locally integrable. You will get a finite result methinks for any subinterval. This is why Stak has to resort to the "property" about the nested intervals. I don't of course know how to prove this property from the general definition of LI given here. I am unable to find on the web any site which provides the connection between the definition of LI shown above and the nested intervals property.

[See separate notes distributions.doc. Note that δ(x) is not even a function, so it can't be a LI function. It is a generalized function.]

Theorem 1: Every locally integrable function f(x) defines a distribution according to // page 32

$$F[\varphi] = \langle f,\varphi \rangle = \int_{-\infty}^{\infty} dx\, f(x)\varphi(x)$$

One thing used in the proof is that a test function $|\varphi_n|$ has some maximum value in the interval M_n. I guess this says you cannot make a C^∞ function that has an infinity somewhere, but this is not proven. Notice that, although the test function space involves a finite interval (a,b), the integral above is (-∞,∞).

[If φ were infinite at some point, it would either be discontinuous there like 1/x at x=0, or the derivative would be discontinuous, like $1/x^2$ at x=0. Then φ could not be C^∞.]

Theorem 2. I you restrict to $f \in C^0$ (continuous), then $\langle f,\varphi \rangle = 0$ for all test functions => f = 0 on the entire real axis.

Corollary 1. If two C^0 functions f_1 and f_2 are different, they generate different distributions.

Proof of corollary: Suppose f_1 and f_2 are two different functions in C^0. Then $f \equiv f_1 - f_2 \neq 0$. By the above theorem 2 we then conclude that $< f_1 - f_2, \varphi > \neq$ for all test functions so that $\langle f_1,\varphi \rangle \neq \langle f_2,\varphi \rangle$ for all test functions φ. Thus we must have $F_1[\varphi] \neq F_2[\varphi]$ and these are then two different distributions.

Notice that the set of locally integrable (**LI**) functions contains C^0 and is thus a larger set of functions than C^0, and the above theorem and corollary are not strictly true for LI functions.

You can talk about two LI functions f_1 and f_2 being **equal almost everywhere**. This means that the integral $\int dx \mid f_1(x) - f_2(x)| = 0$ for any finite interval. This integral is not going to pick up pinhole perforations like 0 on reals and 1 on rationals, the f_2 shown bottom of page 32.

Theorem 3. (p 33) If two LI functions f_1 and f_2 are not "equal almost everywhere" (meaning they are "different" in this sense), then they generate different distributions. If they are equal almost everywhere, then they generate the same distribution.

So this is a new version of the Corollary above. He does not prove Theorem 3. He now says that we will interpret the term "two functions are equal" to mean they are "equal almost everywhere". And of course then the word "different" is being specially defined as well. With these definitions we can say:

Corollary 2: If two LI functions f_1 and f_2 are different, they generate different distributions.

[I think the notion of perforations is related to Lebesgue integration.]

p 33: Now suppose $F[\varphi] = \langle f,\varphi\rangle = \varphi(0)$. The claim is that this is an example of a functional and a distribution that *cannot* be generated by a LI function. If it could, then we need $\langle f,\varphi\rangle = \varphi(0)$ for every test function. But then he rolls out a sample test function φ_a and shows that with it, $|\langle f,\varphi_a\rangle| \le \int_{-a}^{a} dx|f(x)|$. As we let $a\to 0$, if f were LI, then $\int_{-a}^{a} dx|f(x)|\to 0$ which then says $|\langle f,\varphi_a\rangle|\to 0$, which contradicts the fact that $\langle f,\varphi_a\rangle = \varphi_a(0) = 1/e$ (it turns out). We are saying here that $\int dx\, f(x)\varphi(x) = \varphi(0)$ so we know that f(x) is our delta function thing, and we know it is not LI. He writes $\Delta[\varphi] = \langle\delta,\varphi\rangle$ as a notation for this special "delta distribution". [Write now as (δ,φ)]

If a distribution is in fact generated by a LI function f, then it is **regular**, else it is **singular**.

Remark 1 (p 34) We can always "write" a distribution as $F[\varphi] = \langle f,\varphi\rangle = \int dx\, f(x)\varphi(x)$. In the regular case f(x) is a regular function which in fact is LI. But in the singular case, f(x) is not a regular function, it is a "**symbolic function**" or a "**generalized function**" and $\delta(x)$ is an example of such. We shall start using these symbolic functions "on their own". There is of course a correspondence $\langle\delta,\varphi\rangle \leftrightarrow \delta(x)$, where the first thing is the **distribution**, and the second is the related **symbolic function**. When we want to know something about a symbolic function, we can figure it out by going back to its distribution.

Remark 2 (p 34). We can generalize all this to $F:D\to C$ instead of $F:D\to R$.

Comment: I did not notice anything in the above the explained why the space of test functions has to be so strict.

Examples of Distributions (34)

1. If you put in the Heaviside function (regular, not symbolic), then $\langle H,\varphi\rangle$ is the Heaviside distribution and it is just $\int_0^\infty dx\,\varphi(x)$, so we lose half the integration range. Function H is LI.

2. The distribution $\langle\delta,\varphi\rangle = \varphi(0)$ is singular, there is no LI function, but $\delta(x)$ is the symbolic function.

3. The distribution $\langle\delta_\xi,\varphi\rangle = \varphi(\xi)$ is singular, there is no LI function, and $\delta_\xi(x)$ is the symbolic function, and yes, $\delta_0 = \delta$.

4. The distribution <d,φ> = φ'(0) is singular. Write as $\int [-\delta'(x)]\ \varphi(x) = +\int \delta(x)\varphi'(x) = \varphi'(0)$ and you see that the corresponding symbolic function is -δ'(x). Here d = the dipole distribution.

Operations with Distributions (35)

1. Translation. (35)

Notational Question. Go back to formula: $F_{\mathbf{f}}[\varphi] = \langle f,\varphi\rangle = \int_{-\infty}^{\infty} dx\ f(x)\varphi(x)$

where I will now use F = Functional and label it by the function f.

Suppose we have g(x) = f(x-a). Then // later: $[T_{\mathbf{a}}f](x) = f(x-a)$

$$F_{\mathbf{g}}[\varphi] = \langle g,\varphi\rangle = \int_{-\infty}^{\infty} dx\ g(x)\varphi(x) = \int_{-\infty}^{\infty} dx\ f(x-a)\varphi(x) = \int_{-\infty}^{\infty} dx\ f(x)\varphi(x+a)$$

In QM we might say $\langle x|g\rangle = \langle x-a|f\rangle = \langle T_{-\mathbf{a}}\, x|f\rangle = \langle x|\ T_{-\mathbf{a}}{}^{\dagger}\ |f\rangle = \langle x|\ T_{\mathbf{a}}\ |f\rangle = \langle x\ |\ T_{\mathbf{a}}f\rangle$

which says $g = T_{\mathbf{a}}f$ and we have an unambiguous meaning for this notation. By the way,

$T_{\mathbf{b}} = \exp(-ibP_{\mathbf{x}})$ $\qquad$ $P_{\mathbf{x}} = P_{\mathbf{x}}{}^{\dagger}$ and this $T_{\mathbf{b}}{}^{\dagger} = T_{-\mathbf{b}} = T_{\mathbf{b}}{}^{-1}$ = unitary

Thus we could write,

$$F_{\mathbf{g}}[\varphi] = \langle g,\varphi\rangle = \int_{-\infty}^{\infty} dx\ g(x)\varphi(x) = \int_{-\infty}^{\infty} dx\ f(x-a)\varphi(x) = \int_{-\infty}^{\infty} dx\ g(x)\varphi(x+a)$$

$$F_{\mathbf{Taf}}[\varphi] = \langle T_{\mathbf{a}}f,\varphi\rangle = \int_{-\infty}^{\infty} dx\ [T_{\mathbf{a}}f](x)\varphi(x) = \int_{-\infty}^{\infty} dx\ f(x-a)\varphi(x) = \int_{-\infty}^{\infty} dx\ f(x)\varphi(x+a)$$

$$= \int_{-\infty}^{\infty} dx\ f(x)[T_{-\mathbf{a}}\varphi](x) = \langle f,\ T_{-\mathbf{a}}\varphi\rangle$$

Notice if f(x) is δ(x), then Taf(x) = δ(x-a) which is a function translated to the right by a.

So again, we are saying

$$F_{\mathbf{Taf}}[\varphi] = \langle T_{\mathbf{a}}f,\varphi\rangle = \langle f,\ T_{-\mathbf{a}}\varphi\rangle$$

Stak then uses the following simpler but less accurate notation to say the above:

$$F_{\mathbf{Taf}}[\varphi] = \langle f(x-a),\varphi\rangle = \langle f,\ \varphi(x+a)\rangle$$

as in page 35 A. The notation is at least unambiguous, but I feel the operator notation is much more accurate. Example: (you have to keep in mind that $[T_{-a}\varphi]$ is a "new function" obtained from δ)

$$F_{T_a\delta}[\varphi] = \langle T_a\delta,\varphi\rangle = \langle f, T_{-a}\varphi\rangle = [T_{-a}\varphi](0) = \varphi(a)$$

as in p 35 B. Of course his notation gives the same result. But his notation gives us a little more, because we have to "write out" the operator notation to see:

$$\langle T_a\delta,\varphi\rangle = \int_{-\infty}^{\infty} dx\, [T_a\delta](x)\varphi(x) = \int_{-\infty}^{\infty} dx\, \delta(x-a)\varphi(x) = \varphi(a)$$

His notation is good, being $\langle\delta(x-a),\varphi\rangle$, because it directly displays the symbolic function. Fine. He can then compare B to C and conclude that $\delta(x-a) = \delta_a(x)$.

2. Scaling (36)

Again, we could write $g(x) = \langle x|g\rangle = \langle x'|f\rangle$ where $x' = ax$, scaling. We could *define* an operator such that $|x'\rangle = a\,|x\rangle = S_a|x\rangle$. $g(x) = \langle S_a^\dagger x|f\rangle = \langle x|S_a|f\rangle$. This is not a familiar QM operator to me, and is likely non-unitary since changing the size of something, but we could still use it. Then my notation would look like this:

$$F_{Saf}[\varphi] = \langle S_af,\varphi\rangle = \langle f, S_a^\dagger\varphi\rangle$$

and we write this out like so:

$$F_{Saf}[\varphi] = \langle S_af,\varphi\rangle = \int_{-\infty}^{\infty} dx\, [S_af](x)\varphi(x) = \int_{-\infty}^{\infty} dx\, f(ax)\varphi(x) = \int_{-\infty}^{\infty} dy\, f(y)\varphi(y/a)/a$$

$$= (1/a)\langle f, S_{1/a}\varphi\rangle = \langle f, S_a^\dagger\varphi\rangle$$

which shows that $S_a^\dagger = (1/a)\, S_{1/a}$ where I assumed $a>0$. If you just do the math for $a = -b < 0$, you see that the general result is $S_a^\dagger = (1/|a|)\, S_{1/a}$. So he would write

$$\langle S_af,\varphi\rangle = \langle f, S_a^\dagger\varphi\rangle$$

$$\langle f(ax),\varphi\rangle = (1/|a|)\langle f, \varphi(x/a)\rangle \qquad \text{// which is p 36 A.}$$

Here is an example: let $f(x) = \delta(x)$ and let $a = -1$. Then

$$\langle \delta(-x),\varphi\rangle = \langle f, \varphi(-x)\rangle = \varphi(-0) = \varphi(0) = \langle \delta(+x),\varphi\rangle$$

But this tells us that $\delta(-x) = \delta(+x)$ so this symbolic function is symmetric.

3. Multiply by a LI function (36)

In this case, write g(x) = f(x)h(x) where h = LI. Notation? g = [fh] and [fh](x) = f(x)h(x). Then

$$F_{\mathbf{fh}}[\varphi] = \langle fh,\varphi\rangle = \int_{-\infty}^{\infty} dx\, [fh](x)\varphi(x) = \int_{-\infty}^{\infty} dx\, f(x)h(x)\varphi(x) = \langle f,\varphi h\rangle$$

In order for φh to be a test function, we must have h be C^{∞}. Note that h = δ would not meet this requirement. Our first inkling of why test functions so strict.

Example: suppose h = c = constant. then result is just $c\langle f,\varphi\rangle$ which is no surprise.

Example: let f = δ. Then $\langle \delta h,\varphi\rangle = \langle \delta,\varphi h\rangle = [\varphi h](0) = \varphi(0)h(0) = h(0)\langle \delta,\varphi\rangle$. This tells us that

$$[\delta h](x) = \delta(x)h(x) = h(0)\delta(x)$$

if we just compare the symbolic functions. So these two symbolic functions are the same.

Example: let f = d, the dipole distribution function. Then

DIST: $\langle dh,\varphi\rangle = \langle d,\varphi h\rangle = [\varphi h]'(0) = \varphi'(0)h(0) + \varphi(0)h'(0) = h(0)\langle d,\varphi\rangle + h'(0)\langle \delta,\varphi\rangle$

GF: $[dh](x) = d(x)h(x) = -\delta'(x)h(x) = -h(0)\delta'(x) + h'(0)\delta(x)$

So compare these last two results:

$$\delta(x)\,h(x) = \delta(x)\,h(0)$$
$$\delta'(x)h(x) = \delta'(x)\,h(0) - h'(0)\delta(x)$$

Now let's verify the second one by brute force:

$$-\int_{-\infty}^{\infty} dx\, \delta'(x)h(x)\varphi(x) = \int_{-\infty}^{\infty} dx\, \delta(x)[h(x)\varphi(x)]' \qquad \text{// parts}$$

$$= \int_{-\infty}^{\infty} dx\, \delta(x)[h(x)\varphi'(x) + h'(x)\varphi(x)] = h(0)\varphi'(0) + h'(0)\varphi(0)$$

But we know that $\varphi'(0) = \int_{-\infty}^{\infty} dx\, [-\delta'(x)]\,\varphi(x)$ and $\varphi(0) = \int_{-\infty}^{\infty} dx\, \delta(x)\,\varphi(x)$. So we can then compare the symbolic functions to get

$$-\delta'(x)h(x) = h(0)\,[-\delta'(x)] + h'(0)\,\delta(x)$$

So I can imagine lots of little "rules" that we could develop, one for δ"(x)h(x) for example.

What happens if you try to set h(x) = δ(x), not a LI function, in violation of our rules above.

$\delta(x)\,\delta(x) = \delta(x)\,\delta(0)$
$\delta'(x)\delta(x) = \delta'(x)\,\delta(0) - \delta'(0)\delta(x)$

But the objects $\delta(0)$ and $\delta'(0)$ are really undefined as we have seen in lots of examples, so these combinations don't mean anything at all.

Equality of two symbolic functions (37)

They are equal on (a,b) if their distributions are the same for all test functions φ defined on this interval (a,b).

4. Differentiation (37)

Now consider the meaning of

$F_{Df} = \langle Df,\varphi\rangle = \langle f',\varphi\rangle$

A key idea here is this: when you do parts, look what happens:

$$\langle f',\varphi\rangle = \int_{-\infty}^{\infty} dx\, f'(x)\,\varphi(x) = -\int_{-\infty}^{\infty} dx\, f(x)\,\varphi'(x) + f(x)\varphi(x)\Big|_{-\infty}^{\infty}$$

Because φ is a "test function", the parts vanish since test functions 0 outside a finite range. So this is no doubt why this strict condition is put on test functions, to make these parts vanish. Then we get

$\langle f',\varphi\rangle = -\langle f,\varphi'\rangle$

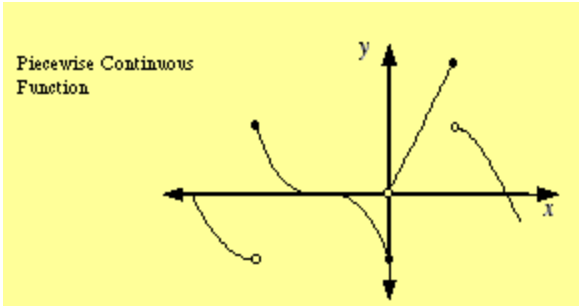
Notice that φ' is a test function if φ is. In general

$\langle f^{(n)},\varphi\rangle = (-1)^n \langle f,\varphi^{(n)}\rangle$ for any positive integer n

Examples of derivatives of distributions (38)

1. Here he shows two ways that in terms of symbolic functions, $dH(x)/dx = \delta(x)$.

2. Here he shows how to think of the derivative of a piecewise differentiable function that has finite jumps at a set of n points. If you follow what happens to the "parts" at each problem point, you end up with result 1.21. Here, the object he notates as [f '] is the conventional derivative that you would get if you ignored the jump points. But f ' is the *correct* derivative of the distribution f. If you integrate this correct f', you will accurately reproduce your original function f(x) including its jumps. Very good.



3. Here he shows again that $-\delta'(x)$ goes with $\varphi'(0)$ and claims you can write $\delta'(x) = d/dx\ \delta(x)$ as a calculus limit of $\delta(x)$ just as you would for a regular function, where you treat $\delta(x+\varepsilon)$ as a "translation" as outlined above.

4. A restatement of doing n derivatives.

5. Now let L be a general linear diff operator as shown where the coefficient functions are C^∞. You of course get $\langle Lf,\varphi\rangle = \langle f,L^\dagger\varphi\rangle$ and he writes this as $\langle f,L^*\varphi\rangle$ using the usual adjoint symbol for this book. Calls it the "formal adjoint".

Consider this simple combination first which he did not do: Let $L = a(x)D$.

$$\langle Lf,\varphi\rangle = \int_{-\infty}^{\infty} dx\ [Lf](x)\varphi(x) = \int_{-\infty}^{\infty} dx\ a(x)\ f'(x)\ \varphi(x) = -\int_{-\infty}^{\infty} dx\ f(x)\ [a\varphi]'(x)$$

$$= -\int_{-\infty}^{\infty} dx\ f(x)\ [a(x)\varphi(x)]' = -\langle f,\ (a\varphi)'\rangle = \langle f,\psi\rangle \qquad \psi = -(a\varphi)'$$

This is generalized to get result A where we start with $Lf = \Sigma_k\ a_{n-k}(x)\ D^k f$ so $\psi = \Sigma_k\ (-1)^k\ D^k(a_{n-k} f)$.

6. If a **distribution is of order n**, then you can write it as the n^{th} derivative of a regular distribution such as H. So $\delta(x)$ = order 1, $\delta'(x)$ = order 2 and so on.

Exercises (40)

1.20. An odd claim that you can make a test function which is flat from -1 to 1.

1.21. Start with the usual test function, then translate and scale it as shown. This gives you a sequence of test functions which approach the delta function! I guess not surprising that you can do that. He then shows if you integrate this test function against any C^0 function u(x) , you again get a test function $\psi_\varepsilon(x)$. But the limit of this sequence of test functions is u! So you can represent any C^0 function as the limit of a sequence of test functions, we just did it. This of course agrees with the δ limit comment above.

1.22 Another way to think about the derivative of a distribution that gives same results.

Convergence of Distributions (42)

Want to think about $t_n \to t$ which means $\langle t_n,\varphi\rangle \to \langle t,\varphi\rangle$ for all test functions φ. And we can also talk about the usual partial sum convergence idea.

Theorem 0. A distribution is infinitely differentiable.

Back on page 37 we showed that $\langle f^{(n)},\varphi\rangle = (-1)^n \langle f,\varphi^{(n)}\rangle$. Since test functions are C^∞, we know that $\varphi^{(n)}$ always exists. QED. So THIS is why we want our test functions to be C^∞ , I finally have an answer: because we want distributions to have this property of theorem 0.

