Stakgold Chapter 3 Notes PhL started 2.1.05 resumed 3.13.09 fin 3.21.09

In 2005 I browsed the first part of this chapter of Stakgold but did not get very far, but did make some notes. In 2009 I started over and updated the existing notes and added to them and finished the chapter.

History: 1
Chapter 3: Linear Integral Equations 2
3.1 Introduction 2
Separable kernels. 3
Hilbert-Schmidt (H-S) kernels. 3
The Fredholm Integral Equations. 4
Examples 4
3.2 The Neumann Series Method 8
The general Volterra case. 9
3.3 The Spectrum of a self-adjoint H-S operator. (p 212) 10
The adjoint of a bounded linear operator (212) 10
Properties of the eigenvalues and eigenfunctions of Symmetric Kernels (212) 10
Existence of a nonzero eigenvalue for a symmetric HS operator: Six Theorems (214) 11
3.4 The solution of the inhomo2 integral equation when K is symmetric and H-S. (p 220) 15
3.5 Extremal Principles (223) 18
Courant minimax principle. 20
3.6 Approximations Based on Extremal Principles (226) 20
Introduction. 20
The Rayleigh-Ritz Procedure (228) 21
The Schwarz Iteration Procedure (p 231) 22
Upper Bounds to Eigenvalues 23
The WBFS Theorem. 23
3.7 Continuity and Uniform Convergence: bilinear series and iterated kernels 24
Bilinear Expansions 26
3.8 Approximation Methods for Solving Integral Equations 27
1. Lattice Integration (242). 27
2. Approximating the kernel as a finite separable sum. 28
3. Approximating u(x) in terms of a finite set of independent functions $v_i(x)$. 28
(1) Best in the Mean. 28
(2) Galerkin's Method (Russian ~1915). 28
4. A Variational Principle (245) 29
5. The Rayleigh-Ritz Equations (248) 30
3.9 Non-symmetric HS operators (250) 30

History:

Fredholm is best remembered for his work on integral equations and spectral theory. In fact much of this work was accomplished during the months of 1899 which Fredholm spent in Paris studying the Dirichlet problem with **Poincaré**, Emile **Picard**, and **Hadamard**. In 1900 a preliminary report on his theory of Fredholm integral equations was published as *Sur une nouvelle méthode pour la résolution du problème de Dirichlet*. **Volterra** had earlier studied some aspects of integral equations but before Fredholm little

had been done. Of course **Riemann**, **Schwarz**, Carl **Neumann**, and Poincaré had all solved problems which now came under Fredholm's general case of an integral equation; this was an indication of how powerful his theory was. Fredholm's contributions quickly became well known to the world of mathematics when **Holmgren** lectured on Fredholm's theory at Göttingen in 1901. **Hilbert** immediately saw the the importance of Fredholm's theory, and during the **first quarter of the 20th century the theory of integral equations became a major research topic**. Fredholm published a fuller version of his theory of integral equations in *Sur une classe d'équations fonctionelle* which appeared in Acta Mathematica in **1903**. Hilbert extended Fredholm's work to include a complete **eigenvalue theory** for the Fredholm integral equation. This work **led directly to the theory of Hilbert spaces** (which was essentially a generalization of linear algebra to ∞ dimensional function spaces.)

Chapter 3: Linear Integral Equations

3.1 Introduction

Integral equations are easier than differentials (1) because usually the operator is bounded and even completely continuous. Also, (2) claim that integral equation includes the BC's that you have to supply separately for a DE, but it is not obvious to me why this is. [think scattering $\psi = \varphi +$ integral, eg] And (3), the integral form lends itself to numerical and variational approaches.

In the previous chapter we worked with an integral operator $y(t) = Ax = \int x(t')\,dt'$, but that is not what we are talking about in this chapter. Instead we want $y(t) = \int k(t,t')x(t')dt'$ and of course we can duplicate the simpler case from this case with a θ function for the kernel k. This integral is on some (a,b) fixed range. This **k** function is called the **kernel**. The type of integral equation with a kernel like this is called a **Fredholm integral equation**. This phrase appears later on page 195. We say that y = Kx so the kernel k generates the linear operator **K**.

Erik **Ivar Fredholm** (April 7, 1866 – August 17, 1927) was a Swedish mathematician who established the modern theory of integral equations. His 1903 paper in Acta Mathematica is considered to be one of the major landmarks in the establishment of operator theory. The lunar crater Fredholm is named after him.

Comment: I think the entire world of linear integral equations can be divided into Fredholm equations or Volterra equations that that is it. The linearity condition limits you, and you have the g = Kf matrix analogy. These two types of integral equations differ only in the nature of the upper integral endpoint.

Long Comment on Matrix Analogy. Notice that this form of operator is exactly the notion of an infinite dimensional square matrix. We could write for the eigenvalue problem

$$\mu\, u(x) = \int dy\, k(x,y)u(y) \quad \Rightarrow \quad \mu\, u_{\mathbf{x}} = S_{\mathbf{x}}\, k_{\mathbf{x},\mathbf{y}}\, u_{\mathbf{y}} \quad \Rightarrow \quad \mu\, u = K\, u$$

Oddly, Stakgold does not mention this fact so familiar from quantum mechanics. In the matrix sense, it is easy to see that we have an "infinite dimensional" matrix. The matrix sense of infinity here is different from the sense of infinity of u(x) being expandable in an infinite set of basis functions. Maybe we could

think somehow of the above matrix sense having delta function basis functions at each point x. Consider for example:

$$\langle u|u\rangle = \Sigma_{m=0}^{\infty}\langle u|\varphi_m\rangle\langle\varphi_m|u\rangle = \Sigma_{m=0}^{\infty} a_m^* a_m$$

$$= S_x \langle u|x\rangle\langle x|u\rangle = S_x \; u_x^\dagger u_x = \int dx\, u^*(x)\, u(x)$$

In the first line, we project function u(x) onto the basis functions $\varphi_m(x)$, whereas in the second line we project onto the "basis functions" which are the values of the function u(x) at each x. It is just the m representation versus the x representation.

$$a_m = \int dx\, \varphi_m^*(x)\, u(x) = \int dx\, \varphi_m^*(x)\, u(x)$$

$$u_{x'} = \int dx\, \delta(x'-x)\, u(x) = \int dx\, \psi_{x'}(x)\, u(x)$$

This last line makes explicit the idea of a set of "basis functions" $\psi_{x'}(x) = \delta(x'-x)$. Although such a set of basis functions is "uncountable" and cannot strictly therefore form a basis, we know it works.

Comment on boundedness. We are only going to care about the case where our integral equation linear operator K is bounded. If it is not bounded, there is not much theory to go on, and not much you can say in a theory sense. Chapter 2 had lots to say about bounded operators, we recall. He refers at first to this bound as M. The lowest bound is $\|K\| = \min_u \|Ku\|$ normalized.

Separable kernels. Such a kernel is $k(x,y) = \Sigma_{i=1}^{N} p_i(x)\, q_i(y)$ where p and q are in L_2. It is a finite sum where each term in the sum factorizes.

Claim: For a separable kernel, the operator K in y = Ku is **completely continuous** (bounded sets to into compact sets). But the reader does not know yet why this is a useful fact, but the reader recalls that there were some theorems about completely continuous bounded operators in Chapter 2.