Theorem 1: If LI $f_n(x)$ converges *uniformly* to f(x) over all finite intervals, then $f_n \to f$ as a distribution. [I guess f(x) is also LI, so all f_n and f are regular distributions in this little theorem.]

Proof: is shown where I marked in the book, a few lines of in-line text, not great detail.

Remark: It might be true that $f_n \to f$ as a distribution even if, as a function, we not only fail to have uniform convergence, but we also fail to have *pointwise* convergence. In our Big Example below, we have $f_n = \sin(nx)$ which pretty clearly does not have pointwise convergence at any x! But we will show that in fact $f_n \to 0$ as a distribution.

Theorem 2: If $t_n \to t$ (distributionally), then $t_n' \to t'$. Proof is a one liner bottom of page 42.

Term by term: You would not use the phrase "differentiate term by term" with regard to a simple sequence, but suppose the sequence is a partial sum thing:

$$t_k(t) = \Sigma_{n=0}^{k} a_k(t) = \text{some distribution}$$

Suppose $t_k \to t = \Sigma_{n=0}^{\infty} a_k(t)$. Then consider $t_k' = \Sigma_{n=0}^{k} a_k'(t)$. We have a finite number of terms here so no issue of whether this term by term differentiation is legal. But now our Theorem 2 tells us that $t_k' \to t' = \Sigma_{n=0}^{\infty} a_k'(t)$. So NOW we can say that for a distribution that is an infinite sum like t, we can in fact *differentiate term by term* and we get the right answer, despite doubts about doing so with an ∞ sum.

The Trick: You want to know whether $t_n \to$ some t as a distribution. If you can show that t_n is some derivative of another s_n sequence which converges uniformly to s, then Theorem 1 combined with Theorem 2 says that $t_n \to t$ as a distribution.

Big Example: Consider $t_n(x) = \sin(nx)$. Here is the symbolic function form for the distribution:

$$\langle t_n, \varphi\rangle = \int_{-\infty}^{\infty} dx\, \sin(nx)\varphi(x) \qquad t_n(x) = \sin(nx) \quad // = d/dx\,(s_n(x))$$

Consider a different sequence s_n as follows:

$$\langle s_n, \varphi\rangle = \int_{-\infty}^{\infty} dx\, [\, -\cos(nx)/n]\, \varphi(x) \qquad s_n(x) = -\cos(nx)/n$$

In the lower line, we know that $s_n(x) \to 0$ uniformly on the full range (any finite interval), in a regular function sense. We can see that $s_n \to 0$ as well. But we see that $t_n = s_n'$. Theorem 1 says that $t_n \to 0$! We can intuitively see this fact since as n increases, the sin chops up the integrand more and more into offsetting pieces. So what we learn in this example is that the function $t_n(x) = \sin(nx)$ does not converge as a normal function even pointwise, certainly not to 0 (faster and faster sine). But the symbolic function $t_n(x) \to 0$. In other words, the first line has $\langle t_n, \varphi \rangle \to 0$ even though the function sin(nx) does not $\to 0$. We say then that $t_n(x) = \sin(nx) \to 0$ "distributionally", though we do not have this convergence "classically". This is just an example of The Trick stated above. It is Example 3 in the following section.

Clarification of the above Big Example 4.4.09. We consider this distribution:

$$\text{" } \sin(nx) \text{ "} = (\sin(nx),\varphi) = \int_{-\infty}^{\infty} dx \sin(nx)\varphi(x)$$

which is finite for any n and is as happy as can be. What happens as $n \to \infty$? We learn from the above analysis that in fact $\lim_{n\to\infty}$ " sin(nx) " = 0. We can then ask what this tells us about the function inside the integral. This function sin(nx) for any finite n is just a regular function, not a generalized function. But suppose we "think of it as" a generalized function, meaning the integral of it is really defined by what is on the LHS. Since we know that limit of LHS = 0, we can "think of" the generalized function $\to 0$ and we can write that $\lim_{n\to\infty} \sin(nx) = 0$ in a "generalized function" sense. I guess we could say that in fact this limit 0 IS a generalized function S(x) = 0. Again: before the limit, sin(nx) is a regular function. But the limit is a generalized function equal to 0.

Examples of the Convergence of Distributions (43-47)

1. Here let $t_k(x) = (1/\pi)\, k/(1+k^2x^2)$ as our symbolic function. We showed earlier $\lim_{k\to\infty}\langle t_k,\varphi\rangle = \varphi(0)$, this was page 22 equation 1.14. Thus, we know that $\langle t_k,\varphi\rangle \to \langle \delta,\varphi\rangle$ so in terms of our symbolic function notation we would say

$$\lim_{k\to\infty} t_k(x) = \lim_{k\to\infty} \{ (1/\pi)\, k/(1+k^2x^2) \} = \delta(x)$$

so this gives us a "representation" of the delta function as a certain limit, all in the distribution world. The sequence of distributions on the left approaches as limit the distribution on the right. [similar to the Big Example above: here, the function $(1/\pi)\, k/(1+k^2x^2)$ is a regular function for any k, but the limit is a generalized function]

2. Same thing but this time

$$\lim_{\alpha\to\infty} t_n(x) = \lim_{\alpha\to\infty} \{ \sin(\alpha x)/(\pi x)\} = \delta(x)$$

and we have another "representation" of the delta function distribution. But $\sin(\alpha x)/(\pi x)$ is the normal integral shown, so we have proven here that

$$\int_{-\infty}^{\infty} d\omega\ e^{i\omega x} = 2\pi\delta(x)$$

Although this thing has been used forever by physics people, its validity is here proven in the distribution sense.

3. This is the Big Example above about $t_n(x) = \sin(nx) \rightarrow 0$ as a distribution though not classically.

4. We are given an $f_n(x)$ which is the δ' thing shown on page 25. At any point x>0, it converges to 0, but does not converge uniformly over the full range of course. Fine. Write this as - d/dx of $(1/\pi)\ n/(1+n^2x^2)$ above which converges to δ, so conclude that $f_n(x) \rightarrow \delta'(x)$. Not exactly sure what this example is showing, except we can interchange limit and d/dx as per our general rule for distributions. [well, this example has shown that $\lim_{n\rightarrow\infty}$ { d/dx of $(1/\pi)\ n/(1+n^2x^2)$} = δ'(x), so similar to Big Example.]

5. Now we have our **first partial sum example**. The limit $(n\rightarrow\infty)$ of the partial sum is $\Sigma_{n=-\infty}^{\infty} c_n e^{inx}$ which converges normally (and uniformly) subject to a certain condition on the $|c_n|$. But for a certain *other* condition, the series *diverges* even pointwise! However, you can write this as the $(\alpha+2)^{th}$ derivative of another series which converges uniformly subject to that same other condition to some well-defined thing g(x). So now you can do that "term by term" differentiation he mentioned earlier regarding such sums. He then shows that $\Sigma_{n=-\infty}^{\infty} c_n e^{inx} = c_0 + g^{(\alpha+2)}(x)$ where both sides are now generalized functions which don't converge normally if the first condition is not met and the second is. We don't know much about g(x) so it might likely be non-differentiable in the classical sense, but of course it is always differentiable in the distribution sense.

6. The Sawtooth. A long example, about two pages! Function f(x) is the sawtooth shown on page 45, and we can see that f'(x) = 1 most everywhere, but we know that page 39 1.21 that the true derivative is f' = [f'] + $\Sigma_k \Delta f_k \delta(x-a_k)$ where we have to throw in the "jumps" . For our new function, the jumps are all distance -1 and are at all integers, so we get

$f'(x) = 1 - \Sigma_{k=-\infty}^{\infty} \delta(x - k)$ // which is a distribution or generalized function etc etc

Meanwhile, without proof, he then claims p 44 A that you can write the sawtooth as a certain Fourier series which seems reasonable, a periodic function, etc. So here is p 44 A:

$f(x) = -(2/\pi) \Sigma_{n=0}^{\infty} [\sin(2n\pi x)/2n]$ $f_k(x) = -(2/\pi) \Sigma_{n=0}^{k} [\sin(2n\pi x)/2n]$ $f_k \rightarrow f$

We know this converges at least pointwise because that is what Fourier says. He does not show the series this is a derivative of, but he claims that whatever that series is, it converges uniformly, so we know that $f_k \rightarrow f$ as a distribution. But this is only the warm-up part of this example. Now compute:

$f'(x) = -2 \Sigma_{n=0}^{\infty} [\cos(2n\pi x)]$ $f_k'(x) = -2 \Sigma_{n=0}^{k} [\cos(2n\pi x)]$ $f_k' \rightarrow f'$

also as a distribution. Comparing to the above , we get this very strange result:

$$- 2 \Sigma_{n=0}^{\infty} [\cos(2n\pi x)] = 1 - \Sigma_{k=-\infty}^{\infty} \delta(x - k)$$ p 45 1.22

The series on the left of course diverges classically. He says in the old days, an equation like this would "create havoc", but when distribution theory came along, now there is "hardly a ripple of excitement". Stak then does a careful review of what the above means not as symbolic functions but as the functionals underlying them. When you add the integrals, everything converges and is nice for all test functions. And in doing this, we have an example of why a test function φ must be zero outside a finite interval! This may be the only place this has showed up.

Now for a grand finale, he rewrites 1.22 as 1.22a which says δ(x-ξ) = sum of trig functions. This is in fact the "**completeness relation**" for Fourier Series analysis. He integrates this thing against some function Z and we then see that Z just on the interval -1 to 1 determines Z everywhere, as apropos for a periodic function. He promises us more completeness relations in Chapter 4, which is many thousand miles away for the poor student just at page 46 in this book!

7. **Poisson Sum Deal**. (*) First, rewrite 1.22 as p 46 A which is itself a rather amazing result:

$$\Sigma_{n=-\infty}^{\infty}\delta(x - n) = \Sigma_{k=-\infty}^{\infty} e^{\pm i2\pi kx}$$ // a distribution equality

As usual, this has a sort of intuitive interpretation where you get +∞ on the right if x is an integer, and this sort of matches the blow up on the left. And away from an integer, all the different phases "cancel" on the right. You can see how a lot of arms must have been waved before this stuff was formalized.

Now, if you integrate this against some function f(x) you get

$$\sum_{-\infty}^{\infty} f(n) = \sum_{k=-\infty}^{\infty} \int_{-\infty}^{\infty} f(x') e^{-2\pi i k x'} \, dx'.$$

and this is **The Poisson Sum Formula** circa 1800 I would guess. Of course we are really supposed to use a test function φ(x) here. By fiddling this result, we get 1.23a and when the dust clears, we end up with 1.23b where Φ is the Fourier integral transform of φ. This is an interesting result all on its own, relating two series, one a sum of a function, the other a sum of its FT, and λ is along for the ride. But then Stak applies this to a specific function φ (which is not a test function, by the way) and we get 1.24. If λ is large, the LHS converges fast. If λ is small, the RHS converges fast. So this gives you a useful computational tool if you have to add up a series which might be converging painfully slowly. Rather amazing, I never heard of this before. Comment: a computer program to add up a series would never "know about" something like this. The computer does not replace analytic work, and does not replace the "theorist", so to speak. // 4.1.09 am now to page 47

(*) I have written a whole separate document "The Poisson Summation Formula.doc", see Math/Transforms/Fourier Transforms.

Exercises (47-51)

A lot of mainline stuff is appearing in these exercises!

Exercise 1.23 (p 47) This has a very *long* teaching section, followed at the very end by a little exercise.

One must understand that in this Stak section, <f,φ> does not mean inner product, it means distribution. Elsewhere I have used (f,φ) for distribution, but here we shall align with Stak <> notation.

Fact Block #1:
Equation A defines a certain classical function $f(x) = (\log x)_+$. This function is discontinuous at x = 0 as my little pencil graph shows. It is a function of a real variable x.
Equation B defines a distribution named "f" as an ε limit of an integral. This limit is needed because the integrand blows up at the ε endpoint, so being careful.
Equation C is just the application of the usual rule for the derivative of a distribution. The ε limit is still there.
Equation D is simply parts integration on the ε integral. He points out that either term in D diverges as you take the limit, but not the sum of the two terms. In the limit the first term is roughly $-\varphi(0)\ln\varepsilon$ which cancels that second term. So the question is: what do you have left after this cancellation? It is finite, and Stak wants to call it "the finite part of the integral taken from 0. " (see example below of a finite part)

Mysterious comment of Stak: (p 48)

"Since $(1/x)_+$ is not LI, it does not generate a distribution?"

What does that mean? Well, an LI function is a "normal function" and for any such function we can obtain a distribution using the usual inner product definition of the distribution. This is called a "regular distribution". Functions which are not normal functions generate "singular distributions". So what Stak really means to say is this:

"Since $(1/x)_+$ is not LI, it does not generate a *regular* distribution?"

Logic Section: I think he is demonstrating how you take the derivative of a distribution. He wants to say this:

$$<d/dx\,[\log(x)]_+ , \varphi > \;=\; < pf(1/x)_+ , \varphi >$$

where

$$< pf(1/x)_+ , \varphi > \;\equiv \text{"the finite part of } \int_0^\infty (1/x)\varphi(x)dx \text{ "}$$

$$\equiv \lim_{\varepsilon\to 0} \left[\int_\varepsilon^\infty dx\ (1/x)\ \varphi(x) \ + \varphi(\varepsilon)\ln(\varepsilon) \right]$$

Here we have several distributions. Here is a list:

$[\log(x)]_+$ = a classical function which is discontinuous at x = 0. It is LI, so it generates a regular distribution
$(1/x)_+$ = a classical function which is discontinuous at x = 0. It is not LI, so does not generate a regular distribution.
$d/dx\,[\log(x)]_+$ = a distribution which is not a regular distribution

$pf(1/x)_+$ = a distribution which is not a regular distribution

Remember that the thing in the first position like $<g,\varphi>$ is the name of the distribution. Now, the name $pf(1/x)_+$ is nothing more than a **NAME** for the distribution defined as the above limit. It is the name of the thing which gives you instructions on how to form the distribution, a set of rules, a prescription. Here that prescription is to take the limit as shown.

The integrand in the prescription is (1/x) on the positive x side, so $pf(1/x)_+$ is a suggestive name for this distribution $d/dx\ [\log(x)]_+$. We sort of mimic the function idea that $\partial_x \ln(x) = 1/x$ for $x > 0$.

So what is the point here? First, we have been able to write down a prescription for the distribution $d/dx\ [\log(x)]_+$ and second, we have given this prescription the suggestive name $pf(1/x)_+$.

Now I suspect that Stak has in mind that we might somehow say

$$< pf(1/x)_+ , \varphi > = \text{"the finite part of } \int_0^\infty (1/x)\varphi(x)dx \text{ "} = \int_0^\infty pf(1/x)_+\varphi(x)dx$$

where now we overload the symbol $pf(1/x)_+$ to be both the name of a distribution, and the name of a pseudofunction. The pseudofunction $pf(1/x)_+$ does not really exist, it is just a "notation". Compare to

$$<\delta , \varphi > = \int_0^\infty \delta(x)\ \varphi(x)dx$$

where δ is the name of a distribution and also the name of a pseudofunction.

Question: Can I install some simple test function $\varphi(x)$ and follow the prescription?

Suppose $\varphi(x)$ is $e^{-|x|}$ even though not a very good test function. Then we have

$$< pf(1/x)_+ , \exp(-x) > \equiv \text{"the finite part of } \int_0^\infty (1/x)\ \exp(-x)dx \text{ "}$$

$$\equiv \lim_{\varepsilon\to 0} [\ \int_\varepsilon^\infty dx\ (1/x)\ \exp(-x)\ + \exp(-\varepsilon^2)\ln(\varepsilon)\ ]$$

Is this really some finite number that I can write down? Maple says that

$$\int_\varepsilon^\infty dx\ (1/x)\ \exp(-x) = Ei(1, \varepsilon) \text{ which just means } E_i(\varepsilon)$$

so we are left with

$$< pf(1/x)_+ , \exp(-x) > = \lim_{\varepsilon\to 0} [Ei(\varepsilon) + \ln(\varepsilon)\ ]$$

We can go look up Schaum page 183 to find that

$Ei(\varepsilon) = -\gamma - \ln\varepsilon + (\text{power series in } \varepsilon)$

so that

$\lim_{\varepsilon\to 0} [Ei(\varepsilon) + \ln(\varepsilon)] = \lim_{\varepsilon\to 0} [-\gamma + (\text{power series in } \varepsilon)] = \gamma = 0.577 = \text{Euler's constant.}$

So there you have it. The "finite part of $\int_0^\infty (1/x)\exp(-x)dx$" is the number 0.577. And

$< pf(1/x)_+ , \exp(-x) > = 0.577$

Now we get to the actual "exercise". Repeat all of the above for a left side log function. Start with the following:

$(\log|x|)_- \equiv \log|x|$ for $x < 0$ and is 0 for $x > 0$.

Here are the equivalent steps to the positive case:

A $f(x) = (\log|x|)_- = 0$ for $x > 0$, $= \log|x|$ for $x < 0$.

B $\langle f,\varphi\rangle = \lim_{\varepsilon\to 0} [\int_{-\infty}^{-\varepsilon} dx\, \varphi(x) \log|x|]$

C $\langle f',\varphi\rangle = -\langle f,\varphi'\rangle = -\lim_{\varepsilon\to 0} [\int_{-\infty}^{-\varepsilon} dx\, \varphi'(x) \log|x|$ $\partial_x \log|x| = +1/x$ (see below)

D $\langle f',\varphi\rangle = -\lim_{\varepsilon\to 0} [-\int_{-\infty}^{-\varepsilon} (1/x)\,\varphi(x)\,dx + \varphi(-\varepsilon)\log|-\varepsilon|]$

So we arrive here:

$$\langle d/dx\,[\log|x|]_- , \varphi \rangle = \langle pf(1/x)_- , \varphi \rangle$$

where

$\langle pf(1/x)_- , \varphi \rangle \equiv$ "the finite part of $\int_{-\infty}^{0} (1/x)\varphi(x)dx$ "

$\equiv \lim_{\varepsilon\to 0} [\int_{-\infty}^{-\varepsilon} (1/x)\,\varphi(x)\,dx - \varphi(-\varepsilon)\log(|-\varepsilon|)]$

$\equiv \int_{-\infty}^{0} pf(1/x)_-\varphi(x)dx$

I think I have the sign right.

Exercise 1.24. This is a repeat of the above, but replace log(x) with $x^{-1/2}$ which also has "a problem" at x = 0. The result here is:

$$d/dx\,[x^{-1/2}]_+ = -\,1/2\ pf(x^{-1/2})_+$$

We are leading up to the famous tick principle part notation, by slow boat.

Exercise 1.25.

Now we are going to assemble our two pieces from Exercise 1.23. Those pieces are

$$\langle d/dx\,[\log(x)]_+ , \varphi \rangle = \langle pf(1/x)_+ , \varphi \rangle$$

$$\langle pf(1/x)_+ , \varphi \rangle \equiv \text{"the finite part of } \int_0^\infty (1/x)\varphi(x)dx \text{ "}$$

$$\equiv \lim_{\varepsilon\to 0} \left[\int_\varepsilon^\infty dx\ (1/x)\ \varphi(x) + \varphi(\varepsilon)\ln(\varepsilon) \right] \equiv \int_0^\infty pf(1/x)_+\varphi(x)dx$$

$$\langle d/dx\,[\log|x|]_- , \varphi \rangle = \langle pf(1/x)_- , \varphi \rangle$$

$$\langle pf(1/x)_- , \varphi \rangle \equiv \text{"the finite part of } \int_{-\infty}^0 (1/x)\varphi(x)dx \text{ "}$$

$$\equiv \lim_{\varepsilon\to 0} \left[\int_{-\infty}^{-\varepsilon} (1/x)\ \varphi(x)\ dx - \varphi(-\varepsilon)\ \log(|-\varepsilon|) \right] \equiv \int_{-\infty}^0 pf(1/x)_-\varphi(x)dx$$

If we just add these two pieces, we are then talking $\langle d/dx\,[\log|x|] , \varphi \rangle$ on $(-\infty,\infty)$. Note that $\log|x|$ is direct addition of the + and - functions we had, so we add the results all the way through:

$$d/dx\,[\log|x|] = pf(1/x)_+ + pf(1/x)_- \equiv pf(1/x) \qquad \text{// agrees p 49 G}$$

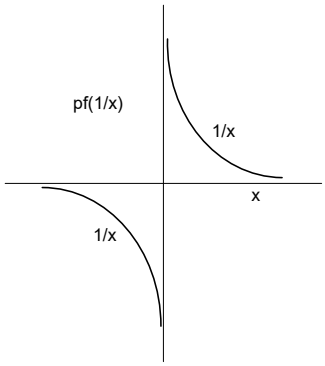
The $\varphi(\varepsilon)\ln(\varepsilon)$ pieces cancel, so the prescription ends up being

$$\langle pf(1/x) , \varphi \rangle = \lim_{\varepsilon\to 0} \left[\left(\int_{-\infty}^{-\varepsilon} + \int_\varepsilon^\infty \right) dx\ (1/x)\ \varphi(x) \right] \qquad \text{// agrees p 49 D}$$

$$\equiv \int_{-\infty}^\infty pf(1/x)\varphi(x)dx \equiv \fint dx\ (1/x)\varphi(x) \qquad \text{// the famous principle part integral.}$$

So like any other distribution, we have a well-defined prescription for what this means.

Note added 6.10.09: What does this symbolic function pf(1/x) "look like". I would have to make a crude plot as follows:



Like δ(x), this function has some kind of singularity at x = 0, it is not really defined there. It is defined really only by its action in an integral, just as with the δ function. This special integral just makes the two halves of the integral zero in on the point x = 0.

Comments: Recall the rule from page 32 1.17: If f is LI, then can write $\langle f,\varphi\rangle = \int f(x)\varphi(x)dx$. I think the function $\ln(x)_+$, even though divergent at x=0, is still LI due to the xlnx -x argument above. But when you take a derivative, the 1/x function is NOT LI for sure. Thus, you are not going to be able to write a distribution for f = 1/x is the normal (regular) manner, which would be this,

$$\langle 1/x,\varphi\rangle = \int_{-\infty}^{\infty} dx\ (1/x)\ \varphi(x) \qquad \text{// because 1/x is not LI}$$

This is similar to the idea that when you write

$$\langle \delta,\varphi\rangle = \int_{-\infty}^{\infty} dx\ \delta(x)\varphi(x)$$

the thing δ(x) is not a LI function and so this is a non-standard (singular) distribution. So the fix up is that just as you have to make up the δ(x) with certain properties, you make up pf(1/x). We sort of *got* to this result by thinking about the ln(x) and then doing a derivative, since we know that the derivative of every distribution is also a distribution!

Exercise 1.26. This shows a simple function whose limit is our same pf(1/x).

Exercise 1.27. And here we derive another famous fact:

$$1/(x\pm i\varepsilon) \ \rightarrow\ \mp\, i\pi\delta(x) + pf(1/x)$$

which is yet another appearance of our pf(1/x). This problem is a little different from the previous ones. The integral shown in A for small α is well defined. We then decompose the contour into the limit piece (principle part) and a half circle around the pole near the origin. I don't see the need for writing things out as in B. If α > 0, pole at z = -iα is in lower half plane, so our straight line contour goes above the pole.

The part at the pole is then a CW half turn which is "the wrong way", so we get an extra minus sign. The half turn then gives

$$\oint_{\mathtt{half}} dz\, \varphi(z)/(z+i\alpha) = -(1/2)2\pi i\, \varphi(0) = -i\pi\varphi(0) \qquad \alpha > 0$$

$$\oint_{\mathtt{half}} dz\, \varphi(z)/(z+i\alpha) = \mp (1/2)2\pi i\, \varphi(0) \qquad \alpha \gtrless 0$$

Thus we get p 50 C as our statement in terms of distributions. Then D is the statement in terms of pseudofunctions. This confirms my assumption above that he does intend things like pf(1/x) to represent pseudofunctions as well as distributions. The pseudofunction statement is just shorthand for the integral notation above it. This would seem to apply on any real interval that includes x=0. Otherwise the limit just becomes 1/x which is a regular function.

Exercise 1.28. Here we use the Poisson formula to derive a formula for the sum of a strange series the kind you find in GR.

Exercise 1.29 and 1.30. Here he derives the completeness $\delta(x-\xi)$ = stuff formulas for the sine only and the cosine only transform.

Differential equations in Distributions (51)

Here we are thinking about, for example, df/dx = g, where f and g are not functions, but are distributions, ie, they are functionals of the special class which we call distributions. We are not talking at this point about f and g being generalized functions, but that is probably coming, such as d/dx $\delta(x) = \delta'(x)$. He is going to use " t " instead of " f " as his generic distribution.

(a) The first case of interest is this:

d/dx " t " = " 0 "

On page 51 most of the page is consumed in showing that the only solutions to the above equation are

" t " = c c = a constant, a number (functional = real number recall)

I will come back soon and review the proof of this fact, but I want to first see where we are going.

(b) The second case of interest is this:

d/dx " t " = " f "

Another long page of stuff. Along the way, he defines this supporting distribution:

" t_0 " = $(t_o,\varphi) = (f, -\varphi_1)$

where φ_1 is a certain peculiar function obtained from φ:

$$\varphi_1 = \int_{-\infty}^{x} dx' \, \varphi(x') - \int_{-\infty}^{\infty} dx' \, \varphi(x') * \int_{-\infty}^{x} dx' \, \varphi_0(x')$$

where

φ_0 = any fixed test function such that $\int_{-\infty}^{\infty} dx' \, \varphi_0(x') = 1$. // any normalized test function

His conclusion is that all the solutions of d/dx " t " = " f " have this form

$$\text{" t "} = \text{" } t_0 \text{ "} + c$$

To me, this is an exceedingly strange result. Classically we might say $t(x) = \int dx \, f(x)$, but Stak's result bears no resemblance.

(c) He now writes the same ODE L he wrote earlier, and writes the same L* adjoint. And now we want to consider the distribution ODE

$$L \text{ " t " } = \text{ " f "}$$

We find at once with no work at all that

$$\text{" t "} = (t, L^*\varphi) = (f , \varphi) \qquad \text{since } (Lt,\varphi) = (f,\varphi) \qquad (1.38)$$

He now distinguishes three cases:

1. The distribution " t " is associated with a (classical) function t(x) which is n times differentiable all in the classical sense, so you don't get any integrals that require generalized functions in them. I think this means that the various " $t^{(i)}$ " derivative distributions are regular, not singular, for i = 1,2..n . We could then write

$$L\left(\int_{-\infty}^{\infty} dx \, t(x) \, \varphi(x) \right) = \int_{-\infty}^{\infty} dx \, f(x) \, \varphi(x)$$

The LHS here is

$$L\left(\int_{-\infty}^{\infty} dx \, t(x) \, \varphi(x) \right) = \Sigma_{k=0}^{n} a_{n-1}(x) \{ D^k \left(\int_{-\infty}^{\infty} dx \, t(x) \, \varphi(x) \right) \} = \Sigma_{k=0}^{n} a_{n-1}(x) \{ D^k \text{ "t" } \}$$

$$= \Sigma_{k=0}^{n} a_{n-1}(x) \int_{-\infty}^{\infty} dx \, t^{(k)}(x) \, \varphi(x) = \int_{-\infty}^{\infty} dx \, \{ \Sigma_{k=0}^{n} a_{n-1}(x) \, t^{(k)}(x) \} \, \varphi(x)$$

$$= \int_{-\infty}^{\infty} dx \, [\, L \, t(x) \,] \, \varphi(x)$$

Then we get as our result that

$$\int_{-\infty}^{\infty} dx \ [\ L\,t(x) - f(x)\ ]\ \varphi(x) \ = 0$$

from which we learn that

$$L\,t(x) - f(x) = 0$$

where this last is a normal linear ODE with functions in it, not distributions. In this situation, we could just solve the classical ODE Lt = f for t(x), and then we would know the solution " t " of our distribution ODE L " t " = " f " . In such a case " t " is called a **classical solution**. Seems good to me.

2. In this case, we still have " t " associated with some real function t(x), but we don't get n clean derivatives. Ie, at least some of these are singular distributions " $t^{(i)}$ " which we would need generalized functions to display in the usual integral form. This fact does not invalidate (1.38) . The solution " t " in this situation is called a **weak solution**. [I would say that " t " = "H(x)" is in the class of " t " that do not have n clean derivatives, but of course this particular " t " is not a solution of our distribution ODE.]

3. Suppose right off the bat "t" itself singular and thus corresponds to no real function, only a generalized function. Perhaps " t " = "δ". Our friend (1.38) is still true. Any solution in this case for " t " is called a **distributional solution**.

We are classifying the nature of the solution "t" which solves L " t " = " f ". The union of these three cases is called a **generalized solution**.

Now without proof he quotes a theorem related to our topic:

Theorem: If $a_0(x) \neq 0$ within L on our interval of interest, and if f(x) is LI, then all solutions of L " t " = " f " are of the classical type. This of course applies to the case f(x) = 0.

Examples of non-classical solution situations (53)

Example 1. Consider the case that L = x d/dx. Then L " t " = " 0 " as the ODE in distributions.

The meaning of L " t " is this:

$$L\ "\ t\ "\ =\ "\ Lt\ "\ = (Lt,\varphi)$$

Then if L = xD, we have

$$L\ "\ t\ "\ =\ "\ xD\ t\ "\ =\ "\ xt'\ "\ = (xt',\varphi)$$

Now how would you solve L " t " = 0 ? We have

$$0 = (xt',\varphi)$$

Suppose the distribution here called xt' is not bounded or continuous, then we cannot use the inner product to delve further. If xt' WERE continuous, we would write

$$0 = (xt',\varphi) = \langle xt',\varphi \rangle = \int dx\, x\, t'(x)\, \varphi(x) = \int dx\, t'(x)\, [x\, \varphi(x)] = (t', x\varphi)$$

and in this case x is in C^{∞} so $x\varphi$ is itself a test function.

Suppose we were to just make the definition that

$$(xt',\varphi) \equiv (t', x\varphi)$$

even if xt' is not LI. Then we can try this solution:

$$" t " = " H " \text{ so } " t' " = " \delta "$$

Then we would have

$$(xt',\varphi) \equiv (t', x\varphi) = (\delta, x\varphi) = [x\varphi(x)]^{x=0} = 0$$

and this shows that " t " = " H " is in fact a solution of the equation L " t " = " 0 ". Note that on the right we could write either " 0 " or 0 since = " 0 " = $(0,\varphi) = 0$.

Let's run this through one more time:

$$L " t " = " Lt " = (Lt,\varphi) = ([xD]t,\varphi) = (xt',\varphi) = (t',x\varphi)$$

Therefore " t " = " H" is a solution:

$$L " H " = " LH " = (LH,\varphi) = ([xD]H,\varphi) = (x\delta,\varphi) = (\delta,x\varphi) = [x\varphi(x)]^{x=0} = 0.$$

Notice that we are using our relation $(x\delta,\varphi) = (\delta,x\varphi)$ even for a singular distribution, and we do this by our definition above.

Now, what does the above look like in terms of generalized functions?

$$(x\delta, \varphi) = \int dx\, [x\, \delta(x)]\, \varphi(x) = (0,\varphi) = \int dx\, [0]\, \varphi(x) = 0$$

so we should then have this generalized function equation:

$$x\, \delta(x) = 0$$

which is a special case of our more general result for generalized functions

$$g(x)\delta(x) = g(0)$$

Now in this example, " H" itself can be differentiated to get " δ ", but we then have a derivative appearing in L which is not classical, so this is the **weak case**. The solution " t " = " H" is a weak solution.

Example 2: Now assume L = x^2 D and still L "t" = 0. Now we have:

$$L\text{ " }t\text{ " } = \text{" }Lt\text{ " } = (Lt,\varphi) = ([x^2D]t,\varphi) = (x^2t',\varphi) = (t',x^2\varphi)$$

Try t = H,

$$L\text{ " }H\text{ " } = \text{" }LH\text{ " } = (LH,\varphi) = ([x^2D]H,\varphi) = (x^2\delta,\varphi) = (\delta,x^2\varphi) = [x^2\varphi]^{x=0} = 0$$

Try t = δ

$$L\text{ " }\delta\text{ " } = \text{" }L\delta\text{ " } = (L\delta,\varphi) = ([x^2D]\delta,\varphi) = (x^2\delta',\varphi) = (\delta',x^2\varphi) = [x^2\varphi(x)]' \text{ eval at } x=0$$

$$= x^2\varphi'(0) + 2x\ \varphi(0) = 0 + 0 = 0$$

If you try t = δ', you end up with $[x^2\varphi(x)]'' = [x^2\varphi'(x) + 2x\ \varphi(x)]' = 2\varphi'(0) + \ldots \neq 0$, so t = δ' is not a solution to our equation $x^2t' = 0$. But we have found two solutions, and the first is weak, and the second is distributional!

The general solution to $x^2t' = 0$ (a distributional ODE) is t = A H + B δ where A,B constants. In generalized function manipulations we would say:

$$t(x) = \text{solution} = A\,H(x) + B\delta(x) \quad \Rightarrow \quad t'(x) = A\delta(x) + B\delta'(x)$$

Then $x^2t'(x) = x^2\,[A\delta(x) + B\delta'(x)\,] = A\,x^2\delta(x) + B\,x^2\delta'(x) = B\,x^2\delta'(x) = -B\,(2x)\delta(x) = 0$

We can guess the next one:

$$x^3t' = 0 \quad \Rightarrow \quad t(x) = A\,H(x) + B\delta(x) + C\delta'(x) \quad \text{// GF}$$

Example 3: $D^2t = \delta''(x)$ where now we are inhomogeneous and f(x) = generalized function. It is easiest I now see to just work directly in terms of generalized functions rather than the distribution. Obviously we have a solution t = δ(x) and this is a distributional solution. To this you could add solutions of the homo which is $D^2t = 0$ and thus t = A + Bx + Cδ(x) is the general solution. The homo has classical solutions, the total general solution is of the generalized solution class.

So now we have "practiced" with three different ODE's and we have obtained solutions of each of the three kinds mentioned earlier. Until I actually studied these "examples", I did not know what I was doing at all.

Fundamental Solutions and Green's Functions (54)

This is a fascinating section, now that I understand it. We start with Lt = δ(x-ε) which we can interpret in two ways. L "t" = "δ_ξ" where things are distributions, or in terms of t and δ being generalized functions.

Since L is degree n, there are n independent solutions to Lt=0 and we could imagine adding all these solutions with n constants to any particular solution we might find.

The next step is to assume that we can find a particular solution which is C^{n-2} but not C^{n-1}. That is to say, the solution t(x) is continuous C^0, C^1 and so on, but not to level C^{n-1} or C^n, so that means we are considering here a "weak solution".

Now his method of solving is interesting. Let u(x) be any homo solution of Lt=0 but only use this on the left side of x = ξ. On the right we will use some other homo solution v(x), but we anticipate that it will depend on ξ, so we call it $v_\xi(x)$. We are thus looking at p 54 A. Now imagine writing out all of L

$(a_0 D^n + a_1 D^{n-1} + ... a_n D^0) t(x) = \delta(x-\xi)$ // the $a_i(x)$ are functions

Integrate this just across the point x = ξ. The $a_i(x)$ are all continuous we assume somewhere, so we get this idea going on in each term

$$\int_{\xi-\varepsilon}^{\xi+\varepsilon} dx\, a_1(x) D^{n-1} t(x) \approx a_1(\xi) \int_{\xi-\varepsilon}^{\xi+\varepsilon} dx\, D^{n-1} t(x) = a_1(\xi) D^{n-2} t(x)\Big|_{\xi-\varepsilon}^{\xi+\varepsilon} = a_1(\xi)\{ D^{n-2}v_\xi - D^{n-2}u \}$$

But the RHS = 0 for this term and all other terms except the leading term due to our continuity of solution assumptions above. That leading term gives (integrating the leading term lowers power of D by one)

$$\int_{\xi-\varepsilon}^{\xi+\varepsilon} dx\, a_0(x) D^n t(x) = a_o(\xi)\{ D^{n-1}v_\xi - D^{n-1}u \} = 1$$

where we have now also integrated the RHS to get 1. You now see why we need to assume $a_o(x) \neq 0$, and we then have a jump rule for the n-1st derivative: $\{ D^{n-1}v_\xi - D^{n-1}u \} = 1/ a_o(\xi)$. In other words, D^{n-1} of the Green's Function has a discontinuity at x = ξ of amount $1/ a_o(\xi)$. In our string example, n = 2 and we had a jump into the first derivative at x = ξ, causing the tight elastic string type Green's picture.

Equation 1.40 restates what we already know, and page 55 A just states that the two functions on the two sides of x=ξ are the same for all derivatives (inclu 0) (we assumed solution C^{n-2} recall) *except* the n-1stone . The claim now is that the functions on the two sides are two *independent* homo solutions due to that last jump condition. You cannot have one solution be a multiple of the other. [Why? Suppose v were a multiple of u: v(x) = k u(x). then v(ξ) = ku(ξ) at x=ξ from p 55A (sol is cont thru C^{n-2}) so k = 1. But then have $v^{(n-1)}(\xi) = u\, v^{(n-1)}(\xi)$ which violates last rule in p 55A.]

We have made various assumptions, but if we can now prove that 1.40 solves our original ODE, then they are justified. If we can show that (Lt,φ) = φ(ξ) at the distribution level, then at the generalized function level we know that Lt = δ(x-ξ) which is our goal to show.

At this point I went off and wrote another section in my "distributions.doc" document and came to certain conclusions there which I will summarize here:

(1) The following five rules or "operations" are valid for singular as well as regular distributions

(t(x+a),φ) = (t,φ(x-a))
(t(αx),φ) = $|\alpha|^{-1}$(t(x),φ(x/α))
(gt,φ) = (t,gφ)
(t', φ) = - (t, φ')

$(Lt,\varphi) = (t, L^*\varphi)$

To show they are "valid" for singular distributions, you consider a sequence $t_n \rightarrow t$ where $t_n(x)$ are perfectly well behaved functions. The rules for t_n are all obvious from simple use of the integration machinery, such as redefinition of the integration variable or parts integration. You can then write out the above rules in the limit using the integral form if you regard certain factors inside the integral as generalized functions. The fifth rule above follows from the 3rd and 4th if you assume that your $a_i(x)$ coefficient functions are C^∞. I give examples of δ and δ'.