Proof: [3.12.09] On page 192-193 he shows that $\|Sf\| \le M \|f\|$, so S must be bounded. $M = \Sigma_i \|p_i\| \|q_i\|$, it happens, and these p and q norms are finite by the definition of a separable kernel. A theorem of Chapter 2 says that a bounded operator transforms bounded sets into bounded sets. If the domain is bounded, so is the range. BUT, our range is a finite dimensional space since it is spanned by the n functions $p_i(x)$ as shown page 192 A. Any bounded finite-dimensional space is compact -- this is the content of either the BW or the HB theorems, see compactness notes. Thus, a separable-kernel mapping S maps bounded sets into compact sets, and this means the mapping S is a **completely continuous operator**, more or less also known nowadays as a **compact operator**.

Hilbert-Schmidt (H-S) kernels. This just means the kernel's double integral is finite on our interval (a,b). We can show that a H-S operator is bounded (= continuous) with a bound M, so that $\| K \| \le M \equiv \sqrt{\text{double integral of K}}$. We are *not* saying that the norm is this integral, it is ≤ this integral. This fact is shown on page 193.

Comment: One gets the impression that an integral equation operator K with a *non*-HS kernel cannot be a bounded operator. That is to say, one gets the impression that *all* linear integral equation bounded operators must be HS operators. So you wonder: could a kernel be bounded but not HS? I think the answer is probably yes, but only a HS bounded operator is compact = completely continuous.

Comment: **Schmidt** was a grad student of **Hilbert** circa 1905-1910. Hilbert was at Konigsberg (this city moved to Russia after WWII) and in 1895 went to U of **Gottingen** to be chairman, the best math department in the world it was said. In slightly different ways, Hilbert and Schmidt figured out most of the integral equation theory.

Basis on a square comment. You can define basis on a square in R^2 using a basis in R^1 in an obvious way. Author shows that such a basis is complete and orthonormal, and retains dual indices which is fine. These are infinite bases, by the way. We can specialize this to choose ϕ_i as one basis, and $\overline{\phi_k}$ as the other basis. In a sense, the function k(x,y) can be expanded in a full infinite set of basis functions in each variable.

So the obvious thing we want to do is expand k(x,y) in this double basis as shown in mid page 194. This is not formally separable because the sum is infinite (but we will approximate it later with a finite sum). The last line on this page shows what error you get if you stop your expansion at some n. Note that a_{jk} are just the double Fourier coefficients. We know that the RHS residual $\to 0$ because of the usual Bessel inequality business (well, see pages 125-126, RL Lemma, etc). So we can get as close as we like to the exact operator by taking more terms in the sum. By a chapter 2 theorem, we have a sequence of operators which are all completely continuous, so the limit K is also. So that is our proof that K is CC. The key fact here is that the double integral was finite, which is what made K be H-S. This is the main point of what we have done so far in this chapter!

Fact: Any H-S K is **completely continuous** as well as bounded. The kernel itself my not be bounded as a function of one or both of its variables, but no bother.

The Fredholm Integral Equations. Think of K = A of our earlier discussion, and B = K - μI as B = A - λI and the domain vector is u(x) [formerly we called it x, as in y = Ax]. We have for B

Bu = 0	Ku = μu	**homo**geneous eigenvalue problem for K	(3.5)
Bu = f	Ku = μu + f	**inhomo**geneous of **second** kind, μ = a free parameter	(3.6)
	Ku = f	**inhomo**geneous of **first** kind	(3.4)

Suppose homo has some solutions u_μ and inhomo2 has a particular solution u_1. Then $u_1 + u_\mu$ are also solutions of inhomo2.

If upper integral endpoint is variable x (a **Volterra equation**), you can easily redefine the kernel (as noted earlier) to get things back into the Fredholm form.

Examples: How we shall look at some real-world examples:

Example 1 : k(x,y) = xy and interval is (0,1)

(a). Looking at the eigenvalue problem where k(x,y) = xy and interval is (0,1). The solution is summarized in the indented area. If $\mu = 1/3$, you have u(x) = Cx where C is any constant. If $\mu=0$, then solution is any u(x) such that $\int x\, u(x) = 0$. For other values of μ, there is no solution. We have a point spectrum with two points. The $\mu = 1/3$ is a **simple eigenvalue** (you want to say multiplicity = 1, ignoring the freedom implied by C), while $\mu = 0$ has a infinite number of solutions, so **infinite multiplicity**.

We know in matrix world that the eigenvectors of a matrix A form a basis for E_N. The claim is made that the eigenvectors of K forms a complete set as well. This set consists of the function x (which goes with the eigenvalue 1/3), and all functions orthogonal to x (an infinite number, going with eigenvalue 0). The number is infinite because there are an infinite number of functions orthogonal to x, you could just think of them as all the other Legendre polynomials.

Just in passing, what would happen if we adopt our infinite matrix view and blindly try to take the determinant to find the eigenvalues:

$$\det (K-\mu\, 1) = | k(x,y) - \mu\, \delta_{x,y} | = 0.$$

I don't right now know how to do such a "determinant", but I suspect this subject shall arise soon.

1 (b) Same kernel, but now do the second kind equation (inhomo2) with $\mu\neq 0$ in there. For $\mu \neq 1/3$, we get the explicit solution shown. In general theory, homo Bu = 0 has no trivial solutions for these μ, so Bu=f must have one solution and that is the one shown. B is therefore invertible for such μ.

Now if $\mu = 1/3$, we can have solutions for f(x) only that integrate with x to 0, so now range is restricted to such functions f(x). But for $\mu = 1/3$, homo has solutions Cx as we found in (a). Now the general solution of inhomo2 is $u(x) = -f(x)/\mu + Cx$ where we in effect are adding the homo solutions.

2 (c) Same kernel, but now do the first kind equation with $\mu = 0$ (inhomo1). In this case, the range is even more restricted: solutions exist only if $f(x) = \alpha x$, in which case $u(x) = 2\alpha$ + any function which integrates to 0 against x. [We get this feeling that as the range becomes more restricted in these examples, the solution gets more degrees of freedom, a dim memory of the alternative theorem somehow.]

Example 2. $k(x,y) = \sin(x)\sin(y) + \theta \cos(x)\cos(y)$ on $(0,\pi)$. eigenvalue equation only

Here, θ is some arbitrary constant (not an angle as letter suggests). Stak looks just at the homo equation which is the eigenvalue problem and things are summarized in the table page 198. In general, there are three eigenvalues. The one $\mu=0$ is any function which integrates to 0 against both sine and cosine, and I think this is an infinite set of functions, perhaps u = sin(nx) or some such, so probably this eigenvalue has infinite multiplicity. The second eigenvalue is $\mu = \pi/2$ and there is one eigenfunction u(x) = sin(x) that goes with it. The third eigenvalue is $\mu = \theta(\pi/2)$ with one eigenfunction u(x) = cos(x). If it happens that $\theta=1$, these two eigenvalues become one degenerate eigenvalue.

Certainly this seems a bizarre example.

Example 3: $k(x,y) = \sin(x)\cos(y)$ on $(0,\pi)$. eigenvalue equation only

Again, results presented in a table. The only eigenvalue is $\mu = 0$ and it does have an infinite multiplicity and solutions are u(x) which are orthogonal to cos(x). But unlike example 1(a), our solution set does not

form a complete basis because cos(x) is excluded from being a solution! Stak claims this can only happen with a non-symmetric kernel.