(2) I then discuss the proof Stak is doing on page 55. He wants to prove that the candidate solution shown in 1.40 does in fact solve $Lt = \delta(x-\xi)$ which means $(Lt,\varphi) = \varphi_\xi(0) = \varphi(\xi)$. He then carries out the proof. The first step in the proof is result 5 above so we have $(Lt,\varphi) = (t, L^*\varphi)$. He then writes this as two separate integrals in order to avoid the problem point ξ, and he properly uses $u(x)$ on the left and $v_\xi(x)$ on the right. This is shown in page 55 B and in my notes. He then starts "swinging" the D operators from right to left. I show that, if we ignore all parts contributions, he ends up with $(Lt,\varphi) = 0$ which is the wrong answer. There is in fact one place where you get parts contributions, and I show in my notes that that parts contribution is exactly $\varphi(\xi)$ and thus our proof is complete. The one place parts contributes is where the ξ parts contribution in the two terms does not cancel, and that happens when you have $D^{n-1}u$ sitting in the parts expressions, because this parts term then gives $(D^{n-1}v_\xi - D^{n-1}u)$ which, instead of being zero, is given as shown in p 55A because this function has a non-zero "jump".

(3) The most general solution to our ODE $Lt = \delta(x-\xi)$ is our Green's Like solution 1.41 + the n homo solutions each with an undetermined constant. Our ODE "system" is then completed by specifying n BC's which determine those n constants (assuming that the BC's allow you to find those n constants). We arrive then at this important conclusion: for n reasonable BC's, there is exactly one solution to $Lt = \delta(x-\xi)$ and it can be written as 1.40 + n properly constanted homo solutions. But, there are n ways to write this unique solution because the function $u(x)$ in 1.40 can be any one of the n homo solutions. It is possible that one of these n ways will result in all the n homo constants being 0. However you write it, the thing you get which matches all the n BC's is called "the Green's Function" for the ODE system including the BC's, and it is unique. [We have not yet written this thing as $g(x|\xi)$ nor have we yet claimed that it is symmetric.] Notice that this Green's Function concept is general to our n^{th} order ODE, though we will probably only use if in the case n = 2 since that includes most problems of interest.

Definition: Any solution of $Lt(x) = \delta(x-\xi)$ is called a "**fundamental solution** with pole at ξ ".

Definition: If you add n BC's, this fundamental solution is unique, and is called a **Green's Function**.

Examples (56)

Example 1: Contradiction: the first claimed solution has a slope discontinuity of 1, which seems right. The second claimed solution has a slope discontinuity of 2, which seems wrong. How can this second solution be right? Resolution: slopes are not ± 1. Slope on left is $(1-\xi)$, slope on right is $-\xi$ and difference in these slopes is 1.

Example 2: Suppose you set u(x) = 0 for your left-side function. The conditions on $v_\xi(x)$ at $x=\xi$ are then pretty simple as shown here. Contradiction: how does he conclude that on the right $v_\xi(x) = v_o(x-\xi)$ if L has constant coefficients (a typical Stak non-obvious statement). At $x=\xi$ we do have $v_\xi(\xi) = 0$ as required. But we have $v_\xi'(\xi) = v_o$ which seems wrong unless n = 2 and $1/a_0 = v_o$. But he seems to be claiming this conclusion for any n. Maybe he just means v_o is some $v_o(x)$ and he has factored out the required 0.
Resolution: v_0 is a function, not a constant, and what he is claiming is that $v_\xi(x) = v_o(x-\xi)$ if coefficients are constant. The underlying reason is that if the coefficients are constants, then L as an operator is invariant under translations, but I don't know how to write that out exactly. But here is an example:

$$(aD_x^4 + b)v_\xi(x) = 0 \qquad x > \xi$$

Let $y = x-\xi$. Then since a and b are constants, we can write this as

$$(aD_y^4 + b)V_\xi(y) = 0 \qquad y > 0$$

We have just rewritten the right-side problem in a new coordinate. We replace v with V because the functional form is different, and we have $v_\xi(x(y)) = V_\xi(y)$. The second problem has no ξ appearing in it, so the solution is some $V_\xi(y) = v_o(y)$. Therefore, $v_\xi(x) = v_\xi(x(y)) = v_o(y) = v_o(x-\xi)$, as he claims.

Now then the total solution is as shown in p 56 A and the conditions on $v_\xi(x)$ at $x=\xi$ are now transferred to $v_o(x)$ at x = 0. The BC's shown are a very particular set that would be appropriate for at least some problem: $v_0(0) = 0$, $v_0'(0) = 0$ ..... and only the n-1st derivative is non-zero. He calls this particular Green's function solution the **causal fundamental solution**. If we had n = 2 and the equation was $m\ddot{x} = \delta(t-\xi)$, then $x_0(0)=0$ and $x_0'(0) = s$ (some initial speed).

Example 3: This just conjecturally talks about how we might solve Lw = f if we have already solved $Lt_\xi(x)= \delta(x-\xi)$. The solution ought to be $w = f(\xi_1)\, t_{\xi 1}(x) + f(\xi_2)\, t_{\xi 2}(x) + \ldots = \int d\xi\, t_\xi(x)\, f(\xi)$. Apply L to both sides and it works. If $t_\xi(x)$ meets your BC's for any ξ, then w(x) also meets those BC's. So it behooves you to find the solution $t_\xi(x)$ that meets the BC's that "come with" the Lw = f problem.

Exercises (57)

Exercise 1.31. Here we consider the problem xt = 0 in distributions. For functions, the answer would be I think that t(x) = 0. Here he shows in a flow of controlled logic that in fact $t(x) = c\,\delta(x)$ is the most general solution to this equation, and of course we know that $x\delta(x) = 0$ as a generalized function rule, because we can check it as $\int dx\, x\, \delta(x) = x(0)$ (x evaluated at x = 0) = 0. So this in itself is good to know. Along the way, we consider $\psi(x) = x\varphi(x)$. We are to show that if $\varphi(x)$ is a test function, then $\psi(x)$ is a test function with $\psi(x) = 0$. That part is pretty obvious. The other way is to show that if $\psi(x)$ is a test function with $\psi(0) = 0$, then it can be written as $x\varphi(x)$ where φ is a test function. I did not prove that but I believe it. This is crucial to his proof that $t(x) = c\,\delta(x)$ for the problem xt = 0.

Now if you consider the problem xt' = 0 you conclude that $t'(x) = c\,\delta(x)$. This is certainly compatible with the conjecture that t(x) = cH(x) + d. If we start with t, we get t'.

Exercise 1.32. (58) We are supposed to solve $x^k t' = 0$. I would start with $x^k t = 0$, or $\langle x^k t, \varphi \rangle = 0$ or $\langle t, x^k \varphi \rangle = 0$. I would then consider $\psi(x) = x^k \varphi(x)$ and show that this means $\psi(0) = \psi'(0) = \ldots \psi^{(k-1)}(0) = 0$ and the previous exercise dealt with k = 1. We then do some kind of trick similar to what he did there, and the conclusion is probably this: $t = c_0 \delta^{(k-1)}(x) + \ldots + c_{k-2}\delta(x)$. All the lower order terms would satisfy this thing I think. That is, if t solves xt=0, it also solves x(xt) = 0. Then the solution of $x^k t' = 0$ would be obtained by saying first that $t' = c_0\delta^{(k-1)}(x) + \ldots + c_{k-2}\delta(x)$ and then $t = c_0 \delta^{(k-2)}(x) + c_1 \delta^{(k-3)}(x) + \ldots + c_{k-2} H(x) + d$. Just my guess.

Exercise 1.33. (58) We are to show that the solution of xt' + t = 0 is $t = c\,\delta(x) + d\,pf(1/x)$. This would be good for me to do, so I will put it on my "to do" list (with 10^9 other things).

Exercise 1.34. (58) Consider Lw=f with a particular set of BC's that $w(0) = \ldots = w^{(n-1)}(0) = 0$. This looks like p 56 A but with $a_0 = \infty$. It turns out that the correct solution for w is the Volterra form! Another one to put on the list. That is the difference between a real class and what I am doing. I want to finish my pass through this chapter and not take all year to do it.

1.4 Preliminary Results on 2nd order linear ODEs (58)

The coefficient functions are now a_0, $a_1$and a_2 and are just assumed continuous. In Lu = f, we assume f is piecewise continuous. And throughout we assume that $a_o(x) \neq 0$ over the interval.

Theorem 1 (existence and uniqueness for initial value BCs):

If $a_o(x) \neq 0$ anywhere in your interval I, then Lu=f with the "initial-value" BC's (but at any point x_0) has a solution, and that solution is unique.

No proof given. This theorem is for n = 2, but it is also true for any n, and this is the subject of Exercise 1.36 which merely states the theorem for the n case and then you are supposed to prove it.

Remarks (59) He gives one example which violates the $a_o(x) \neq 0$ condition and finds a solution but it is not unique. Then in a second example of this type, he finds there is no solution at all. So $a_o(x) \neq 0$ is obviously a crucial assumption of the theorem.

Corollary 1: If the initial-value BC';s are all 0, only solution is $u \equiv 0$.

Corollary 2: Effect of scaling the BC's to get a scaled solution.

Dependence and Independence and Wronskians (59)

He first defines the n^{th} order **Wronskian** of a set of n functions, just the det of the usual matrix and of course it is a function of x. Do not confuse this det of matrix with the Jacobian which is a determinant of a matrix of derivatives of a function that does f: $R^n \rightarrow R^n$ so there are n variables and n component functions. I think these two determinants are very highly non related, just both happen to have the same det of matrix form. In our n=2 case, the Wronskian is pretty simple, 1.46.

Abel's Formula for the Wronskian (60)

It is trivial to show that, if u and v are two solutions of Lu=0, then $W(u,v:x) = C\, e^{-m(x)}$ where m(x) is a solution to $m'(x) = a_1(x)/a_o(x)$. Notice that this solution has a free constant such that $m(x) = \int^x dx\, a_1(x)/a_o(x) + K$. Once you select K, you have selected your m(x), and then C is no longer a free constant but is set by your functions u and v. This thing is **Abel's Formula** (Abel 1802-1829 TB death, so this formula is probably circa 1824. The word "abelian" is due to Abel, a Norwegian, email Tom.)

So one point is that the x=dependence of W(u,v:x) is entirely in this m(x) thing. It is the integral of the function $a_1(x)/a_o(x)$ which is C^1 on the interval, has no poles, has no infinite points. [The function 1/x is not C^1 at x=0, so not on interval (0,1) I suppose.] Therefore, if $W(x=x_0) = 0$, then W = 0 everywhere in the interval. So we have just proven

Theorem 2: If W of two solutions vanishes at one point, it vanishes on the entire interval.

He then goes on to prove another theorem which requires a little more work to show:

Theorem 3: solutions u and v are dependent $\Leftrightarrow W(u,v:x_0) = 0$ for some point x_0 in I
$\Leftrightarrow W(u,v:x) = 0$ for all x in I (Thm 2)

Going to the right is easy, because then v = Ku and det = 0. Going the other way requires some simple linear algebra and the uniqueness theorem 1 above. All OK.

Exercises (61)

1.35. I think I have the answer today: the two functions don't solve the same Lu=f ODE, so the theorem does not apply.

1.36. States existence and uniqueness of solution to Lu=f for initial-value BC's but **order n**. It is also claimed that the Abel formula is true also for order n from which it follows that W=0 at some point implies W=0 at all points.

The Homogeneous Differential Equation (61)

The theorem says there are two indep equations for Lu = 0 (order 2) and proof is given.

Next, on p 62 he restates all our known facts for the case n = 2 for of $Lt = \delta(x-\xi)$. Formally the $a_i(x)$ have to be C^∞ in which case there is always a weak solution for t. If not C^∞, we can just regard 1.49 as defining the nature of our solution, then see what happens.

Now take 1.49 on its own. We can construct a solution using our u(x) and $v_\xi(x)$ method earlier. This solution is outlined in p 62 A. If we think of this as a problem for $v_\xi(x)$ over the whole interval, we have your n = 2 initial-value BC's at the point $x = \xi$, and that gives us existence and uniqueness of $v_\xi(x)$on the interval, and thus $t(x|\xi)$ exists and is totally well defined. Since we never proved the uniqueness theorem (just quoted it above), I don't know if C^∞ of the coefficient functions is required for the proof. I guess not. I suspect C^1 is enough for the $a_i(x)$. [well, there is the test function issue still]

We then assume our superposition idea works for the Lw = f equation.

Now below my pencil line on page 63 we have a little side discussion which I am having trouble following. // I now inject a section from my meta notes which I think makes this clearer, with bar separators, but maintain all the raw notes:

The ∞ issue and the Volterra-form solution to Lu=f.

We are on page 63 and the logic is very tangled, but I think I now see the forest for the trees. Our usual Lz=f solution is $z(x) = \int_a^b d\xi\ g(x|\xi)f(\xi)$ + homo solutions, where we construct g from u(x) and $v_\xi(x)$ which are both solutions of Lu=0. Recall that each choice of u(x) results in a different form for g(x|ξ), but any form will work in our formula for z(x). Recall that the "causal" solution had u(x) = 0 and was as good as any other solution. Our *concern* is that if (a,b) = (-∞,∞), the integral might diverge. On the bottom of page 63 the *upshot* is that, *in this case*, we can write the solution as $z(x) = \int_\alpha^x d\xi\ v_\xi(x)f(\xi)$, where $v_\xi(x)$ satisfies the usual BC's of our "causal" fundamental solution which are these: $v_\xi(\xi) = 0$ and $v_\xi'(\xi) = 1/a_0(\xi)$. The constant α here means that z(x) satisfies the homo initial-value BC's at x = α (obvious that z(α) = 0, but also true that z'(α) = 0). So you can pick the constant point x=α as you see fit. Due to the Volterra type upper endpoint x, and picking a finite α, we have avoided any possible divergence of the integral, which alleviates the concern just mentioned.

In the derivation of this form of z(x), Stak refers to α as "a" which confused me on the (a,b) issue. He should have used some other symbol like α.

So here is how Stak comes up with the Volterra solution form. He writes the two obvious u=0 "causal" solutions and calls them t_1 and t_2. He picks a constant α (his a). For x > α he uses the t_1 form for g. For x < α he uses the t_2 form for g. In each case, the solution z(x) is the identical Volterra form $z(x) = \int_\alpha^x d\xi\ v_\xi(x)f(\xi)$ which of course does not mention t_1 or t_2. So why not use this solution for x anywhere in the range (-∞,∞)! For any finite x and α, there are no ∞ divergence issues. I now quote from the raw notes, using a = α :

---

back to raw notes:

We have t_1 being our standard solution with u(x) = 0 called the causal Green's on page 56.

The second form t_2 is not obvious since we did not talk about this possibility before. Well, suppose we start all over on page 54 putting u(x) on the left instead of the right side. The two functions match in all derivatives at x=ξ except for the jump. The jump will have the opposite sign! So for example we would have that $v_\xi(\xi) = u(\xi)$ but $v_\xi^{(n-1)}(\xi) = u^{(n-1)}(\xi) - 1/a_o(\xi)$. Now if we set u=0, we get $v_\xi(\xi) = 0$ and so on and then finally $v_\xi^{(n-1)}(\xi) = -1/a_o(\xi)$. So if we want to use the same set of relations shown on page 55a, we can make that all work by replacing $v_\xi \rightarrow -v_\xi$, and so that explains t_2 on page 63.

The next piece of logic is unclear, but I have finally "decoded it". My old pencil notes say "I dig it" but it took me a long time here to dig it again. And I agree he should have added the pencil conclusion I wrote long ago at the bottom of page 63.

(1) Suppose the interval is (a,∞). Then we get, for x in the interval,

$$z_1(x) = \int_a^\infty d\xi\, t_1(x|\xi)\, f(\xi) = \int_a^\infty d\xi\, \theta(x-\xi)\, v_\xi(x) f(\xi) = \int_a^x d\xi\, v_\xi(x) f(\xi) \qquad x > a$$

$z_1(x)$ is clearly a solution of our ODE $Lz = f$ on (a,∞). The form of t_1 makes it have that upper Volterra looking endpoint. So we have a fine solution for $x > a$. AND, we have avoided the possibility of the integral diverging at $\xi = \infty$ since we stop at x.

(2) Now, can we get some solution on the interval $(-\infty,a)$? This is why he gets into the z_2 stuff because we need a solution when $x < a$ as well. So here we go:

$$z_2(x) = \int_{-\infty}^a d\xi\, t_2(x|\xi)\, f(\xi) = -\int_{-\infty}^a d\xi\, \theta(\xi-x)\, v_\xi(x) f(\xi) = -\int_x^a d\xi\, v_\xi(x) f(\xi)$$

$$= \int_a^x d\xi\, v_\xi(x) f(\xi) \qquad \text{all this for } x < a$$

OK, we I finally have it figured out. So $z_2(x)$ and $z_1(x)$ *are the same formula*, so let's just call it $z(x) = \int_a^x d\xi\, v_\xi(x) f(\xi)$ and this then works for ALL x. So this is our solution on $(-\infty,\infty)$.

(3) Notice that $z(a) = 0$. If we go ahead and compute $z'(x)$ we get what he shows, and then $z'(a) = 0$ because (1) the integral has the same end points so it is zero, same as the way we got $z(a) = 0$; (2) $v_\xi(\xi) = 0$ for any ξ, so $v_a(a) = 0$, so the second term is also 0. So I am happy with $z(a) = 0$ and $z'(a) = 0$ which are then some BC's for z(x).

(4) Therefore, we may conclude that $z(x) = \int_a^x d\xi\, v_\xi(x) f(\xi)$ is a solution to $Lz = f$ on the interval $(-\infty,\infty)$ with the boundary conditions $z(a) = 0$ and $z'(a) = 0$, where $v_\xi(x)$ is the solution of $Lv = 0$ with the (different) boundary conditions that $v_\xi(\xi) = 0$ and $v_\xi'(\xi) = 1/a_o(\xi)$. There is only one such solution $v_\xi(x)$ to $Lv = 0$ with those BC's. We have found a particular solution of $Lz = f$. We know there are no solutions to $Lz = 0$ with BC's $z(a) = 0$ and $z'(a) = 0$ (our corollary earlier, ie, only solution is $z \equiv 0$), so there are no homo solutions to add, so our z(x) is then THE ONLY solution of $Lz = f$. This is the **Theorem p 63**.

Comment: It is not exactly obvious *how* you meet the BC's $v_\xi(\xi) = 0$ and $v_\xi'(\xi) = 1/a_o(\xi)$ with a solution of the homo equation $Lv_\xi(x) = 0$. We have to have a general solution for all values of ξ, because all these values will then appear in the integral. I suspect he will return to this theorem later on, and I am glad it has now been decoded.

1.5 Boundary Value Problems (64)

We set up our general 2nd order problem on (a,b) with the same conditions as before on the $a_i(x)$ and on f(x). But the new feature here is that we write Goldstein-style "homogeneous" BC's $B_i(u) = 0$ where each is a linear function of u(a), u'(a), u(b) and u'(b). The footnote explains why we don't want to see things like u"(a) in such a BC. Obviously the two BC's want to be independent. And there are the usual two special cases: (1) B_1 involves only endpoint a, and B_2 only endpoint b. This is the unmixed = pure BC.
(2) the "initial value problem" BC's. He shows this at the left endpoint only. So all this is fine.
Now on page 65 we have some comments:

1. Finally I see the definition of his "*completely* homogeneous system" . It means Lu=0 and our two homogeneous boundary conditions $B_i(u) = 0$. So the Lu=0 is the inhomo ODE, and the BC's are the inhomo BC's. So, if you have two solutions of Lu=f, they have to differ by a "completely" homo solution! This difference solution Δ has to have $B_i(\Delta) = 0$ so that the homo BC's for Lu=f are not altered! So this is just dandy.

2. If the completely homo has no solutions (other than 0), then Lu=f has at most one solution. This follows from item 1. Later we will show that this one solution exists.

3. If the completely homo has non-trivial solutions, things become a bit unpleasant. The Lu=f problem in this case can has either (1) no solution (2) many solutions. I imagine this is saying L has a nullspace, so the alternative theorem restricts the range, so certain f in the range won't have solutions. [meta insert:]

Fact: If complete homo Lu=0 has solutions, then operator L has a nullspace of some dimension N>0, and we know this will restrict the range of Lu=f (from our knowledge of later chapters in Stakgold). If f is not in the range, there will be *no* solutions to Lu=f. If f is in the range, there will be N+1 solutions given by a particular solution + the nullspace solutions. So "none or many".

4. This is a linearity fact regarding solutions for u_1 and u_2 of Lu=f for f_1 and f_2. The lincom of these solutions $u = c_1u_1 + c_2u_2$ solves the case $f = c_1f_1 + c_2f_2$ and then $B_i(u) = 0$ as desired.

Now we have a seeming change in topic, we are done with the above "observations". He states the Green's Function problem now with the homo BC's. He states this problem with and without δ functions. He is not yet saying that the Green's function solution exists or is unique. But here it is:

First, let's clearly define two problems on interval (a,b):

(1) Complete Homo: $Lu = 0$ with $B_1(u) = 0$ and $B_2(u) = 0$ 1.54

(2) Green's Problem: $Lg = \delta$ with $B_1(u) = 0$ and $B_2(u) = 0$ 1.55

where L is second order with continuous a_i and where B_i are as shown in 1.51. In both the above problem, we have Homo boundary conditions.

Big Theorem: (p66 top) If Complete Homo has only the trivial solution, *then* the Green's Problem has a solution and that solution is unique.

We know that any difference of two solutions Δ must satisfy complete homo 1.54, and if it has no solutions, we know there can not be more than 1 solution. So our problem is to show existence. Stak will do this be constructing an explicit solution in a few distinct cases. (end for 4/6/09)

1. The Case of Unmixed BC's. He first argues that there exist functions $u_1(x)$ and $u_2(x)$ which are non-trivial solutions of Lu=0, where u_1 satisfies only the B_1 unmixed condition of (1.52) [the a end] and u_2 satisfies only the single B_2 condition at the b end. How do we know such solutions even exist? For $u_1(x)$ suppose we take the specific BC at end a which says $u_1(a) = \alpha_{12}$ and $u_1'(a) = -\alpha_{11}$ which does make $B_1(a) = 0$ as I label it in 1/52. Now, back on page 58 bottom we had a Theorem which said that Lu=f with the "initial value" BC's had a unique solution (though this theorem was not proven; we apply it to the present case by setting f = 0). In our current circumstance, $u_1(a) = \alpha_{12}$ and $u_1'(a) = -\alpha_{11}$ is the "initial value" BC's for $u_1(x)$, so $u_1(x)$ must exist and be unique, by that p 58 theorem. So in this way we "construct" unique functions u_1 and u_2 that do what we need, though other pairs of functions likely exist as well, with different ways to solve 1.52 at the end points.

We now claim that u_1 and u_2 are independent. If they were not, then $u_1 = K\,u_2$ and then u_1 or u_2 would satisfy the BC's at *both* ends, B_1(a end) = 0 and B_2(b end) = 0. But by assumption of our theorem here, there can be no such solutions because Complete Homo is assumed to have only the trivial solution. Therefore u_1 and u_2 are independent.

The next step is to write g in the form p 66A. We know we can do this if we meet the two usual conditions that g be continuous at ξ, and that the first derivative have the usual jump there, as shown in B. We now have a matrix equation $M\begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} C \\ D \end{pmatrix}$ [this is p 66 B] which has a solution (we want it to have a solution!) when $\det M \neq 0$ which in this case we see means Wronskian(at $x=\xi$) $\neq 0$. But we know this is true since u_1 and u_2 are independent. Thus a solution $\begin{pmatrix} A \\ B \end{pmatrix}$ DOES exist! We then just write it out using $\begin{pmatrix} A \\ B \end{pmatrix} = M^{-1}\begin{pmatrix} C \\ D \end{pmatrix}$ and the result is 1.56, we are all done: we have constructed an explicit solution, and we already know there can only be one solution.

2. The Case of General BC's (67)

(a) initial value

I am changing the batting order here a little. Let's first ignore the "general BC" situation and try to show existence for the "initial value" BC case. Back on page 58 (as quoted above), we had a Theorem which claimed existence and uniqueness for Lu = f where I presume f had to be a non-generalized function. So here we are trying to show existence (we already know uniqueness) for Lu = δ. The problem is stated with these initial value BC's in 1.57. We take as our trial Green's solution page 67 A. It certainly meets both our BC's at the a end since u = 0 near that end. We then have to assure the usual Green's continuity and jump conditions are met as shown in p 67 B. But these are the same as p 66 B with A = -A. We make the same arguments and we see that since u_1 and u_2 are independent, we solve for A and B, and we obtain the final result (1.58) for our Green's. If we regard the long expression in 1.58 as $v_\xi(x)$, we recover our earlier

"causal" form on page 56 (lines are written in reverse order). And we know we can write this in the alternate t_2 form shown in 1.59a depending on where "a" lies, as discussed at length in the page 63 notes above.

So, we have now shown that, if Complete Homo has only the trivial solution (which makes u_1 and u_2 independent), we can construct an explicit solution 1.58 to the "initial value" version of the Green's Problem. [we are now at top of page 68].

(b) general case

Now how can we handle the "general" BC situation shown in 1.51? We are told to consult exercise 1.53 on page 79. In turn, on this little treasure hunt, that exercise refers us to that "causal fundamental solution" thing which was mentioned back on page 56. But this just means the Green's thing where you set u(x) = 0 -- this is where that was done for the first time. So our "hint" is that we should take as a trial Green's constructed solution for our "general BC case" the following form

$$g(x|\xi) = t_{causal}(x|\xi) + A\, w_1(x) + B w_2(x)$$

where NOW we have $w_1(x)$ solve homo Lu= 0 with general condition $B_1(w_1) = 0$ and w_2 does B_2 (and these are general, not end a and end b). Note that G is supposed to be the Lu=δ solution with the completely general (but still homo in that $B_i = 0$) BC's.

How do we know $w_1(x)$ exists? For now I just appeal to the fact that Lu=0 is supposed to have two independent solutions with two constants, and we can surely solve *one* equation $B_1(w_1) = 0$ to find a pair of constants that works. Similarly for w_2. And w_1 and w_2 are again independent because otherwise the Complete Homo would have $w_1 = k w_2$ as a solution, which violates our starting assumption.

At this point, a voyage down memory lane, back to page 54. Recall there that we wrote a general form for Greens as shown p 54 A where we started with some homo solution u(x), and found $v_\xi(x)$ more or less. Then in Example 1 page 56 we saw a case where we have two different ways to write the Green's which correspond presumably to two different ways to choose the starting homo solution u(x). And we took note there that the difference between these two solutions must be a solution to the homo equation. Back on that page, we were not concerned with any "boundary conditions", but now we are.

So, now our equation above makes more sense. I suppose we pick some solution like maybe $u_3(x)$ which we know satisfies the Lu=0 homo (without regard to BC's), and from that we construct the thing called $t_{causal}(x|\xi)$ above. I suppose we could take $u_3 = u_1$ if we wanted (and we don't worry that earlier our u_1 was selected to meet a certain a-end BC).

Now consider, based on the above expression for g:

$$B_1(g) = B_1(t) + A\, B_1(w_1) + B\, B_1(w_2) = B_1(t) + B\, B_1(w_2) = 0$$

$$B_2(g) = B_2(t) + A\, B_2(w_1) + B\, B_2(w_2) = B_2(t) + A\, B_2(w_1) = 0$$

Since we know t, we know $B_1(t)$. Since we know the w_i we compute $B_1(w_2)$. We know this is not zero otherwise w_2 would be a solution of the Complete Homo. So we find:

$$B = - B_1(t) / B_1(w_2)$$

$A = - B_2(t)/ B_2(w_1)$

So once you have determined w_1 and w_1 and t, you compute A and B as just shown, and then we have constructed our existence solution $g(x|\xi) = t_{causal}(x|\xi) + A\ w_1(x) + Bw_2(x)$. The continuity and jump conditions are taken care of for us by $t_{causal}(x|\xi)$.

So it would seem then that I have constructed a solution for the "general case", so why do we need to consider those special cases like "unmixed" or "initial value" ? I think I have "done enough" on this general subject in these "raw notes". I certainly believe the theorem, so let's just "move on"

Summary of Solutions of various ODE systems:

Page 68: we now consider several classes of ODEs:

(0) homo Lu = 0 with homo BC's $B_1 = 0$ and $B_2 = 0$ // the "complete homo problem"

(a) inhomo Lu=f with homo BC's $B_1 = 0$ and $B_2 = 0$

(b) inhomo Lu= f with inhomo BC's $B_1 = \alpha$ and $B_2 = \beta$ // the "complete inhomo problem"

(c) the eigenvalue problem $Lu = \lambda u$ with homo BC's $B_1 = 0$ and $B_2 = 0$

Case (0). This is 1.54 on page 65 and it *may* or *may not* have non-trivial solutions. The issue is to find the nullspace of L with the homo BC's, but nullspace is not yet mentioned in this chapter. If it has only the trivial solution (no nullspace in my language), then we know that $Lu = \delta$ with has a unique Green's solution even with our fully general BC's with $B_1=0$ and $B_2=0$. In this case, we are well on our way to knowing the solution to the inhomo with these same homo BC's:

Case (a): This is the case just mentioned. The solution is the "usual" form 1.61 since g already meets our homo BC's. The "proof" is given here that this form satisfies Lu=f and satisfies the homo BC's !!!

Case (b): Now same as (a) but use inhomo BC's so $B_1 = \alpha$ and $B_2 = \beta$. We take the same g which solved case (a), and we try for our solution $u(x) = 1.61 + c_1w_1(x) + c_2w_2(x)$ where the w_i are those same functions. [Aside: above we used these w_i in $g(x|\xi) = t(x|\xi) + A\ w_1(x) + Bw_2(x)$ in attempting to solve the *Green's* Problem $Lu=\delta$ with general homo BC's and we found A and B. Here we are using the w_i to solve not the Green's Problem but the *inhomo* problem Lu = f with *inhomo* BC's. It is a different problem, so we don't get the same A,B solution, but one which is similar.] This is the idea that we add these homo solutions in order to meet our new BC's $B_1 = \alpha$ and $B_2 = \beta$. If $\alpha = \beta = 0$, we know that the two c_i are 0. In general, we solve for the c_i and write the solution as in 1.63. This is done very simply in p 68 B and gives 1.63, very hard to argue with!

Case (c): This is just case (a) with $f = \lambda u$ and we get our integral equation for u.

At this point, I wrote meta notes for Chapter 1 up to this point, just to stabilize the foundation a bit.

Adjoint, Symmetric, and Self-Adjoint Systems (69)

We start off here computing $\int dx\ [\ vLu - uL^*v\ ]$ where L* is the usual thing we get doing parts integration to transfer all D's from u to v. As usual, $L = a_o D^2 + a_1 D + a_2$. If you write out L* long hand, you get as shown bottom of page 69. Now, when you do the parts integration, you pick up various parts which are collected into something called J(u,v) evaluated at b less at a. So we can write 1.65 and 1.66 which seem to have certain historical names.

Now, if you want to have L = L* just in the sense of operator equality, you can do it with $\mathbf{a_o' = a_1}$ which is pretty simple, and then you can write L as in p 70 A. In this sense of equality, J is said to be "**formally self-adjoint**".

Now having read Chapter 2, we know that for L = L* to be fully "self adjoint", they have to have the same domains and we need D = D*. This lets us simplify the alternative theorem and not have to put * on everything, and then we get nullspace of L + perp space of L and all that nice stuff.

Now let's skip over to page 71 and consider these two systems:

$Lg = \delta(x-\xi)$	$B_i(g) = 0$	space of all g ∋ $B_i(g) = 0$ is called D
$L^*h = \delta(x-\xi)$	$B_i^*(h) = 0$	space of all h ∋ $B^*_i(h) = 0$ is called D*

(1) Here the * on B_i just means these are "some other" boundary conditions appropriate for the "adjoint problem". This * does not mean CC, and it does not mean adjoint. So $B_i^*(h) = 0$ are the "adjoint boundary conditions".

(2) D is really the domain of the first problem, because all solutions must respect the BC's. We don't care about functions that don't respect $B_i(g) = 0$. And D* is the domain of the second problem. Notice that we think of ξ as some fixed number, and functions like $g(x|\xi)$ are treated as $g_\xi(x)$.

(3) g is called the Green's, and h is called the Adjoint Green's

(4) Now I have said so far that the BC's $B_i^* = 0$ come out of thin air. In fact, they have *connection* to the first set of BC's $B_i = 0$. They are not the same equations, but they have a connection. Here is the connection: if u is in D, and if v is in D*, then $J(u,v)|_a^b = 0$. So in general, you first define D as the space of functions such that $B_i(u) = 0$ and this is compatible with the Green's system above. Then you define space D* as follows: if u is in D, then if v is in D*, $J(u,v)|_a^b = 0$. All v's that make this true span D*. Stak shows with an example that the condition $J(u,v)|_a^b = 0$ GIVES YOU the two adjoint boundary conditions which I write above as $B^*_i = 0$. So these $B^*_i = 0$ are not arbitrary. For a given $B_i(u) = 0$ pair, the condition $J(u,v)|_a^b = 0$ gives you the $B^*_i(v) = 0$ pair. I guess this is not easy to show in general so he had to use an example bottom of page 70.

So we now assume that this is what D* means for the Adjoint Green's system. On page 71 without doing very much he shows then that $h(x|y) = g(y|x)$. So they are the same but you swap the two arguments. This is (1.72).

Now, if L = L* (formal) and if D* = D, then L is fully self-adjoint. In this case, the two systems shown above (Green's and Adjoint Green's) are really the same system so h(x,y) = g(x,y). But then from what we just showed, we know that g(x,y) = g(y,x). Thus:

Fact: The Green's Function is symmetric when L is fully self-adjoint.

This takes us to the pencil line on page 71 bottom part. We have worked on Green's, now we want to work on inhomo Lu = f. We are not assuming self adjoint anything here. By simple fiddling, he first shows page 71 A. This gives the solution of Lu=f as a strange integral of the h green's and some J stuff.

If $\alpha = \beta = 0$ in 1.74, then u is in D. But h is in D*. But then $J(u,v)|_a^b = 0$. Thus, we can write our solution u(x) as shown either way in page 72 A. The second is of course our usual formula.

Now, if α,β are not 0, then go back to p 71 A and write out as:

$$u(\xi) = \int dx\, h(x|\xi)f(x)dx - J\,(\,u(\xi),\, h(x|\xi)\,)\,|_a^b \;=\; \int dx\, g(\xi|x)f(x)dx - J\,(\,u(x),\, g(\xi|x)\,)\,|_a^{x=b}$$

and this is 1.75. In this case, due to 1.74, we cannot say that u is in D (since $\alpha,\beta \neq 0$). This is our "usual formula" with an extra term.

We now have a Stakgoldian long list of didactic Examples.

Example 1: We assume L = L* formally only, and

(a) I agree that Abel becomes (1.76) which is now very simple indeed: $a_o(x)W(u,v;x) = C$.

(b) Now I have to prove something that is not obvious to me at all:

Theorem: if $B_i = 0$ are the "pure" BC's, then D = D*.

Proof: If $B_i(u) = 0$, then we have to show that the equations $B_i^*(u) = 0$ are exactly the same two equations. So this forces me to do more work on the general problem of relating these two sets of BCs.

$B_1(u) = Au(a) + A'u'(a) = 0$
$B_2(u) = Bu(b) + B'u'(b) = 0$

Assume formal self-adjoint so that $J(u,v) = a_0(x)[\, v(x)u'(x) - u(x)v'(x)\,]$

We seek to make $J(u,v)|_a^b = 0$ which says

$$a_0(b)[\, v(b)u'(b) - u(b)v'(b)\,] \;-\; a_0(a)[\, v(a)u'(a) - u(a)v'(a)\,] \;= 0$$

What conditions on v and v' at the end points makes this be true? Here are two:

$a_0(b)[\ v(b)u'(b) - u(b)v'(b)\ ] = 0 \qquad (*)$
$a_0(a)[\ v(a)u'(a) - u(a)v'(a)\ ] = 0$

but we need conditions on v which don't depend on u because the space D* should be determined by constants and values of v only, and cannot be a function of u. But from our original conditions we have:

$u'(a) = -\ A/A'\ u(a)$
$u'(b) = -\ B/B'\ u(b)$

So our conditions (*) then become

$a_0(b)[\ v(b)\{ -\ B/B'\ u(b)\} - u(b)v'(b)\ ] = 0$
$a_0(a)[\ v(a)\{ -\ A/A'\ u(a)\} - u(a)v'(a)\ ] = 0$

$a_0(b)\ u(b)\ [\ v(b)\{ -\ B/B'\ \} - v'(b)\ ] = 0$
$a_0(a)\ u(a)\ [\ v(a)\{ -\ A/A'\} - v'(a)\ ] = 0$

So here then are two conditions which then make $J(u,v)|_a^b = 0$ if u is in D:

$v(a)\{ -\ A/A'\} - v'(a) = 0$
$v(b)\{ -\ B/B'\ \} - v'(b) = 0$

$v(a)\{ -\ A\} - A'v'(a) = 0$
$v(b)\{ -\ B\ \} - B'v'(b) = 0$

$Av(a) + A'v'(a) = 0 \qquad B_1^*(v) = 0$
$Bv(b) + B'v'(b) = 0 \qquad B_2^*(v) = 0$

So these last two equations are two condition s B_i^* which make $J(u,v)|_a^b = 0$ and which put the function v into the space D*. BUT, these are the same as conditions B_i, so D* = D. QED. So we have just proven this theorem:

Theorem 1: If L = L* formally, and if $B_i(u) = 0$ are pure BC's, then $B_i^*(u) = 0$ are the exact same equations, and then D = D*, and then L is fully self-adjoint.

While here, let's try to do the general case:

We won't assume formally self adjoint so now we have

$J(u,v) = a_o(x)\ [\ v(x)u'(x) - u(x)v'(x)\ ] + a_3(x)u(x)v(x) \qquad a_3 \equiv a_1 - a_o'$

Here are our two general case starting conditions.