Suppose kernel were k(x,y) = p(x) q(y) in this example. For $\mu \neq 0$, solution would have to be $u(x) = C\ p(x)$ and we would have $\int dy\ q(y)\ p(y) = \mu$. In example 3, this integral was zero, which contradicted our $\mu \neq 0$ assumption. If the kernel is "symmetric" so p = q, then this integral is $\int dy\ p^2(y)$ and cannot then be 0, so we don't have this problem, and there is a viable eigenvalue $\mu \neq 0$ (the value of the integral of pq). This extra eigenvalue then provides the missing basis function and then the set of eigenfunctions form a complete set.

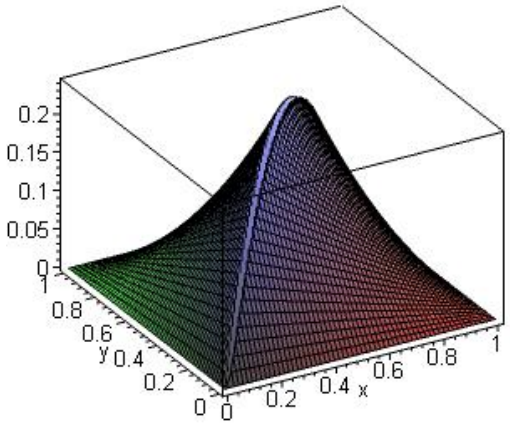
Example 4: k(x,y) = separable $= \Sigma_{i=1}^{n} p_i(x)\ q_i^*(y)$. p_i and q_i each lin indep sets

For $\mu \neq 0$, the eigenvalue problem converts to an nxn *matrix* eigenvalue problem shown on the left side of 3.8. At once we assume that k is **symmetric** so $q_i = p_i$. There will be n eigenvalues μ_i (not necessarily all different, but none 0), and each will have an independent eigenvector c_i which corresponds to a solution of the original eigenvalue problem Ku = μ u . So this gives us n eigensolutions of our problem. For $\mu = 0$, the solutions u(x) are all functions which are orthogonal to ALL the n functions p_i^* . If I think of the first n Legendres, say, I then have u(x) being all Legendres with ℓ>n. So in this problem, we first have the n basis functions associated with $\mu_i \neq 0$, and these are linear combinations of the $p_i(x)$. Then we have sort of ∞-n eigenfunctions associated with $\mu = 0$, and combined, we have a full basis. Again, this seems always to happen when the kernel is symmetric.

Example 5: k(x,y) = Green's function of the string problem 3.9 (symmetric)

This will be a long example. The string problem is the ODE 3.9 with T = 1, as treated in Chapter 1. We solve this using the Green's method and the Green's is k(x,y) shown in p 199A. This k(x,y) is symmetric, as this plot from Maple verifies:

```
f := piecewise(x < y, x*(1-y), x > y, y*(1-x)):
plot3d(f,x=0..1, y=0..1,grid = [60,60]);
```



Comments. Here we see an ODE problem directly related to an integral equation problem. The relation is provided by the Green's function method. The integral equation does in fact "build in" the two boundary conditions of the ODE problem as claimed in the opening of this chapter. For example, u(0) = $\int k(0,y)u(y)dy = \int 0(1-y)u(y)dy = 0$. In this simple example, we know the eigenvalues λ and eigenfunctions u(x) of the ODE problem. Therefore, we also know the solutions of the HS kernel problem since μ = 1/λ. Down the road we are not going to know the solutions of the "related ODE", and it will have an unbounded operator Lu like $-D^2$ in this example, and it is a "hard problem". But the corresponding HS integral equation has a "nice" bounded operator K, and we have more tools at our disposal to solve a bounded operator problem, that at least is the claim. Also, the operator K is compact which I guess will also help, not just bounded. So there is some theory coming soon I suspect, we are just warming up with these examples.

So we have solved the homo equation (eigenvalue) because we know the ODE solution in this example. What about the inhomo2 equation where we have both μ and f(x) ? We know that the homo solutions are $C\sin(n\pi x)$ so we try converting our inhomo2 problem to one involving the Fourier coefficients a_n for f(x) and b_n for u(x), where we mean Fourier relative to the sine solutions of the homo. We then find that the a_n and b_n are related simply as in 3.13 so in Fourier space the problem is greatly simplified.

We are now going to study three cases.

The first is when μ is NOT one of the homo eigenvalues. In this case we can write the b_n as in 3.13a and we have thus found our solutions u(x) as shown there, so we are "done". Stak sidetracks a moment on the convergence of this 3.13a series. For large n, it is not obvious that the thing is converging very well. But he then rewrites the series as in 3.14 where f(x) now appears on its own, and the second term there is seen to have good convergence for large n.

The second case is when μ = one of those homo eigenvalues, call it $\mu = 1/(\pi m)^2$. Now there is only a possible solution if f(x) is orthogonal to $\sin(m\pi x)$. In this case, all the b_n are as in 3.13a except for b_m which is a free parameter, so our solution for u(x) is as shown in p201 A.

The third case is when μ = 0, so we really have the inhomo1 problem now. From 3.13a we seem to have $b_n = a_n(n\pi)^2$ and a corresponding solution u(x) shown in 3.16. But this does not look particularly convergent, having an n^2 in the numerator. If you pick some arbitrary f(x) [like f(x) = 1] , the resulting solution u(x) as given in this series might not converge. In other words, there are restrictions on the range vector f(x) and Stak interprets this restriction in terms of the string problem (mid page 202). [I regret to say that I have not recently read Chapter 1 and I seem to have no notes on it. I would like to defer that if possible. He says that the series might still converge in the sense of a distribution and delta function.]

End of Example 5.

Example 6: k(x,y) = Green's for $L = -D^2 + 1$ on $(-\infty,\infty)$

Again, we write the integral equation version (3.17), and the ODE version (p202B) of the problem with μ = 1/λ. The Green's in this case is as shown p 202A, which is $k(x,y) = e^{-|x-y|}/2$, which is **symmetric**. On the square this thing diverges so **k(x,y) is not H-S**, **but** he is going to claim that still **K is bounded** so we have still a nice problem. [now at page 203, have done 12 pages of Chap 3.] He shows on page 203 that $||Kf|| \le ||f||$ which says $||K|| \le 1$ and is therefore bounded. It is left to us to show that in fact $||K|| = 1$.

So our integral homo equation is Ku = μu as in 3.17. If we look at the ODE, we see that the only possible solutions are expos, but on an infinite square they are not L_2 functions. The conclusion is that

there are **no eigenvalues μ at all**, not even a single eigenvalue. The point spectrum is empty for this problem.

He looks next at inhomo2. He replaces u(x) with z(x) as in p 203A. We end up with an ODE for z(x) as shown in 3.20. The Greens equation for 3.20 is p 204A. The Green's is shown to be p 204B, where now we have decaying expos in both directions with A and B as yet unknown. The usual Green's "matching" tells us A and B so we get 3.21 as the Green's for ODE 3.20 and so the solution to 3.20 in terms of this Green's is p204C. We then have our solution for u(x) as in D. The conclusion is that as long as μ is not on the real interval (0,1), this is a viable solution for I guess any f(x). For all these μ, things are "**regular**", so μ is in the resolvent set. The range (0,1) for μ belongs to the **continuous spectrum**. Recall Chap 2:

Theorem 3: The *continuous* spectrum of a *symmetric* operator A is limited to the real axis.

so in this example, we do have a piece of the real axis being part of the continuous spectrum. Recall also that

Theorem 4: The entire spectrum of a *self-adjoint* operator lies on the real axis, and there is no residual spectrum, meaning that case 4 never occurs.