$B_1(u) = Au(a) + A'u'(a) + Bu(b) + B'u'(b) = 0$
$B_2(u) = \alpha u(a) + \alpha'u'(a) + \beta u(b) + \beta'u'(b) = 0$

We seek to make $J(u,v)|_a^b = 0$ which says

$$a_o(b)\,[\,v(b)u'(b) - u(b)v'(b)\,] + a_3(b)u(b)v(b) - a_o(a)\,[\,v(a)u'(a) - u(a)v'(a)\,] - a_3(a)u(a)v(a) = 0$$

Now what? We want to get rid of the u's. So write the $B_i = 0$ equations as follows:

$$\begin{pmatrix} A' & B' \\ \alpha' & \beta' \end{pmatrix}\begin{pmatrix} u'(a) \\ u'(b) \end{pmatrix} = -\begin{pmatrix} A & B \\ \alpha & \beta \end{pmatrix}\begin{pmatrix} u(a) \\ u(b) \end{pmatrix} \qquad \begin{pmatrix} W & Y \\ w & y \end{pmatrix} \equiv -\left[\begin{pmatrix} A' & B' \\ \alpha' & \beta' \end{pmatrix}\right]^{-1}\begin{pmatrix} A & B \\ \alpha & \beta \end{pmatrix}$$

As long as $\det\begin{pmatrix} A' & B' \\ \alpha' & \beta' \end{pmatrix} \neq 0$, we can solve this for

$$\begin{pmatrix} u'(a) \\ u'(b) \end{pmatrix} = \begin{pmatrix} W & Y \\ w & y \end{pmatrix}\begin{pmatrix} u(a) \\ u(b) \end{pmatrix}$$

or

$$u'(a) = W\,u(a) + Y\,u(b)$$
$$u'(b) = w\,u(a) + y\,u(b)$$

where W,Y,w,y are all just constants, each a function of all 8 original constants. So here again is our required condition $J(u,v)|_a^b = 0$,

$$a_o(b)\,[\,v(b)\{u'(b)\} - u(b)v'(b)\,] + a_3(b)u(b)v(b)$$
$$- a_o(a)\,[\,v(a)\{u'(a)\} - u(a)v'(a)\,] - a_3(a)u(a)v(a) = 0$$

$$a_o(b)\,[\,v(b)\{w\,u(a) + y\,u(b)\} - u(b)v'(b)\,] + a_3(b)u(b)v(b)$$
$$- a_o(a)\,[\,v(a)\{W\,u(a) + Y\,u(b)\} - u(a)v'(a)\,] - a_3(a)u(a)v(a) = 0$$

$$u(a)\,\{\,w\,a_o(b)\,v(b) + a_o(a)\,v'(a) - a_3(a)\,v(a) - W\,a_o(a)\,v(a)\,\}$$
$$+ u(b)\,\{\,y\,a_o(b)\,v(b) - a_o(b)\,v'(b) + a_3(b)\,v(b) - Y\,a_o(a)\,v(a)\} = 0$$

Now we can take the two {...} as our two $B_i^* = 0$ conditions (which do not depend on u)

$$w\,a_o(b)\,v(b) + a_o(a)\,v'(a) - a_3(a)\,v(a) - W\,a_o(a)\,v(a) = 0$$
$$y\,a_o(b)\,v(b) - a_o(b)\,v'(b) + a_3(b)\,v(b) - Y\,a_o(a)\,v(a) = 0$$

$$-[\,a_3(a) + W\,a_o(a)]\,v(a) + [a_o(a)]\,v'(a) + [w\,a_o(b)]\,v(b) = 0$$
$$-[-a_3(b) + y\,a_o(a)]\,v(b) - [a_o(b)]\,v'(b) - [Y\,a_o(a)]\,v(a) = 0$$

$$[\,a_3(a) + W\,a_o(a)]\,v(a) - [a_o(a)]\,v'(a) - [w\,a_o(b)]\,v(b) = 0 \qquad B_1^*(v) = 0$$
$$[\,a_3(b) - y\,a_o(a)]\,v(b) - [a_o(b)]\,v'(b) - [Y\,a_o(a)]\,v(a) = 0 \qquad B_2^*(v) = 0$$

There is an asymmetry here in that each of these equations is missing one of the four terms, but I think the math is correct. I have just proven the following theorem:

Theorem 2: Given order-2 operator L with BC's $B_i(u) = 0$ for any equation Lu=f, we know that the conditions $B_i(u) = 0$ define a space $u \in D$. If we then define space $v \in D^*$ by $J(u,v)|_a^b = 0$ assuming that $u \in D$, we find that the two conditions $B_i^*(v) = 0$ shown above put $v \in D^*$. In general, the two pairs of equations are very different, and we have $D \neq D^*$. In any event, given the $B_i(u) = 0$, we know from above exactly how we can find the $B_i^*(v) = 0$ conditions. We did have to make a certain det(...) $\neq 0$ assumption about some of the coefficients.

Special Case. How suppose we apply the above Theorem to the case L = L* formally. Then $a_3 = 0$ and the above conditions become

$$[\ W\ a_o(a)]\ v(a)\ -[a_o(a)]\ v'(a) -\ [w\ a_o(b)]\ v(b)\ = 0 \qquad B_1^*(v) = 0$$
$$[\ y\ a_o(a)]\ v(b)\ -[a_o(b)]\ v'(b) -\ [Y\ a_o(a)]\ v(a)\ = 0 \qquad B_2^*(v) = 0$$

Things are not simplified very much here.

Example 1 continued. Use our new Abel rule in 1.56 which we developed for the pure boundary condition situation. The denominators are both replaced with just C and we get the very simple 1.77 for our constructed Green's Function $g(x|\xi)$, where u_1 and u_2 were the two functions we constructed. The result is blatantly symmetric, which fits with our theorem that self-adjoint L have symmetric g. We then use our result (1.61) to write out the solution to Lu=f with our $B_i(u) = 0$ homo BC's. So this was a pretty useful example I would say. It concerned *any* L=L* with pure homo BC's and we saw how this problem has L = *fully* self-adjoint. However, we must assume that the complete homo has no solutions.

Example 2. Here $L = D^2$ and the BC's are u(a) = 0 and u'(b) = 0 where $a,b = 0,\ell$. The complete homo has no solutions. From pencil page 69 bottom, we see that L = L* formally, and we have pure BC's so we know this L is fully self-adjoint. The Green's is then given by 1.77 and we have $a_0 = 1$. The result is the same as what Stak finds by a simple brute force solution, which is p 73A, clearly symmetric. It is easy to graph this Green's function, but I don't have any string-style physical interpretation.

Example 3. Here we do some "initial value" BC stuff at endpoint a = 0. So our $B_i(u) = 0$ are of the initial-value form. He does a little trick to show that the $B_i^*(v) = 0$ equations are the same as the $B_i(u) = 0$ except they are evaluated at the right endpoint. Thus, the $B_i^*(v) = 0$ are different, so $D \neq D^*$. In general, he claims, no initial-value homo BC problem can be self-adjoint.

Digression: What can I say in general about initial-value conditions?

$$J(u,v) = a_o(x)\ [\ v(x)u'(x) - u(x)v'(x)\ ]\ + a_3(x)u(x)v(x) \qquad a_3 \equiv a_1 - a_o'$$

$$u(a) = 0 \qquad B_1$$
$$u'(a) = 0 \qquad B_2$$

We thus have our D defined. We seek D* by looking at $J(u,v)|_a^b = 0$

$a_o(b)\,[\, v(b)u'(b) - u(b)v'(b)\,] + a_3(b)u(b)v(b) - a_o(a)\,[\, v(a)u'(a) - u(a)v'(a)\,] - a_3(a)u(a)v(a) = 0$

$a_o(b)\,[\, v(b)u'(b) - u(b)v'(b)\,] + a_3(b)u(b)v(b) = 0$

$$[-a_o(b)\,v'(b) + a_3(b)v(b)]\,u(b) + [a_o(b)\,v(b)]\,u'(b) = 0$$

$\Rightarrow$ $v(b) = 0$ and then $v'(b) = 0$, just as he said, I just repeated it here in my notation.

So regardless of whether L is formally self-adjoint or not, for "homo initial value BC's" we know that $D \neq D^*$. So even if $L = L^*$, we have $D \neq D^*$. Thus, I agree with his claim

Theorem: No L having initial-value homo BC's is self-adjoint.

Suppose we have inhomo initial value? Then

$u(a) = \alpha$ $\qquad$ $B_1 = \alpha$
$u'(a) = \beta$ $\qquad$ $B_2 = \beta$

$a_o(b)\,[\, v(b)u'(b) - u(b)v'(b)\,] + a_3(b)u(b)v(b) - a_o(a)\,[\, v(a)\beta - \alpha v'(a)\,] - a_3(a)\alpha v(a) = 0$

$$u'(b)\{\, a_o(b)\,v(b)\,\} + u(b)\,\{-a_o(b)\,v'(b) + a_3(b)\,v(b)\} = a_o(a)\,[\, v(a)\beta - \alpha v'(a)\,] + a_3(a)\alpha v(a)$$

Now we cannot eliminate u and there are no $B_i^*(v) = k_i$ that don't involve u. So I guess we never talk about inhomo BC's in this context.

Example 3 continued. In our reading, we have dealt with the initial-value problem already in two places:

First was page 63: the Theorem gives the Volterra form for the solution in terms of $v_\xi(x)$ where this function is a solution of Lu=0 with its own little initial values as shown in 1.50. [We know such a problem has a unique solution $v_\xi(x)$].

The second place is page 67 where we are looking at the Green's problem with initial conditions as in 1.57. Our initial value BC's are homo, so result (1.61) applies to our Lu=f problem. This with (1.59) gives us the Volterra solution p 73 B, but this is the same as 1.50 which also states the IV conditions on $v_\xi(x)$.

So up to this point this example has shown that : (a) no initial-value homo BC problem can be self-adjoint: (b) We can write the solution to Lu = f in our Volterra form p 73 B.

Now he wants to treat the non-homo initial-value BC situation with α and β. The result we already know from (1.63) where u_1 and u_2 are homo solutions with certain partial BC's. But here he redefines these two functions slightly so we have 1.63 now looking like p 73 C with p 74A. Each set of conditions here is initial value, so we know u_1 and u_2 exist as solutions to Lu = 0. It is then easy to directly verify that u(x) satisfies the two BC's with α and β. Recall the general idea that you can always add homo solutions to a Green's solution in trying to meet the BC's.

So he has now done (c) find solution to Lu=f with inhomo initial-value BC's α,β.

There is more to go still. Suppose now, in addition to the specifications for case (c) above, we also have L = L* formally so we can use the Abel thing for the Wronskian. It happens that our p 74A

conditions cause $W(u_1,u_2; 0) = 1$ since $\det \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1$. But we know $W(x)a_0(x) = C$ from 1.76, so we then know that $W(u_1,u_2; 0)\ a_0(0) = C = W(u_1,u_2; x)\ a_0(x) = a_0(0)$ so install this into our initial-value Green's solution 1.58 where the denominator then becomes just $a_0(0)$ and we then have result p 74 B for $v_\xi(x)$. Then our solution for Lu=f which we just wrote as p 73 C can be made more explicit in (1.79).

So this says that, (d) given an arbitrary formally self-adjoint L problem with inhomo initial-value BC's, we first solve the homo Lu=0 for u_1 and u_2 subject to IV conditions p 74A, and then our full solution to Lu=f is given by 1.79. Another chunk-style example!

Example 4.

Digression to Appendix A.3 on Heat Conduction

Page 328 A has a heat flow balance equation for a 3D region of space. Some "source" adds or removes heat according to a function f(**r**,t) which is per volume and per time. During a Δt of time, the LHS is the heat added to the volume. The rightmost term is the heat flow out through the bounding surface, and the middle term is the increase in temperature of the volume due to the heat capacity. So basically

heat added = increase in stored energy - flow out through the boundary

Using Gauss's Law as in B, we end up with balance equation C. This applies to any volume, and in particular to a differential volume, we end up with a PDE for "heat conduction" which is A.10. At each point in space, there is some thermal conductivity k(**r**,t), and there is some heat capacity c(**r**,t) of whatever the material is. Result A.10 is the general result allowing that k and c are functions of space and time. If either or both of these are constants, we get a simpler equations.

I have to observe in passing that, although I have a Ph.D. in Physics, I don't remember ever working with this heat conduction equation. Yes, thermal conductivity is mentioned in passing in Zemansky and Reif, but these books don't operate at the PDE math level. It's not "modern physics", it is classical physics, but not mechanics, not E&M, might appear in plasma or star books. [Amazon has lots of books on Heat Conduction, it is a whole field in itself.]

Here is the equation:

$$\partial u/\partial t - 1/c\nabla\bullet(k\nabla u) = f/c \qquad (*) \qquad u = u(\mathbf{r},t), \text{ same for k, c, f}$$

where f,k,c,u are all scalar fields, u = temperature. If k = constant, you get $k\nabla^2 u$ and this thing becomes the **diffusion equation** which is a little better known to me, but only slightly. Stakgold deals with some of these things in his Volume 2. In particular, a static solution for k=constant looks like this:

$$\nabla^2 u = (1/k)\, f \qquad \text{compare to } \textbf{Poisson} \text{ from electrostatics:} \qquad \nabla^2\varphi = -4\pi\rho$$

So here we have three classical physics equations that are really the same: diffusion, electrostatics and heat conduction. I think they are all based on the continuity equation. The functions you solve for in these three cases are: matter density, electrostatic potential, and temperature. The sources in the three cases are matter source, charge density, and heat source. Coefficients are diffusion coefficient, "unity", and thermal conductivity.

Example 4 Continued:

Our problem is a thin rod which we treat as a 1D steady-state problem in dimension x. We apply a heat source f(x) along the rod, and we want it to be in steady state, so $\partial u/\partial t = 0$. The heat capacity c cancels out in (*) above, and the fancy term in the heat conduction equation in 1D becomes

$$\nabla\bullet(k\nabla u) = \partial_x (k\nabla u)_x = \partial_x (k\,\partial_x u) \qquad \Rightarrow \qquad -\partial_x [k(x)\,\partial_x u(x)] = f(x) = L\,u(x)$$

So here is an actual real-world ODE problem for us to ponder. Looking first at p 70 A, we see that L is formally self-adjoint and we have $a_o(x) = -k(x)$, the thermal conductivity function, and $a_2(x) = 0$. For boundary conditions, we force each end of the bar to be at "0 temperature" so u(0) = 0 and u(1) = 0. These are pure BC's, so our ODE system is fully self-adjoint. The system for the Green's is shown in C. Our problem falls exactly under the conditions of Example 1, g is as given there in terms of u_1 and u_2 where u_1 satisfies $u_1(0) = 0$ and u_2 satisfies $u_2(1) = 0$. The homo equation is really $(k\,\partial_x u) = K$, a constant. Stak solves simply for $u_1(x)$ and $u_2(x)$ as integrals of 1/k as shown in D (I did it all). We can then compute the Wronskian as shown in E and from this we find the Abel constant C and then the Green's is as shown in F. So, we have a complete Green's solution for this interesting 1D problem. We could then look at Lu=f with the same homo pure BC's and the answer would be the integral in (1.61), "the usual". We could then vary the ending temperatures to be α and β and then have something like (1.63) but I might have to tune the constants shown there.

Example 5. (p75) Let's take a gander at the normal Bessel equation (Schaum p 136)

$$L' = x^2D^2 + xD + (x^2 - n^2) \qquad L'u = 0$$
$$= xD(xD) + (x^2 - n^2)$$

$$L = L'/x = D(xD) + (x - n^2/x) \qquad \text{note that } L'u=0 \Rightarrow Lu=0 \text{ away from } x=0.$$

Set n = 0 (order = 0) and then we have L = D(xD) + x and this is our example. It is formally self-adjoint as per p 70 A. We pick (a,b) to the right of x=0 to avoid potential problems at x=0. And we apply the unmixed u=0 at the two endpoints, so we have a Green's problem p 75 A. The pure BC's mean we are fully self-adjoint.

Digression: Appendix B.1 and B.2 on Bessel Functions

B.1 is exactly Lu = 0 as I wrote it above. A very clear short appendix, J_α = Bessel, N_α = Neumann, H = Hankel's, rules for picking two independent solutions.

Since we are self-adjoint, we can use the simple Abel thing as shown p 330 A. We know the LHS is independent of z, so compute at small positive z where series are simple. I agree with B and C (see scribble on bottom of previous page), so I like all three Wronskian results given here. These Wronskians have always seems foreign to me, but here we are putting them to use!

Example 5 continued: Stak writes two functions Z_1 and Z_2 as usual where Z_1 satisfies the left BC and Z_2 the right one, shown in B. We are again in the bailiwick of Example 1 and the solution has the form

shown in p 75 C. We know from 1.76 that C = a_o W(Z_1,Z_2) and a_o = x here. But all we need to so this is W[J,N] which we just derived in the Appendix as B.7, so I agree with p 75 D. Therefore, our constant C is as shown in E = xD.

Again, we assume that Lg=0 with homo BC's has no solutions, but interestingly, if a and b take on certain values such as both roots of J_0, then there are non-trivial solutions just as there would be in the trig world for $D^2u = 0$.

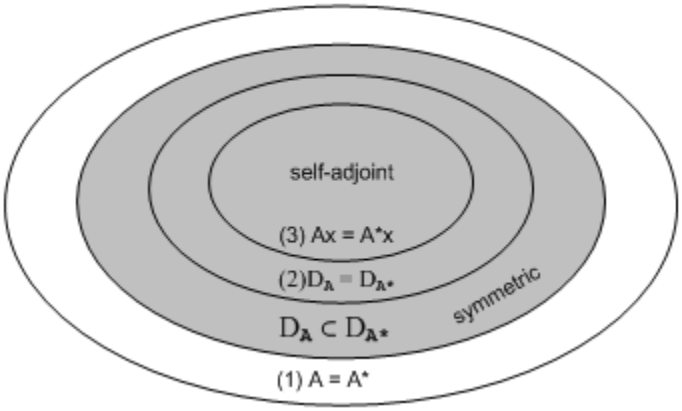
Comment on inner product. Looking back at 1.65, called Green's Formula, we have

$$\langle v,Lu\rangle = \langle L^*v,u\rangle + J(u,v)|_a^b$$

where the last term is in general not zero for some arbitrary u,v,L you pull out of your hat. However, if we insist that u ∈ D and v ∈ D* where D and D* were defined earlier (that is to say, D is all functions solving the BC's $B_i(u) = 0$, and then D* is all functions satisfying the adjoint ones $B^*_i(v) = 0$ which we then know makes $J(u,v)|_a^b = 0$), THEN for such u and v, we have <v,Lu> = <L*v,u>. In chapter 2, we refer to (u,v) as an admissible pair. D is the domain of L and D* is the domain of L*.

So we conclude that you can write <v,Lu> = <L*v,u> as long as u,v are an admissible pair, which means that u ∈ D and v ∈ D*.

Example 6: This then brings up the subtle **distinction** between **symmetric** and **self-adjoint**, and in Chapter 2 I made this picture:



where A = L, $D_A = D_L$= D and $D_{A^*} = D_{L^*}$ = D*. The outermost oval corresponds to the "formally self-adjoint" situation which means <Ax,y> = <x, Ay> for admissible pairs. If D* contains D, then L would be symmetric. You would describe this situation by saying that $J(u,v)|_a^b$ = 0 for all u and v in D. This might also be 0 for some u ∈ D and v ∈ D* with v NOT in D, but we don't care about that for "symmetric". The main idea of "symmetric" is that you have a nice little world D and for x,y in this little world D, you can say <Ax,y> = <x,Ay>. For math people, <Ax,y> = <x, Ay> exhibits the "symmetry".

But we can write also <y,Ax>* = <x,Ay>. In matrix notation we would then write A_{yx}^* = A_{xy} and then we have our famous Hermitian matrix, which math people refer to as symmetric even when there is a *.

As the oval picture above shows, it is tougher to be self-adjoint than to be merely symmetric. First, you need to have D = D*. And you need L = L* when acting on any f(x) in D. But for differential

operators, since L = L* exactly as an operator, condition (3) is nothing new, so perhaps we would dispense with the inner most oval and call self-adjoint the next outer one just with D = D*.

In Example 6, Stak thinks about $L = D^2$ with a very restrictive BC set which cranks down so hard on D that D = φ. From the alternative theorem, and since L is formally self-adjoint, we know that this means D* = entire space. So this is an example of $D \subset D^*$ and then L is "symmetric" . But not a very exciting example if L can only act on an empty space D. Too bad he did not have a better example, I bet he added this Example 6 late in the game just to have *something* with $D \subset D^*$.

Exercises: (76)

Exercise 1.37.

As part of this exercise, we redo the General Case BC analysis done above in a different way.

We won't assume formally self adjoint so now we have (ie, keep $a_3 \neq 0$)

$$J(u,v) = a_o(x)\,[\, v(x)u'(x) - u(x)v'(x)\,] + a_3(x)u(x)v(x) \qquad a_3 \equiv a_1 - a_o'$$

Here are our two general case starting conditions.

$$B_1(u) = Au(a) + A'u'(a) + Bu(b) + B'u'(b) = 0$$
$$B_2(u) = \alpha u(a) + \alpha' u'(a) + \beta u(b) + \beta' u'(b) = 0$$

Let's write these equations using two matrices as follows: (see later for relation to α_{ij} of Stak)

$$\begin{pmatrix} A & A' \\ \alpha & \alpha' \end{pmatrix}\begin{pmatrix} u(a) \\ u'(a) \end{pmatrix} = -\begin{pmatrix} B & B' \\ \beta & \beta' \end{pmatrix}\begin{pmatrix} u(b) \\ u'(b) \end{pmatrix} \qquad \text{or} \qquad A\begin{pmatrix} u(a) \\ u'(a) \end{pmatrix} = -B\begin{pmatrix} u(b) \\ u'(b) \end{pmatrix}$$

$$\text{or} \qquad A\,\mathbf{a} = -B\,\mathbf{b} \qquad \Rightarrow \qquad \mathbf{a} = -A^{-1}B\,\mathbf{b}$$

where now A and B are 2x2 matrices, and **a** and **b** are column vectors in the obvious manner.
Define a new matrix C as follows:

$$C = \begin{pmatrix} C & C' \\ c & c' \end{pmatrix} \equiv -A^{-1}B \qquad \Rightarrow \qquad \mathbf{a} = C\,\mathbf{b}$$

$$u(a) = C\,u(b) + C'\,u'(b)$$
$$u'(a) = c\,u(b) + c'\,u'(b)$$

We seek to make $J(u,v)|_a^b = 0$ which says

$$a_o(b)\,[\, v(b)u'(b) - u(b)v'(b)\,] + a_3(b)u(b)v(b) - a_o(a)\,[\, v(a)u'(a) - u(a)v'(a)\,] - a_3(a)u(a)v(a) = 0$$

Let's get rid of u(a) and u'(a) using the above formulas:

$$a_o(b)\,[\,v(b)u'(b) - u(b)v'(b)\,] + a_3(b)u(b)v(b)$$
$$- a_o(a)\,[\,v(a)u'(a) - u(a)v'(a)\,] - a_3(a)u(a)v(a) = 0$$

$$a_o(b)\,[\,v(b)u'(b) - u(b)v'(b)\,] + a_3(b)u(b)v(b)$$
$$- a_o(a)\,[\,v(a)\{\,c\,u(b) + c'\,u'(b)\} - \{\,C\,u(b) + C'\,u'(b)\}v'(a)\,] - a_3(a)\{\,C\,u(b) + C'\,u'(b)\}v(a) = 0$$

$$u(b)\,\{ - a_o(b)\,v'(b) + a_3(b)\,v(b) - a_o(a)\,c\;\; v(a) + a_o(a)\,C\,v'(a) - a_3(a)C\,v(a)\}$$
$$+ u'(b)\,\{\,a_o(b)\,v(b) - a_o(a)\,c'\,v(a) + a_o(a)\,C'\,v'(a) \;- a_3(a)\,C'\,v(a)\,\} = 0$$

Now this suggests taking our B_i^* as the two bracketed quantities:

$$- a_o(b)\,v'(b) + a_3(b)\,v(b) - a_o(a)\,c\;\; v(a) + a_o(a)\,C\,v'(a) - a_3(a)C\,v(a) = 0$$
$$a_o(b)\,v(b) - a_o(a)\,c'\,v(a) + a_o(a)\,C'\,v'(a) \;- a_3(a)\,C'\,v(a) = 0$$

Rewrite as:

$$B_1^*(v) = [- a_3(a)C - a_o(a)\,c]\,v(a) + a_o(a)\,C\,v'(a) + a_3(b)\,v(b) - a_o(b)\,v'(b) = 0$$
$$B_2^*(v) = [- a_3(a)\,C' - a_o(a)\,c']\,v(a) + a_o(a)\,C'\,v'(a) + a_o(b)\,v(b) = 0$$

and these are then our adjoint boundary conditions $B_i^*(v) = 0$ which we can compare to our starting boundary conditions $B_i(u) = 0$ which were

$$B_1(u) = Au(a) + A'u'(a) + Bu(b) + B'u'(b) = 0$$
$$B_2(u) = \alpha u(a) + \alpha'u'(a) + \beta u(b) + \beta'u'(b) = 0$$

Now let's go to the L = L* case where we know that $a_3 = 0$. Then we have

$$B_1^*(v) = - a_o(a)\,c\,v(a) + a_o(a)\,C\,v'(a) - a_o(b)\,v'(b) = 0$$
$$B_2^*(v) = - a_o(a)\,c'\,v(a) + a_o(a)\,C'\,v'(a) + a_o(b)\,v(b) = 0$$

We can write these in matrix notation as (first change signs in the lower equation above)

$$a_o(a)\begin{pmatrix} -c & C \\ c' & -C' \end{pmatrix}\begin{pmatrix} v(a) \\ v'(a) \end{pmatrix} = a_o(b)\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} v(b) \\ v'(b) \end{pmatrix}$$

where we recognize our friend σ_1 sitting there. Mult from the left by σ_1 to get

$$a_o(a)\begin{pmatrix} c' & -C' \\ -c & C \end{pmatrix}\begin{pmatrix} v(a) \\ v'(a) \end{pmatrix} = a_o(b)\begin{pmatrix} v(b) \\ v'(b) \end{pmatrix}$$

Now go look up the rule for inverting a 2x2 matrix:

$$C = \begin{pmatrix} C & C' \\ c & c' \end{pmatrix} \qquad \Rightarrow \qquad \det(C)\,C^{-1} = \begin{pmatrix} c' & -C' \\ -c & C \end{pmatrix}$$

Thus our equation above becomes

$$a_o(a)\,\det(C)\,C^{-1}\begin{pmatrix} v(a) \\ v'(a)\end{pmatrix} = a_o(b)\begin{pmatrix} v(b) \\ v'(b)\end{pmatrix}$$

$$\begin{pmatrix} v(a) \\ v'(a)\end{pmatrix} = [a_o(b)/a_o(a)\det(C)]\; C\begin{pmatrix} v(b) \\ v'(b)\end{pmatrix}$$

$$\mathbf{a} = [a_o(b)/a_o(a)\det(C)]\; C\,\mathbf{b}$$

But our original BC's were **a** = C **b**. Thus, the final and initial BC's will be the same iff

$$[a_o(b)/a_o(a)\det(C)] = +1$$

But from above we have C = $-A^{-1}B$ so that AC = -B and det(A)det(C) = det(-B) = det(B). Thus, our condition becomes

$$a_0(b) = a_0(a)\,\det(C) = a_0(a)\,\det(B)/\det(A) \quad \Rightarrow \quad a_0(b)\,\det(A) = a_0(a)\,\det(B)$$

This is the conclusion of Exercise 1.37 in fact, and we have

$$A = \begin{pmatrix} A & A' \\ \alpha & \alpha'\end{pmatrix} = \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22}\end{pmatrix} = \alpha \qquad B = \begin{pmatrix} B & B' \\ \beta & \beta'\end{pmatrix} = \begin{pmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22}\end{pmatrix} = \beta$$

$$a_0(b)\,\det(\alpha) = a_0(a)\,\det(\beta)$$

Exercise 1.38. A routine drill to compute the $B_i^*(v)$ for a certain problem.

Exercise 1.39. Considers L with n ≠ 2. For real a_i functions, it is true that when n = odd, you cannot have L = L*. We know the general formula for L* from 1.36 on page 52. The leading term in the sum shown there is k = n so is $(-1)^n D^n(a_o(x)\varphi)$ and this contains the term $(-1)^n a_o(x)D^n\varphi$ which cannot agree with the leading term of L if n is odd -- we have a minus sign. Of course in QM we let the a_i be complex and this is then no longer true, such as for momentum.

Let's do part of this. Suppose L = D and u(0) = α u(1). We know that in this case we have $a_0 = 0$, $a_1 = 1$, $a_2 = 0$. Since $a_o' \neq a_1$, we know L is not formally self-adjoint, and in fact L* = -L. We can write $a_3 = a_1 - a_o' = 1$. We seek to make $J(u,v)|_a^b = 0$ which says

$$a_o(b)\,[\,v(b)u'(b) - u(b)v'(b)\,] + a_3(b)u(b)v(b) - a_o(a)\,[\,v(a)u'(a) - u(a)v'(a)\,] - a_3(a)u(a)v(a) = 0$$

or

$$+u(b)v(b) - u(a)v(a) = 0 \quad \Rightarrow \quad +u(1)v(1) - u(0)v(0) = 0$$

$$\Rightarrow +u(1)v(1) - \alpha\,u(1))v(0) = 0 \quad \Rightarrow \quad u(1)\,\{\,v(1) - \alpha\,v(0)\,\} = 0$$

So D* is the set of functions with $v(0) = \alpha^{-1}\,v(1)$ which is not the same as D.

Exercise 1.40. Claims you can find s(x) such that sL = formally adjoint.

Exercise 1.41. Seems obvious that (L*)* = L. Then $D_{(L^*)^*} = D_L = D$.

Exercise 1.42. A solid compute-everything exercise for a specific L.

Exercise 1.43 - 1.50. Typically we are given some L and some BC's, and we are supposed to find the Green's Function and the adjoint boundary conditions and maybe the adjoint Green's Function.

Exercise 1.51. This concerns "deflections of a beam", something I would like to learn about. // I have now learned about beams, see separate beams.doc which has lots of stuff in it.

So in this problem, our beam is clamped at both ends. An isolated unit shear force would then be $\delta(x-\xi)$ and we could hopefully solve the 4th order ODE for the Greens $g(x|\xi)$.

Since M' = V, we could represent a "unit torque" as $M = \delta(x-\xi)$, but since M' = V this would be equivalent to a unit shear force of $\delta'(x-\xi)$.

We should think of ξ as our Exercise 1.50 parameter. So imagine we solve the problem $Lg = \delta(x-\xi) = f(x; \xi)$. Then 1.50 shows that $L[\, dg/d\xi] = df/d\xi = -\delta'(x-\xi)$, or $L\{ - dg/d\xi \} = \delta'(x-\xi)$. Therefore, the solution to the "unit torque" problem is $- dg/d\xi$ where g is the solution to the unit shear force problem.

Exercise 1.52. The Green's function for a δ source is called the **impulse response** by EE's when the variable is time. This problem discusses the related **step response** where the source is $H(t-\tau)$ and the response function he calls $Z(t|\tau)$.

Exercise 1.53. We consider an arbitrary L with some general homo BC's like 1.51. We know we can write the Green's as any Greens (regardless of BC's) + a homo solution with constants, and then we adjust those constants. Here we are supposed to solve this general situation and use the simple-forced causal Green's as our Green's to start with. I think I could do this problem.

1.6 Alternative Theorems and Modified Green's Functions (79)

System of n equations in n unknowns: the matrix problem of order n. (79)

This is just a preview of Chapter 2 concepts. If A is a real matrix with rank = n-k, then nullity = k, and both A and A* have nullspaces of dimension n. Then Au = f only has a solution if f is in the perp space of A. That means that f must be orthogonal to all vectors which span the nullspace, $A\varphi_i = 0$. The term consistency condition is used.

Example: A 3x3 matrix (n=3) has a 2x2 nonzero subdeterminant so rank = 2, nullity = 1. The only non-trivial nullspace solution is found to be vector(1,-1,1). But vector f = (1,3,1) is not orthogonal to the nullspace factor, so f is not in the perp space of A, so there is no solution of Ay = f.

Linear Integral Equations. (81)

Here we look at $\lambda Ku = u$ and also $\lambda K^*v = v$, which I think of as the EV and adjoint EV equations. But swing over to the left so that $(\lambda K-1)u = 0$ and $(\lambda K^*-1)v = 0$ so then for operator $(\lambda K-1)$, this is a homo

and an adjoint homo equation. Let B = λK-1, for example, so Bu = 0 and B*v = 0. Then consider the inhomo equation Bu = f. He now states "**the alternative theorem**" (but does not call it Fredholm's alternative): Here are the two alternatives, (a) or (b):

(a) Bu=0 has no solutions, so neither does B*v=0. In this case, Bu=f has a unique solution for any f.

(b) Bu=0 has solutions, then so does B*v = 0, and Bu=f has many solutions if f is in B perp and none otherwise.

This just says : the nullity is 0, or the nullity is > 0.

Second Order ODEs (82)

[p 82] He starts off with the time-dependent string problem with a driving term. If we assume an $e^{-i\omega t}$ factor in the solution [ie, we pick some ω to drive with] , our ODE in just x becomes $L = -D^2-\omega^2$ where ω is a parameter, the assume frequency of the driving source. If Lu=0 has solutions (meaning ω^2 is an eigenvalue), then we have a nullspace, so we will have consistency conditions on our driving force function F(x) as in 1.86 -- one such condition in fact for each eigenvalue. The eigenvalues we know are resonances.

[p 83] For ω^2 = an eigenvalue, we know that $\omega_k = k\pi/L$ and solution of $Lu_k = 0$ is $\sin(k\pi x/L)$. He likes to call the eigensolutions $Z_k(x)$ or $Z_\omega(x)$. So in mid-page, he states the Alternative Theorem for our operator L. If ω is such that we are not on a resonance, then there is no nullspace, and we have a unique solution. Otherwise there is no solution unless we meet the consistency requirement that F(x) be in the perp space of L_ω. If we are at some $\omega = \omega_k$, it is not clear that our solution has the $e^{-i\omega t}$ form we have assumed. We expect a solution which grows in time and blows up somehow.

At the pencil line we start really solving this problem. Since string is glued at the origin, we can have only Fourier sine terms, so we could expand as in 1.88, replacing variable x with Fourier variable k, and then the inversion is as shown. This is a Fourier sine series, not a Fourier transform, by the way. We imagine that our string is repeated along the x axis so we have a periodic thing appropriate to a Fourier Series.

[p 84] We apply our Fourier series inversion formula p 83 A to the entire partial ODE in 1.87. Then F(x) becomes F_k in his notation, and u(x) becomes a_k. This gives p 84 B. We then do parts twice to move D^2 over onto the sine factor, where it produces $-(k\pi/L)^2$ and leaves us with an integral which is a_k. We then end up with C. We have Fourier'd on x, not on t.

Where does he get the BC's shown in C for the $a_k(t)$? Looking at p 83 A, I think to get these he has to assume that the string is flat at t=0 and has no velocity either. So maybe it is completely at rest, and we suddenly turn on our f(x,.t) at t=0. He has fudged this and just shows that BC's on $a_k(t)$ as initial value BC's.

But now we have an ODE in variable t, with initial-value conditions.

Review. In Example 3 of the last section, we studied the generic initial-value problem in fact with inhomo BC's α,β. We found a way to write the solution of Lu=f as the Volterra integral of $v_\xi(x)$ against f, plus some extra homo terms involving certain u_1 and u_2. We were then in fact able to express $v_\xi(x)$ in terms of u_1 and u_2 as in page 74 B. Then the complete solution was shown in B which was 1.79. We want to apply that solution to our current problem with time t as variable and with $\alpha = \beta = 0$, so we get just the first integral in 1.79.

Our equation of interest is p 84 C where $L = \partial_t^2 + c$ and we have $La_k = f = F_k e^{-i\omega t}$. For this problem, what are those functions u_1 and u_2 ? I think they are these:

$u_1(t) = \cos(k\pi t/L)$ $\quad u_1(0) = 1$ $\quad u_1'(0) = 0$

$u_2(t) = \sin(k\pi t/L) * (L/\pi k)$ $\quad u_2(0) = 0$ $\quad u_2'(0) = k\pi/L * (L/\pi k) = 1$

which matches page 74 A. Then

$u_1(\xi)\, u_2(x) - u_1(x)\, u_2(\xi) = ?$ // change variable names $\xi \to \tau$, $x \to t$

$u_1(\tau)\, u_2(t) - u_1(t)\, u_2(\tau) = (L/\pi k) \{ \cos(k\pi\tau/L) \sin(k\pi t/L) - \cos(k\pi t/L) \sin(k\pi\tau/L) \}$

$= (L/\pi k) \{ \sin(k\pi t/L)\cos(k\pi\tau/L) - \sin(k\pi\tau/L)\cos(k\pi t/L) \}$

$= (L/\pi k) \sin(k\pi(t-\tau)/L)$

so then 179 with $a_0 = 1$ becomes

$$u(t) = (L/\pi k) \int_0^t d\tau \sin(k\pi(t-\tau)/L) F_k e^{-i\omega t}$$

$$a_k(t) = (LF_k/\pi k) \int_0^t d\tau \sin(k\pi(t-\tau)/L) e^{-i\omega\tau}$$ //which is p 84 D exactly,

$$= (F_k/\omega_k) \int_0^t d\tau \sin(\omega_k(t-\tau)) e^{-i\omega\tau}$$

$$= (F_k/\omega_k) \int_0^t d\tau (1/2i) \{ e^{i\omega k(t-\tau)} - e^{-i\omega k(t-\tau)} \} e^{-i\omega\tau}$$

$$= (F_k/\omega_k) (1/2i) \{ e^{i\omega kt} \int_0^t d\tau [e^{-i(\omega+\omega k)\tau}] - e^{-i\omega kt} \int_0^t d\tau [e^{-i(\omega-\omega k)\tau}] \}$$

$$= (F_k/\omega_k) (1/2i) \{ e^{i\omega kt} [-i(\omega+\omega_k]^{-1} [e^{-i(\omega+\omega k)t} - 1] - e^{-i\omega kt} [-i(\omega-\omega_k]^{-1} [e^{-i(\omega-\omega k)t} - 1] \}$$

$$= (F_k/\omega_k) (1/2) \{ e^{i\omega kt} (\omega+\omega_k)^{-1} [e^{-i(\omega+\omega k)t} - 1] - e^{-i\omega kt} (\omega-\omega_k)^{-1} [e^{-i(\omega-\omega k)t} - 1] \}$$

$$= (F_k/2\omega_k) \{ (\omega+\omega_k)^{-1} [e^{-i\omega t} - e^{i\omega kt}] - (\omega-\omega_k)^{-1} [e^{-i\omega t} - e^{-i\omega kt}] \}$$

$$= (F_k/2\omega_k) \{ (\omega+\omega_k)^{-1} [- e^{i\omega kt}] - (\omega-\omega_k)^{-1} [- e^{-i\omega kt}] + e^{-i\omega t} [(\omega+\omega_k)^{-1} - (\omega-\omega_k)^{-1}] \}$$

$$= (F_k/2\omega_k) \{ (\omega+\omega_k)^{-1} [- e^{i\omega kt}] - (\omega-\omega_k)^{-1} [- e^{-i\omega kt}] + e^{-i\omega t} [(\omega+\omega_k)^{-1} - (\omega-\omega_k)^{-1}] \}$$

$$= (F_k/2\omega_k) \{ [- e^{i\omega kt}/ (\omega+\omega_k) + e^{-i\omega kt}/ (\omega-\omega_k) + e^{-i\omega t}(-2\omega_k)/ (\omega^2 - \omega_k^2) \}$$

which agrees with **p 84 E**. Now take the limit that $\omega \to \omega_k$. We have to be *careful* with the last two terms, so write them out:

$$1/(\omega-\omega_k)\{ e^{-i\omega kt} - 2\omega_k e^{-i\omega t}/(\omega+\omega_K)\} = \{ e^{-i\omega kt}(\omega+\omega_K) - 2\omega_k e^{-i\omega t}\}/[(\omega-\omega_k)\,2\omega_k]$$