Probably example 6 is self-adjoint that that is why there is no residual spectrum (I am guessing).

Exercises (p 204). I notice that I have circled 3 of these problems. I probably did these problems in Winter 1972 (37 years ago!) in my UCB Ted Dushane course math 224A. My history XLS says I have "partial notes" somewhere for this course! As I look at those few notes now, I realize that I must have just kept a few sample pages of notes for each course, and then thrown out 90% of each course's notes. The current "me" would never throw out something like that! I am amazed that I did that.

3.2 The Neumann Series Method

We take our **inhomo2** equation and redefine f to be f/μ and set $\lambda = 1/\mu$ we end up with

$$u = f + \lambda K u$$

In my physics experience, this is a famous integral equation for scattering. The function f(x) is some kind of incoming "plane wave" and u(x) is then the ψ(x) that solves some integral version of the Schrodinger Equation, an example being the Lippmann-Schwinger equation. It may be that $k(x,x') = G_0(x,x')V(x')$ in this application (non-symmetric?), I am not sure, but let's not get distracted on that.

Stak does not mention it, but we have as usual a formal solution of the above equation:

$$u = (1 - \lambda K)^{-1} f = (1 + \lambda K + \lambda^2 K^2 + ...)\, f \qquad \text{which is 3.25}$$

The meaning of the operator K^n is just what you would think, for example,

$$k^2(x,y) = \int k(x,z)k(z,y)\, dz$$

This is all displayed in full detail on page 206 (but not my formal solution idea), and then on page 207 we get several theorems:

(1) If the series for u converges, it is indeed a solution of our integral equation $u = f + \lambda Ku$.

(2) The condition for convergence is that $|\lambda|\ ||K|| < 1$, not too surprising, so we have a circle of convergence in the λ plane, see p 207 A.

(3) The series solution is unique.

(4) If $f = 0$, the series solution is $u(x) = 0$. So for $|\lambda|\ ||K|| < 1$, since the series solution is the only solution of our equation, the equation $u = \lambda Ku$ can have no eigenvalues. Thus, if there are eigenvalues, they must have the property that $|\lambda|\ ||K|| \geq 1$ which says $||K|| \geq 1/|\lambda| = |\mu|$ so only eigenvalues must be $|\mu| \leq ||K||$.

The downside of the Neumann series solution: (1) hard to compute all those K^n kernels; (2) there may be a solution which is not a converging Neumann series.

We are now going to get two big examples.

Example 1: Here we use our favorite $k(x,y) = xy$. // p 208

The iterated kernels are just $k^n(x,y) = xy(1/3)^n$, and $||K|| = 1/3$ it is claimed, obvious that $\leq 1/3$, So the Neumann series converges for $|\lambda|\ ||K|| < 1$ or $|\lambda| < 3$ and is given by (3.26). The sum can be done to get (3.27) and then 3.27 allows us to analytically continue the solution beyond our radius 3 disc and gives our problem's solution for any λ except $\lambda=3$ where it blows up.

The general Volterra case.

Example 2: The general **Volterra** case. // p 209

(a) We first look at the homo equation (3.28) = (3.29). We assume that $|k| \leq M$ on our rectangle where we sort of make up a b. Note that this is a much stronger condition that saying $||K|| \leq M'$ as we do for H-S. With this assumption, Stak shows that the homo equation has no eigenvalues! The proof consists of showing that $|u(x)| < \varepsilon$ for any x, and ε can be arbitrarily small, we must have $u(x) = 0$ and so there is no non-trivial solution. This proof occupies page 209. He then claims the conclusion also follows with the less strict assumption that K is H-S.

(b) This treats the special case $\mu = 0$. We conclude that equation (3.28) has no eigensolutions for $\mu = 0$ unless it happens that $k(x,x) = 0$ for some x. Stak claims this special case is related to "singular points" for the corresponding ODE, and we are told to consult exercise 3.10.

(c) Now we look at the inhomo2 equation so we have μ or λ and f(x). The Volterra sense can be hidden by saying $k(x,y) = 0$ as needed so we now have the regular form our integral equation. Stak examines the Neumann series in this case and finds that the terms for large n get small in ratio uniformly for any x in our interval. In fact we have "convergence in the mean" (L_2 convergence) for ANY λ. Stak does not mention what $||K||$ might be for this effective kernel, but we are still assuming $|k| < M$. I guess we do know that $||K|| \leq abM$ or maybe even half this. [I am not quite sure where the Volterra aspect enters into this case (c). Well now I see: in equations 210 A, the K operator has that upper x endpoint which is

where the powers $(x-a)^n$ come from and where the important 1/n! is coming from. So his comment about faking k was just to show we have the standard inhomo2 "form".]

So the conclusion here is that for *any* λ we get a convergent Neumann series solution (3.33). So the Volterra situation is "nice" in this regard. The conclusion can be generalized to k being any H-S kernel. If equation is non-Volterra, we instead get a radius of convergence for |λ|.

Exercise 3.10. Here we consider the most general inhomo 2nd order ODE with function coefficients. Stak proves that the boundary conditions u and u' determined at point x=a force a *unique* solution to the ODE. This seems sort of intuitive if we were to numerically integrate the equation away from x = a. It is assumed that all the functions are "nice".

How does this proof go? He recasts the ODE as an integral equation for u"(x) and this is seen to be a Volterra integral equation of inhomo2 type. The kernel k(x,y) is stated in p 211 A. We just showed in part (c) that the solution of such a Volterra equation is unique and is given by the Neumann series which converges for all λ. I guess the assumption |k| < M is met if we assume some finite square, functions are not allowed to have poles if they are continuous ie nice. And of course $a_o(x) \neq 0$ in the interval of interest. Thus, the solution u" must be unique, and then use equation B to compute u'(x) using u'(a), and then use equation to compute u(x) using u(a) and we thus end up with our unique solution u(x), as claimed. This is an example of how the theory of integral equations tells you something interesting about an ODE.

I think we are now headed back to the theory department.

3.3 The Spectrum of a self-adjoint H-S operator. (p 212)

{ at this point, I have done 21 pages of Chap 3, done on 3/13 and 3/14/09. enough for today }

The adjoint of a bounded linear operator (212)

3.15.09. The first section on page 212 talks about a "symmetric" kernel. The term symmetric means Hermitian in my matrix sense: $k(x,y) = \overline{k}(y,x)$ where overbar means C.C. Luckily, if the kernel k is symmetric as I have just written it, then the operator K is self-adjoint, we don't have to worry about the fancy distinctions of Chapter 2. As expected, $\langle Ku,v\rangle = \langle u,Kv\rangle$ for self-adjoint and we have the usual result that $\langle Ku,u\rangle$ is real (remember backwards from usual bra/ket notation).

Properties of the eigenvalues and eigenfunctions of Symmetric Kernels (212)

The next section **pp212-214** provides very simple proofs for some **theorems**, some of which were given in Chapter 2, but not all.

If K is symmetric: (the first two are the same as those of the finite matrix world)

(1) Eigenvalues of Ku = μu are all real. (p 213)

(2) Eigenfunctions of different eigenvalues are orthogonal.

If K is symmetric and is *also* H-S:

(3) the multiplicity of any non-zero eigenvalue is finite. (**simple** if m=1, **degenerate** if m> 1) (p 214)

(4) the # of eigenvalues if either (1) finite, or (2) infinite, in which case the only limit point is $\mu_i \rightarrow 0$.

The proofs of these theorems are all very straightforward.

Existence of a nonzero eigenvalue for a symmetric HS operator: Six Theorems (214)

Now we think about eigenvalue bounds. Using the usual simple steps, it is easy to show that $|\mu| \leq ||K||$. If we have a H-S kernel, then there exists a finite ||K||. But now we call upon a Chapter 2 theorem which I quote:

Theorem 6: If A is symmetric *and* completely continuous, then: (p 190)
(1) At least one of these numbers is an eigenvalue: ||A|| or –||A||.
(2) There is no eigenvalue μ for which $|\mu| > ||A||$.

What exactly is this saying? Suppose μ_1 is the eigenvalue having the largest absolute value of all the eigenvalues. We don't know the sign of μ_1. But we know that $|\mu_1| = ||A||$.

Recall that earlier in Chapter 3, in the intro section, we jumped through a lot of hoops just to show that a H-S kernel integral operator K was "completely continuous". We showed this first for a separable kernel, then we took the limit and showed it was true for any H-S k(x,y) which, we know, has a finite bound ||K||. So this then allows us to use Theorem 6 above and to make this claim:

Theorem 1 and its Corollary (p 215)

There is at least one eigenvalue for K, and we know this about it:

$|\mu_1| = ||K||$

The only thing we don't know about it is whether μ_1 is positive or negative (it is real). Notice that this is consistent with our weaker result just stated, $|\mu| \leq ||K||$.

Now we are going to develop a different fact, bottom of page 214. Let's just write out some of these facts:

$|| Ku ||^2 = \langle Ku,Ku \rangle$ $|| Ku || \leq ||K|| \, ||u||$ // definition of norm ||K||

But for the eigenvalue problem

$|| Ku || = || \mu u || = |\mu| \, ||u||$

Thus we have

$|| Ku || = |\mu| \, ||u|| \leq ||K|| \, ||u||$ => $|\mu| \leq ||K||$ // as already quoted above.

But now we have $||K|| = |\mu_1|$ so we can say

$$\| Ku \| \le \|K\| \|u\| \qquad \Rightarrow \qquad \| Ku \| \le |\mu_1| \ \|u\|$$

Meanwhile, we can say this:

$$\langle Ku,u \rangle = \mu \langle u,u \rangle = \mu \|u\|^2$$

$$|\langle Ku,u \rangle| = |\mu| \langle u,u \rangle = |\mu| \|u\|^2$$

$$|\langle Ku,u \rangle| / \ \|u\|^2 = |\mu|$$

This last is true for each eigenvalue and its corresponding eigenfunction. Certainly if we go with μ_1 we will have it be true, so the LHS can get as large as $|\mu_1|$. Now I guess the claim is that this is the most that the LHS can be. How would I prove this? Let $u = \Sigma a_i \varphi_i$ that is normalized. Then

$$\langle Ku,u \rangle = \langle K \Sigma a_i \varphi_i, u \rangle = \langle \Sigma a_i \mu_i \varphi_i, u \rangle = \langle \Sigma_i a_i \mu_i \varphi_i, \Sigma_j a_j \varphi_j \rangle$$

$$= \Sigma_i \Sigma_j \mu_i a_i a_j^* \langle \varphi_i, \varphi_j \rangle = \Sigma_i \mu_i |a_i|^2 < \Sigma_i |\mu_1| |a_i|^2 = |\mu_1| \Sigma_i |a_i|^2 = |\mu_1|$$

This I have shown that $\langle Ku,u \rangle \le |\mu_1|$ for general u, and we know $\langle Ku,u \rangle = |\mu_1|$ for u = the eigenfunction. So the upshot of all this is that I have shown that:

$$|\langle Ku,u \rangle| / \ \|u\|^2 \le |\mu_1| \qquad \text{and } |\langle Ku,u \rangle| / \ \|u\|^2 = |\mu_1| \quad \text{when } u = \varphi_1$$

What this really means is this: $\max |\langle Ku,u \rangle| / \ \|u\|^2 = |\mu_1|$ or

$$||| K ||| = |\mu_1|.$$

So the upshot of this Theorem 1 can be simply stated:

$$||| K ||| = |\mu_1| = \|K\|$$

Sorry it took so long! So the maximization problem Stak always talks about is the |||K||| thing.

We are now after Theorem 1 on page 215. We let φ_1 be the eigenfunction that goes with μ_1. We then define a kernel and operator $k^{(2)}$ and $K^{(2)}$ as shown where we merely subtract $\mu_1 \varphi_1(x) \bar{\varphi}_1(y)$ from the kernel. I agree with the facts quoted top of page 216: (the subtraction piece is symmetric!)

(a) if u is orthogonal to φ_1, then $K^{(2)} u = Ku$.

(b) If u is a multiple of φ_1 , then $K^{(2)} u = 0$

I would say just based on this that $K^{(2)} = KQu$ where Q projects out the perp space of φ_1.

(c) the range of $K^{(2)}$ is the perp space of φ_1. This is what I just said above.

(d) Suppose μ is a non-zero eigenvalue of $K^{(2)}$. Then:

$$K^{(2)}u = \mu u$$

which says that u must be in the perp space of φ_1 where we know $K^{(2)}u = Ku$, so that

$$Ku = \mu u$$

We have shown that if μ is a non-zero eigenvalue of $K^{(2)}$, then it is also an eigenvalue of K. In effect, the only difference between $K^{(2)}$ and K is that the μ_1 eigenvalue has been shifted to 0 for $K^{(2)}$.

Theorem 2

What happens next seems to be a recursion of what we did for K, but now we do it for $K^{(2)}$. We first notice that $K^{(2)}$ is in fact symmetric if K was, and is also H-S (I think this follows since φ_1 is normalized in L_2) . Thus we can just apply Theorem 1 to $K^{(2)}$ and we conclude that ($K^{(2)}$ is also completely continuous) :

There exists some μ_2 such that $|\mu_2| = \| K^{(2)} \|$

The corresponding eigenfunction we call φ_2 which maximizes $< K^{(2)}u,u>$ with $||u|| = 1$.

Since this μ_2 is also an eigenvalue of K, and since μ_1 we know is the largest ev of K, it must be true that $|\mu_2| \leq |\mu_1|$. And also $\| K^{(2)}u \| \leq |\mu_2| \, ||u||$ as in theorem 1.

Assuming $\mu_2 \neq 0$, we know that φ_2 is orthogonal to φ_1.

Theorem 2A. (We are still assuming K is symmetric and H-S since Theorem 1 assumed that.)

This is a restatement of Theorem 2 where we don't have to mention $K^{(2)}$. We assume at the start that we have already found μ_1 and φ_1 . Then there must exist μ_2 and φ_2 of K such that φ_2 is orthogonal to φ_1. Thus function φ_2 maximizes $|<Ku,u>|$ if, *while* you maximize, you impose $<u,\varphi_1> = 0$. This max problem will yield the value $|\mu_2|$ and as noted above $|\mu_2| \leq |\mu_1|$. Of course μ_2 can be degenerate.