This is a 0/0 situation, so do $\partial_\omega/\partial_\omega$ to get

$$= [e^{-i\omega kt} + 2i\,\omega_k t\, e^{-i\omega t}]/2\omega_k = e^{-i\omega kt}/(2\omega_k) + i\,t\,e^{-i\omega t}$$

Then our result is

$$= (F_k/2\omega_k)\{ - e^{i\omega kt}/(2\omega_k) + e^{-i\omega kt}/(2\omega_k) + i\,t\,e^{-i\omega t}\}$$

$$= (F_k/2\omega_k)\{ [-1/(2\omega_k)](e^{i\omega kt} - e^{-i\omega kt}) + i\,t\,e^{-i\omega t}\}$$

$$= (F_k/2\omega_k)\{ [-1/(2\omega_k)]2i\sin(\omega_k t) + i\,t\,e^{-i\omega t}\}$$

$$= (F_k/2i\omega_k)\{ [1/(2\omega_k)]2\sin(\omega_k t) - t\,e^{-i\omega t}\}$$

$$= (F_k/2i\omega_k)\{ \sin(\omega_k t)/\omega_k - t\,e^{-i\omega kt}\}$$

which agrees with **p 84 G**. We are applying $\omega = \omega_k$ to our system, and the first term oscillates, but the second term blows up linearly in time! So if you hit a resonance, solution will blow up UNLESS it happens that the matching $F_k = 0$. But this is exactly our consistency condition p 83 B. I don't know how to interpret this in terms of the alternative theorem, however, maybe just a coincidence. If you have now power in the mode F_k, then of course you can't drive the mode and solution won't blow up. [the Tacoma Narrows bridge showed this linear increase in time]

[p 85] Yet another repetition. We write our three usual equations bottom of page 84, then once again we state the **Alternative Theorem** in terms of these equations. But this time we are stating it for a general n=2 linear differential operator L system (with homo BC's), whereas on page 83 it was stated in the context of an example. The only thing to note here is that Lu=0 could have a nullspace of dimension 0, 1 or 2. Thus, there could be 0,1, or 2 consistency conditions needed to put f of Lu=f in the perp space of L.

// here 4.16.09

Examples:

Example 1. If T is temperature in a bar, we know that heat flow is given by dQ/dt = k (dT/dx) since a temperature gradient drives heat flow. At the end of an insulated bar, dQ/dt through the end surface is 0, so we have dT/dx = 0 or, in this problem, y'(0) = 0 and y'(L) = 0, so self -adjoint. A solution is that the bar is at a constant temperature, ie, a solution to $-D^2y = 0$ with these BC's.

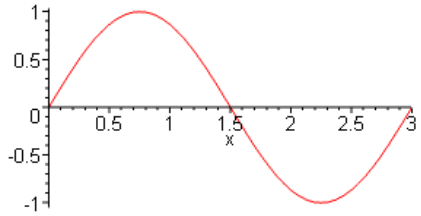
(1) physically, we can see that $\int dx\, f(x) = 0$ for static solution, since heat has nowhere to go.

(2) nullspace has nullity = 1, so range function f has have $f.\varphi_i = 0$ and φ_i = constant, so same result.

Stak goes on to solve this problem for a certain zero-area sine source. Here are my plots of the source, and of the resulting temperature with the constant A = 0. Note zero slope at both temp ends:

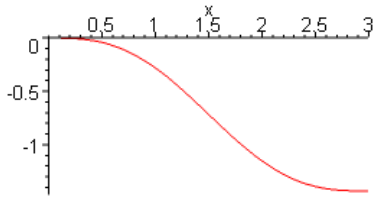
f(x) source

```
plot(sin(2*Pi*x/L),x=0..L);
```



T(x)

```
plot( ((L/(2*Pi))^2)*sin(2*Pi*x/L) -
L*x/(2*Pi),x=0..L);
```



Example 2. Here we are given an off-the-shelf L which is NOT formally self-adjoint. But BC's are unmixed, so we know that D = D* so the $B_i^* = 0$ are the same as the $B_i = 0$. We have constant coefficients and L and L* differ by the sign of the D term. Claim is that Lu=0 has solution $e^{-x}\sin(2x)$ and so has nullity 1. We know L*u=0 has same nullity, but solution will be different: $e^{+x}\sin(2x)$. If we look then at Lu=f, f in perp means that f(x) integrates to zero against the L*u=0 solution because that is what the rule means: perp space of L • nullspace L* = 0, or $R_L = N_{L^*}{}^{\perp}$. If f(x) meets this condition, you find a particular solution and then add multiple of the Lu=0 solution (but Stak does not solve for a particular solution).

Modified Green's Function

Go back to the Example 1 rod temperature problem above with y'(0) = 0 and y'(L) = 0, $L = -D^2$. As a review, because Lu=0 has u=C as a solution, we have consistency condition $\int dx\ f(x) = 0$ for Lu=f. But this then means that our Green's equation Lg = δ has no solution!!! Hay disastre! Stak nicely gives the physical interpretation of the problem: supply heat at a single point, it will have a certain sign, heat flow cannot stabilize since the bar is insulated, so there is no static solution. Very good. If we chose some other boundary conditions, we might have only the trivial for Lu=0 and then this problem would not arise, as in the case of our string problem.

More generally, we know that it is possible that given an L with some BC's, Lu=0 may have non-trivial solutions, ie, L has some nullspace. For n=2 ODE, this non-trivial nullspace can have dimension 1 or 2. We know then that the equation Lu=f will have no solution if f is not in the perp space of L, or will have many solutions if it is, many meaning particular + homo solutions. But the Green's problem is of the form Lu=f, and we write Lg = δ. So what happens if Lu=0 has non-trivial solutions, *and* δ is not in the perp space of L? Then there is no Green's function, so our whole Green's function method breaks down! Just to confirm what we are saying here, recall on page 66 the Theorem which says g exists and is unique *if* Lu=0 has no non-trivial solutions. The "repair" of this breakdown is the "modified" Green's function method.

Suppose u_1 is a non-trivial solution of Lu=0 where L is fully self-adjoint. Then Lu=f has a solution if $\int dx\ f(x)u_1(x) = 0$. Consider Lg = δ. Our g will only exist if $\int dx\ \delta(x-\xi)u_1(x) = u_1(\xi) = 0$ for all ξ, but this

would really mean $u_1 = 0$, but we are saying $u_1 \neq 0$, so we *always* have this malfunction of there are non-trivial solutions and we are talking Green's.

The Modified Green's method for self-adjoint L and nullity = 1

The proposed fix is to define an "associated" Green's problem which is this:

$$Lg_M = \delta(x-\xi) - u_1(x)u_1(\xi) = f \qquad \text{where } u_1 \text{ is normalized so that } \int u_1^2 dx = 1.$$

Here, we assume for the moment we have only nullity 1 which I think is common. Now consider

$$\langle f,u_1\rangle = \int dx\, [\delta(x-\xi) - u_1(x)u_1(\xi)]\, u_1(x) = u_1(\xi) - u_1(\xi) * 1 = 0$$

so now we have a "new f" in effect which meets our required consistency condition. This "new f" gives rise to some "new Green's function" which we call g_M.

So we now have two problems: (1) how do we solve for g_M ? (2) If we know g_M, how can we solve the equation Lu = f ?

Solving for g_M. If we integrate our g_M equation across the problem location $x = \xi$ (as usual), on the LHS we pick up our usual $a_o(\xi)$ jump $(Dg_m) = a_o(\xi)\,[\, g_m'(\xi+) - g_m'(\xi-)]$. On the RHS we get the usual "1" from the δ, and we pick up an additional $\int dx\, u_1(x)u_1(\xi) = u_1^2(\xi)\, dx \to 0$. We assume that $u_1(x)$ is smooth in our interval so this works. Thus, our g_M just has "the usual" jump condition, and it will also have the usual continuity at ξ. To construct g_M we go away from $x = \xi$ and we know that

$$Lg_m(x|\xi) = - u_1(x)u_1(\xi) \qquad x \neq \xi$$

and we somehow need to solve this equation subject to the BC's which are the same for all our little subproblems here, $B_i(g) = 0$. In general, this is in itself an ODE of the form Lu = f and we have to just solve it by some method to get $g_m(x|\xi)$. In the long example in the text, u_1 = constant C and $L = -D^2$ so this equation becomes $-D^2 g_m = -C^2$ where in this case $C = \sqrt{1/L}$. Then $g_M'' = 1/L$ which tells us that away from the problem point we have $g_M = A + Bx + x^2/2L$. We then have something to work with, and we can solve for the various unknown constants. But in the general case, I think you just have to do a brute force solution of the above equation, and Stakgold does not say much about this.

We do know that g_M will *not be unique* because we can always say

$$g_{m2}(x|\xi) = g_m(x|\xi) + A(\xi)\, u_1(x) \qquad \text{since } Lg_{m2} = L_{gm} .$$

so we always have (nullity = 1) one free constant floating around.

Symmetric g_M. The claim is that one of these many g_M solutions will be symmetric. On page 89 he claims we can cause this to happen by forcing $\int dx\, g_m(x|\xi)\, u_1(x) = 0$. Here is how you show this is true. First,

$$Lg_M(x|\xi_1) = \delta(x-\xi_1) - u_1(x)u_1(\xi_1)$$
$$Lg_M(x|\xi_2) = \delta(x-\xi_2) - u_1(x)u_1(\xi_2)$$

$$g_M(x|\xi_2)Lg_M(x|\xi_1) = g_M(x|\xi_2)\delta(x-\xi_1) - g_M(x|\xi_2)u_1(x)u_1(\xi_1)$$
$$g_M(x|\xi_1)Lg_M(x|\xi_2) = g_M(x|\xi_1)\delta(x-\xi_2) - g_M(x|\xi_1)u_1(x)u_1(\xi_2)$$

$$\int dx\, g_M(x|\xi_2)Lg_M(x|\xi_1) = \int dx\, g_M(x|\xi_2)\delta(x-\xi_1) - \int dx\, g_M(x|\xi_2)u_1(x)u_1(\xi_1)$$
$$\int dx\, g_M(x|\xi_1)Lg_M(x|\xi_2) = \int dx\, g_M(x|\xi_1)\delta(x-\xi_2) - \int dx\, g_M(x|\xi_1)u_1(x)u_1(\xi_2)$$

$$\int dx\, g_M(x|\xi_2)Lg_M(x|\xi_1) = g_M(\xi_1|\xi_2) - u_1(\xi_1)\int dx\, g_M(x|\xi_2)u_1(x)$$
$$\int dx\, g_M(x|\xi_1)Lg_M(x|\xi_2) = g_M(\xi_2|\xi_1) - u_1(\xi_2)\int dx\, g_M(x|\xi_1)u_1(x)$$

Now, if we assume $\int dx\, g_m(x|\xi)\, u_1(x) = 0$, we knock out the right term in each equation. If we can show that the LHS's are the same, then we have shown that g_M is symmetric! But L = L* since we have assumed L is self adjoint. So the LHS's are the same as long as the parts are 0. But parts are $J(g_M(x|\xi_2), g_M(x|\xi_1))|^b_a$ and we know D = D* and we know that g_M satisfies $B_i(g) = 0$ so I think we are OK on these parts being 0. So, g is symmetric if we impose $\int dx\, g_M(x|\xi)u_1(x) = 0$ which then removes our degree of freedom on g_M. In our example, u_1 = constant, so condition was $\int dx\, g_M(x|\xi) = 0$.

Making use of g_M to solve Lu=f.

We assume up front that $\int dx\, f(x)u_1(x) = 0$, otherwise there is no solution. Notice that when we multiply by y(x) below, we are at that point assuming there is a solution y(x)! If no solution, y(x) = 0 and our little processing falls apart.

We have these two equations which we "process" things in such a way to "expose" y(ξ).

$$Lg_M = \delta(x-\xi) - u_1(x)u_1(\xi)$$
$$Ly = f \qquad \text{// see want to find inhomo solution y(x)}$$

$$y(x)\, L\, g_M(x|\xi) = y(x)\delta(x-\xi) - y(x)u_1(x)u_1(\xi) \qquad \text{// mult by y(x)}$$
$$g_M(x|\xi)\, Ly = g_M(x|\xi)\, f(x) \qquad \text{// mult by } g_M(x|\xi)$$

$$y(x)\, L\, g_M(x|\xi) - g_M(x|\xi)\, Ly = y(x)\delta(x-\xi) - y(x)u_1(x)u_1(\xi) - g_M(x|\xi)\, f(x) \qquad \text{// subtract}$$

$$\langle y, Lg_m \rangle - \langle Ly, g_M \rangle = y(\xi) - u_1(\xi) \int dx\, y(x) u_1(x) - \int dx\, g_M(x|\xi) f(x) \qquad \text{// integrate}$$

$$0 = y(\xi) - u_1(\xi) K - \int dx\, g_M(x|\xi) f(x)$$

$$y(\xi) = \int dx\, g_M(x|\xi) f(x) + K u_1(\xi)$$

$$y(x) = \int d\xi\, g_M(\xi|x) f(\xi) + K u_1(x)$$

$$y(x) = \int d\xi\, g_M(x|\xi) f(\xi) + K u_1(x) \qquad \text{// if we use the symmetric } g_M$$

So our solution is really just "the usual" Green's function solution but with g_M in place of g, and as usual we add an arbitrary homo solution to the result.

Generalize to nullity = 2

Let's try this

$$Lg_M = \delta(x-\xi) - u_1(x)u_1(\xi) - u_2(x)u_2(\xi) = f \qquad \text{where } u_i \text{ are orthonormal}$$

Here, we assume for the moment we have only nullity 1 which I think is common. Now consider

$$\langle f, u_1 \rangle = \int dx\, [\delta(x-\xi) - u_1(x)u_1(\xi) - u_2(x)u_2(\xi)]\, u_1(x) = u_1(\xi) - u_1(\xi) * 1 - 0 = 0$$

$\langle f, u_2 \rangle = 0$ in the same way

We then have to go off and solve this equation for g_M

$$Lg_m(x|\xi) = - u_1(x)u_1(\xi) - u_2(x)u_2(\xi) \qquad x \neq \xi$$

which is a side problem we have to deal with. Assuming we find g_m (with BC's of course), then let's try to generalize our algebra above:

$$Lg_M = \delta(x-\xi) - u_1(x)u_1(\xi) - u_2(x)u_2(\xi) \qquad Ly = f \qquad \text{// see want to find inhomo solution } y(x)$$

$$y(x)\, L\, g_M(x|\xi) = y(x)\delta(x-\xi) - y(x)u_1(x)u_1(\xi) - y(x)u_2(x)u_2(\xi) \qquad \text{// mult by } y(x)$$

$$g_M(x|\xi)\, Ly = g_M(x|\xi)\, f(x) \qquad \text{// mult by } g_M(x|\xi)$$

$$y(x)\, L\, g_M(x|\xi) - g_M(x|\xi)\, Ly = y(x)\delta(x-\xi) - y(x)u_1(x)u_1(\xi) - y(x)u_2(x)u_2(\xi) - g_M(x|\xi) f(x)$$

// subtract

$\langle y, Lg_m \rangle - \langle Ly, g_M \rangle = y(\xi) - u_1(\xi) \int dx\, y(x)u_1(x) - \int dx\, g_M(x|\xi)\, f(x)$ // integrate

$- u_2(\xi) \int dx\, y(x)u_2(x)$

$0 = y(\xi) - u_1(\xi)\, K_1 - u_2(\xi)\, K_2 - \int dx\, g_M(x|\xi)\, f(x)$

$y(\xi) = \int dx\, g_M(x|\xi)\, f(x) + K_1\, u_1(\xi) + K_2\, u_2(\xi)$

$y(x) = \int d\xi\, g_M(\xi|x)\, f(\xi) + K_1\, u_1(x) + K_2\, u_2(x)$

$y(x) = \int d\xi\, g_M(x|\xi)\, f(\xi) + K_1\, u_1(x) + K_2\, u_2(x)$ // if we use the symmetric g_M

There are now two conditions we impose to get a symmetric g_M which set the two free constants:

$\int dx\, g_m(x|\xi)\, u_1(x) = 0 \qquad \int dx\, g_m(x|\xi)\, u_2(x) = 0$

Generalize to nullity = m for order n ODE (where m ≤ n)

$Lg_M = \delta(x-\xi) - \Sigma_{i=1}^{m}\, u_i(x)u_i(\xi)$ where u_i are orthonormal

$\langle f, u_i \rangle = 0$ i = 1,2..m consistency conditions

$\int dx\, g_m(x|\xi)\, u_i(x) = 0$ i = 1,2..m to make g_M be symmetric

$y(x) = \int d\xi\, g_M(x|\xi)\, f(\xi) + \Sigma_{i=1}^{m}\, K_i\, u_i(x)$ // if we use the symmetric g_M

Application to the Example

$L = -D^2$ with unmixed, so self-adjoint // p 86 (1.89)

$u_1(x) = \sqrt{1/L}$ so that u_1 is normalized on (0,L) // p 87 A

$Lg_M = \delta(x-\xi) - 1/L$ // p 87 (1.91)

$\int dx\, f(x) = 0$ consistency // p 87 B

$\int dx\ g_M(x|\xi) = 0$ condition to make g_M be symmetric // p 88 A

$y(x) = \int d\xi\ g_M(x|\xi)\ f(\xi) + K$ // if we use the symmetric g_M // p 89 A

In this example, we can find g_M by solving $-D^2 g_M = -1/L$ for $x \neq \xi$ which is an easy problem to solve. But in general, this is not an easy problem! He solves for g_M in the "usual manner" and determines constants in the usual manner and we are left with one free constant, as we expect. Then that one free constant is nailed when we impose the symmetric-g condition, see p 88 B for the symmetric g_M.

Exercises

I have done 1.54 and 1.55 in the notes above. I think the thin ring problem is close to what we have already done. Others are to find the modified Green's in various cases.

References

Classified into three areas: ODE, Green's and Distributions.

THE END of this very long 91 page chapter that too me forever to get through.