Theorems 3 and 3A

Now we can see that you can just keep iterating this procedure to bring out μ_3, μ_4 and so on. So now Stak restates theorems 2 and 2A above for an arbitrary level of this recursion process! The case n=0 is the top level and yields φ_1 and μ_1. Then the case n=1 gives the $K^{(2)}$ level we just did. Notice that the 2A version of the theorem involves maximizing $|<Ku,u>|$ subject to u being orthogonal to ALL the lower φ_i which would mean φ_1, φ_2.... φ_n. So at the n=1 level, only need be orthogonal to φ_1. Of course the eigenvalues we obtain by this method are in descending order in magnitude.

There are two ways that our recursive procedure can end:

(1) After a finite number of steps, we find some $\mu_{n+1} = 0$. This means that the $K^{(n)}$ no longer change, so you just "sweep out the nullspace" as he says. The claim is that the φ_m we get here form a complete set. What is not clear is this: will there be an infinite number of φ_i for the nullspace? Well, $\mu = 0$ is the only eigenvalue for which you are allowed to have an infinite multiplicity, so then yes, there could be an infinite number of EF's here. But it might also be a finite number. In this case, we get a "complete" set of functions φ_m which is a finite set. It will turn out below that this set is complete only in a certain sense.

Another thing unclear: when you hit a degenerate eigenvalue $\neq 0$, do you get all of its independent eigenfunctions? I guess you would find that the eigenfunctions u that max our norm thing allow for m independent functions somehow. How does this work? I guess at some point we will see some examples of this fancy process.

(2) We never hit some $\mu_{n+1} = 0$ and we just keep recurring forever. In this case, we know that the infinite set of eigenvalues must approach 0 since that is known to be the only limit point.

At the bottom of page 217, two claims are now made:

(1) Our procedure produces all non-zero eigenvalues and their eigenfunctions!

(2) This set of eigenfunctions does form a complete set, provided we include all $\mu=0$ eigenfunctions.

Page 218. The text here proves Theorem 4 which I will now state: [the theorem in the book is stated after it is proven]

Theorem 4: Consider the set of basis functions φ_i that emerges from our little "procedure" above, but include in this set only those eigenfunctions that go with non-zero eigenvalues. This set might be finite or infinite, and in the latter case, the limit point is $\mu_i \rightarrow 0$. Is this set of basis functions complete? The answer is : yes, it is complete for the set of functions which can be written as Kf where f is in L_2 . In other words, the set of φ_i is complete for the range of K. This explains the mystery of how a finite set can be complete.

Bottom of page 218: a little proof that our procedure does not "miss" any non-zero eigenvalues. If it did, he calls that eigenvalue α and its function ψ and he gets a contradiction. So this is the proof of little item (1) appearing above.

Page 219 top: Now again we are going to prove a theorem, then state it. The first paragraph shows very simply that an arbitrary function f, though perhaps not in the range of K, can be written as the sum of two functions which he calls h and the other is the Fourier series of f. The component h is in the nullspace of K and thus is associated with $\mu=0$ eigenvalues. Ie, h IS an eigenfunction for $\mu = 0$. The other piece is the Fourier series part which does lie in the range of K. Therefore, when you include the $\mu=0$ eigenfunctions, *any* function f can be expanded on the set of all eigenfunctions of K, where we include both the $\mu \neq 0$ ones and the $\mu=0$ ones. The second term in 3.40 includes only the projections onto the non-zero eigenvalue eigenfunctions φ_i.

Theorem 5: The full set of eigenfunctions of K forms a complete set on L_2.

Again, all the time here we are talking about K being symmetric and Hilbert-Schmidt. Here is an obvious corollary:

Theorem 5A. If K has no value-zero eigenvalues, then the set of non-zero eigenvalue eigenfunctions forms a complete set, and vice versa. Ie, if these form a complete set, then there are no zero eigenvalues.

The last text on page 219 leads up to Theorem 6. The idea here is that we consider a "self-adjoint" "differential system" of second order, which is to say, we have some second order differential operator L which is self-adjoint (and I think this self-adjointness takes account of the two boundary conditions). You can always "talk about" the eigenvalue problem for L, which we write $Ly = \lambda y$ with the BC's. Stak then shuffles this problem just slightly to show that you can restate the eigenvalue problem in terms of a slightly different version of L called M such that the new EV problem is guaranteed to have no zero-value eigenvalues. I think this is just a technical point so we can apply our theorems. So then we look at the integral equation which is associated with our EV differential equation and that is written in 3.42a. Here θ is the eigenvalue which is guaranteed not to be 0 as in 3.42. Then when we rewrite as 3.42b, we are guaranteed that $\mu \neq \infty$ since $\theta \neq 0$. Now the claim is that the kernel in 3.42b, which is the Green's function of our problem with M, is symmetric *and* the corresponding operator G is Hilbert Schmidt. This is not obvious to me, but it must be a property of the Green's Function in general for a second order ODE problem. Granted that this is true, we can apply theorem 5 which says the eigenfunctions of this operator G form a complete set, and there are no zero eigenvalues by construction. But the eigenvalues and functions of G are the same as those of the ODE eigenvalue problem. We therefore reach this conclusion:

Theorem 6: The eigenfunctions of any second order self-adjoint ODE system form a complete set on L_2. I have long known this to be true somehow, but here we have an actual proof.

This theorem will no doubt be used in Chapter 4 on ODE stuff, but in Chapter 3 we are still worrying about integral equations.

3.4 The solution of the inhomo2 integral equation when K is symmetric and H-S. (p 220)

This is the equation $Ku = \mu u + f$. It is the eigenvalue equation with some function f tacked on. Let φ_i be just the eigenfunctions for the non-zero eigenvalues, so φ_i might not be a complete set. Still, make these expansions of u and f, both of which may be invalid due to φ_i not complete

$$u = \Sigma_i a_i \varphi_i \qquad f = \Sigma_i b_i \varphi_i \qquad \Rightarrow \qquad g \equiv \mu u + f = \Sigma_i (\mu a_i + b_i)\varphi_i$$

Now suppose there exists a solution u of $Ku = \mu u + f$. We know that g is in the range of K and can be expanded on the φ_i, so our RHS expansion for g shown above must be valid with some c_i. Now the logic eludes me, but let's try: *assume* that the expansion $u = \Sigma_i a_i \varphi_i$ is valid (which we don't know) . Then we can say that $Ku = \Sigma_i a_i K\varphi_i = \Sigma_i a_i \mu_i \varphi_i$. But this must equal g, so we then have

$$\Sigma_i a_i \mu_i \varphi_i = \Sigma_i (\mu a_i + b_i)\varphi_i$$

Even though φ_i might not be a complete set, this equation could be satisfied if this were true:

$a_i\mu_i = (\mu a_i + b_i)$ or $a_i(\mu_i - \mu) = b_i$

I would say that: *if* Ku = μu + f has a solution u, and *if* this solution can be expanded $u = \Sigma_i a_i\varphi_i$, then the above relationship between a_i and b_i must be true.

Well, here may be a way out. Suppose ψ_i are the eigenfunctions with $\mu_i = 0$. Then we know we can say this for sure

$u = \Sigma_i a_i\varphi_i + \Sigma_i a'_i\psi_i$

$Ku = \Sigma_i a_i K\varphi_i + \Sigma_i a'_i K\psi_i = \Sigma_i a_i\mu_i\varphi_i + 0$

since $K\psi_i = 0$ since ψ_i have zero-value eigenvalues. Thus, we don't have to make our second assumption and we can say: if *if* Ku = μu + f has a solution u, then the a_i and b_i must be related as above.

Case 1: Suppose $\mu \neq 0$ and $\mu \neq$ some u_j (μ = the constant in our equation). Then we write this, and a_i never blows up due to the denominator being 0,

$a_i = b_i/(\mu_i - \mu)$

We can then write our solution by considering

$g \equiv \mu u+f = \Sigma_i(\mu a_i + b_i)\varphi_i = \Sigma_i(\mu\, b_i/(\mu_i - \mu) + b_i)\varphi_i = \Sigma_i\, b_i \{ \mu/(\mu_i - \mu)+ 1\} \varphi_i$

$= \Sigma_i\, b_i \{ \mu + (\mu_i - \mu)\}/(\mu_i - \mu)\varphi_i = \Sigma_i\, b_i\, \mu_i /(\mu_i - \mu)\varphi_i$

where we recall that we know the series for g is valid. Now solve this for u:

$\mu u+f = \Sigma_i\, b_i\, \mu_i /(\mu_i - \mu)\varphi_i$

$u+f/\mu = \Sigma_i\, b_i\, \mu_i /\mu(\mu_i - \mu)\varphi_i$

$u = - f/\mu + \Sigma_i\, b_i\, \mu_i /\mu(\mu_i - \mu)\varphi_i$ **// agrees with 3.45**

and now we see why we had to disallow $\mu = 0$ as well as $\mu = \mu_j$.

Page 221 top: Stak wants to prove that the series in 3.45 in fact converges. Looking back at the Riesz-Fisher Theorem (RFT) from Chapter 2, it is true that is we can show that the series on the left converges, then the one on the right must converge. Stakgold leaves much unsaid here:

(1) How do we know that $\Sigma\, |b_k|^2$ converges? We have $f = \Sigma_i b_i\varphi_i$ and when we started we were not sure this was a complete expansion, but maybe we know at this point it is. Assuming it converges, then the

RFT does tells us that $\Sigma\ |b_k|^2$ also converges, so let's go with that for now. [Also, I suppose Bessel's Inequality tells us $\Sigma\ |b_k|^2\ \leq ||\ u\ ||$ so that would be another reason for $\Sigma\ |b_k|^2$ to converge.]

(2) What happens in the tail of the series on the left? Since limit point $\mu_k \rightarrow 0$, the tail of this series looks like this: (we will use $|\mu_i| \leq 1$ in the tail;)

$$1/|\mu|^4\ \ \Sigma_i |\mu_i|^2\ |b_i|^2 \leq 1/|\mu|^4\ \ \Sigma_i\ |b_i|^2\ \ = \text{converges}$$

So this strongly suggests to me that the left series converges, to the right series must also by RFT.

(3) I have no doubt that plugging 3.45 into 3.43 gives an identity.

Theorem A (p221). In the above work we have shown that: if $\mu \neq 0$ and $\mu \neq$ some μ_j ev of K, then the inhomo2 equation $Ku = \mu u + f$ may be be solved for u as in 3.45 where b_k are the Fourier coefficients of f. Normally we might say such a solution was not unique since we could add to it any solution of the homogeneous equation $Ku = \mu u$. But we have said that μ is not an eigenvalue of this equation, so there are no homogeneous solutions to add, so our mechanically computed solution is the only one.

Case 2a: Maintain $\mu \neq 0$, but suppose μ is in fact some *simple* eigenvalue $\mu = \mu_m$ of $Ku = \mu u$. Looking at 3.44 we get for this one value k=m that $a_m\ 0 = b_m$. If $b_m \neq 0$, then no a_m exists and $Ku = \mu u + f$ then has no solution at all. But if it happens that $b_m = 0$ (which means $\langle f,\varphi_m\rangle = 0$), then $a_m\ 0 = b_m = 0$ so a_m can be any number you want. This notion $b_m = \langle f,\varphi_m\rangle = 0$ is called a "consistency condition" for this case. If it is true, then we have a solution with a_m being arbitrary. So our solution in this case is

$$u = -\ f/\mu + \{\ a_m\varphi_m\ +\ \ \Sigma_{i\neq m}\ b_i\ \mu_i\ /\mu(\mu_i - \mu)\varphi_i\ \} \qquad\qquad a_m = \text{arbitrary} \quad \langle f,\varphi_m\rangle = 0$$

Notice that $a_m\varphi_m$ is a solution of the homogeneous equation $Ku_m = \mu_m u_m$. Thus we have our usual result that we have a particular solution plus and solution of the homogeneous equation. In terms of the operator $A = K - \mu I$, we now have a 1D nullspace for A and our inhomo2 equation is $Au = f$. Our notion of the "alternative theorem" suggests that this corresponds to some restriction of the range of B and of K. That restriction is $\langle f,\varphi_m\rangle = 0$. Remember that for a self-adjoint operator, we don't have to worry about the distinction between A and A* and the (∞ dim) alternative theorem says that $\overline{R_A} = (N_{A^*})^{\perp}$ and $(R_A\ )^{\perp} = N_{A^*}$ so as the nullspace grows, the range shrinks, even though the sum of the nullity and rank is not a finite number as it is in the matrix case.

Case 2b: Same as above but now μ_m is a degenerate eigenvalue. Again $a_m\ 0 = b_m$ for each of the degenerate eigenfunctions in the m group. We could write this as $a_m^{(i)}\ 0 = b_m^{(i)}$. If any one of the f projections $b_m^{(i)} = 0$, we are dead meat and there are no solutions because $a_m^{(i)}$ no exist. But if all of these things are 0, then we get the solution as follows

$$u = -\ f/\mu + \{\ \Sigma_i\ \ a_m^{(i)}\varphi_m^{(i)}\ +\ \ \Sigma_{i\neq m}\ b_i\ \mu_i\ /\mu(\mu_i - \mu)\varphi_i\ \} \qquad a_m^{(i)} = \text{arbitrary} \quad \langle f,\varphi_m^{(i)}\rangle = 0$$

where now our consistency condition applies to each of the degenerate eigenfunctions $\varphi_m^{(i)}$. Now our range is more restricted, and our nullspace for A is larger.

Case 3: $\mu = 0$ = not an eigenvalue of K. Now 3.44 says $a_k = b_k/\mu_k$ and our solution is $u = \Sigma_i a_i \varphi_i$, but there may be convergence issues. We have $u = \Sigma_i a_i \varphi_i = \Sigma_i (b_i/\mu_i)\varphi_i$ and now μ_i is on the bottom and we may have the accumulation point $\mu_i \rightarrow 0$ which looks troublesome indeed. So the solution here may or may not converge.

Case 4: $\mu = 0$ = eigenvalue of K. We then have $Ku = \mu u + f = f$ and $K\varphi_0 = 0$. Our equation 3.44 now includes at least one term with $\mu_0 = 0$ let's call it, so we have $a_0\ 0 = b_0$ so as in the previous cases, we can only possibly have a solution if $b_0 = 0$ which means $\langle f,\varphi_0 \rangle = 0$. For all the other eigenvalues, we still have a solution series $u = \Sigma_{i \neq 0} a_i \varphi_i = \Sigma_{i \neq 0} (b_i/\mu_i)\varphi_i$ so we have exactly the same issues with convergence of this series.

3.5 Extremal Principles (223)

This subject comes up about every 20 pages through this book, so it must somehow be important. Let's first look back and review what was said in earlier sections of Chapter 2.

(1) In the E_n section, we had this stuff:

$\|\|A\|\| = \max I_2$ over all x	$I_2 = \|\| Ax \|\| / \|\|x\|\| = [\langle Ax,Ax\rangle / \langle x,x\rangle]^{1/2}$
$\|\|A\|\| = \max I_1$ over all x such that $\|\|x\|\| = 1$	$I_1 = \|\| Ax \|\|$
$\|\|\|A\|\|\| = \max I_4$ over all x	$I_4 = \|\langle Ax, x\rangle\| / \|\|x\|\|^2$
$\|\|\|A\|\|\| = \max I_3$ over all x such that $\|\|x\|\| = 1$	$I_3 = \|\langle Ax, x\rangle\|$

Theorem: for a symmetric operator, $\|\|A\|\| = \|\|\|A\|\|\| = |$ the largest eigenvalue$| = |\mu_1|$

For a non-symmetric matrix, he claims *without proof* that this is all you can show

$|\mu_1| \leq \|\|A\|\|$ and $|\mu_1| \leq \|\|\|A\|\|\|$

(2) Then in the ∞ dimension section we had six theorems, but these only apply to *bounded* operators:

Theorem 1: If A is bounded and $D_A = H$, then $\|\| A \|\| = \|\| A^* \|\|$.

Theorem 2: If A is bounded and $D_A = H$, then $\|\|\|A\|\|\| \leq \|\|A\|\|$

Theorem 3: If λ is in any of the three kinds of spectra for A, then $| \lambda | \leq \|\| A \|\|$.

Theorem 4: If A is symmetric, $\|\|\|A\|\|\| = \|\|A\|\|$.

Theorem 5: If A is symmetric, there exists a domain sequence x_k with $\|\|x_k\|\| = 1$ which, in the limit, would become an eigenvector of $Ax = \lambda_1 x$ where $\lambda_1 = \pm \|\|A\|\|$. BUT, this theorem allows that the limit itself might not exist, so theorem is stated as a $\lim Ax_k \rightarrow \lim \lambda_1 x_k$.

Theorem 6: If A is symmetric *and* completely continuous, then the largest eigenvalue is ± || A ||.

(3) Then earlier in Chapter 3 we had these facts:

$$|\mu| \le \|K\|$$

Theorem 6: If A is symmetric *and* completely continuous, then the largest eigenvalue is ± || A ||.

There is at least one eigenvalue for K, and we know this about it: $|\mu_1| = \|K\|$ // sym HS

But we also showed in page 215 Theorem 1 that $|\mu_1| = |||K|||$.

(4) Now with the above as review, what is new in this section? First, he divides the eigenvalues in this way according to their sign:

$$\mu_1^- \ \mu_2^- \quad \text{..... (0)} \quad \text{.....} \ \mu_2^+ \ \mu_1^+$$

We don't at first know which μ_1 has the largest absolute value. His first demonstration is this:

$$\mu_1^+ = |||\ K\ |||$$ // 3.47

This is a completely new result because it says something about an eigenvalue without an absolute value on it. Earlier we knew that $|\mu_1| = |||K|||$ so I guess what 3.47 is saying is that the largest absolute value eigenvalue is in fact the largest positive one!

Then in 3.48 we have the recursion idea that the next most positive eigenvalue maximizes that same $|||K|||$ thing but subject to the condition that the φ_2 be orthogonal to φ_1 and we then hit μ_2+. OK, I guess this is what we are doing here. In this way we produce a descending sequence of positive eigenvalues and their eigenfunctions. As we go down, we might terminate at 0, or we might have negative ones.

The claim is then on page 224 that the same thing works for the negative eigenvalues starting with the largest negative one but then we are talking some kind of $|||K|||_{min}$ thing instead of the usual max. This is expressed in 3.49 and 3.50.

Comment: You can see generally how the above works. If you have computer scan through the entire L_2 function space, your <Ku,u> will eventually hit the max value μ_1^+ when it finds φ_1^+. And if you then scan orthogonal to this, the max in that case will be φ_2^+ which will hit μ_2^+ , and so on. On the negative side, a scan with no conditions looking to minimize <Ku,u> will hit μ_1^- when it finds φ_1^- , and so on. It really is quite simple and reasonable, and of course suggests a search method for finding eigenvalues and eigenfunctions. The trick is to select a smart trial function parameterization.

Def: if all eigenvalues are non-negative, then <Ku,u> ≥ 0 and K is said to be **non-negative.**

Def: if all eigenvalues are positive, then <Ku,u> > 0 and K is said to be **positive**.
[aside: of L is positive and <Ku,u> = 0, then u = 0]

Def: if all eigenvalues are non-positive, then <Ku,u> ≤ 0 and K is said to be **non-positive**.

Def: if all eigenvalues are negative, then <Ku,u> > 0 and K is said to be **negative**.

Def: Otherwise, the operator K is said to be **indefinite**.

Def: Given u = real, if Ku = real, then K is said to be **real**.

Fact: For integral operator, K is real iff k is real.

Courant minimax principle. Here is what I think this theorem says. First, if you were to maximize <Ku,u> while requiring that all trial solutions u be perp to the higher solutions 1..n, then we know that you will find that $[\max_{u} <Ku,u>] = \mu_{n+1}$. Suppose you were to try to do this process with some *arbitrary set* of n functions instead of the right ones. Then the claim is that you always get $[\max_{u} <Ku,u>] > \mu_{n+1}$ and you only hit the value if you use the exact right set of functions. He calls $\max_{u} <Ku,u> = \nu_{n+1}$. Thus, the theorem says this: $\min_{\psi} \{ \max_{u} [<Ku,u>] \} = \mu_{n+1}$.

Proof: OK we know that if we use the *right* functions, we get [max <Ku,u>] = μ_{n+1} . If we are looking for min [max <Ku,u>] over all function sets ψ_i, that min cannot lie above μ_{n+1}, I agree. We have a specific case which gives [max <Ku,u>] = μ_{n+1} . A *minimum* cannot be larger than a value we have already found. So at this point we know that $\nu_{n+1} \leq \mu_{n+1}$.

Now for the next part. Start with some set ψ_i. He constructs directly a function u that is orthogonal to all the ψ_i. He constructs this u out of the higher φ_i functions. For the specific function u so constructed, he shows that [<Ku,u>] ≥ μ_{n+1}. If we are looking then for *max* [<Ku,u>] over all u, that max value has to be larger than the specific value we found with our specially constructed u. Thus we can conclude that max [<Ku,u>] ≥ μ_{n+1} which says $\nu_{n+1} \geq \mu_{n+1}$.

Combining our two results, we conclude that $\nu_{n+1} = \mu_{n+1}$.

I see on the web that this principle is often stated in terms of E_n and matrices.

3.6 Approximations Based on Extremal Principles (226)

Introduction.

We work here only with K that are: real, symmetric, H-S and non-negative.

Stak shows that you can construct various functions F(u) which, when u = φ_1 , have the property that $F(\varphi_1) = \mu_1$. He gives three specific examples he calls R(u), V(u) and W(u). He then shows that if you write u = $\varphi_1 + \varepsilon \eta_1$ for a functional variation, R(u) has the property that dR/dε = 0 and so this function is "stationary" against variation, and that is why in the plot on page 227 he shows R having zero slope at ε = 0. His other two sample functions do not have zero slope, he claims, and they are also shown in the drawing. So if you are doing some kind of trial function u that is not perfect, if you are off from ε=0 due to your imperfection, your F(u) will likely be closest to μ_1 if you use R(u), since it has zero slope. So the point is that if you want to do a variational thing to find μ_1, then R(u) is a good thing to use.

This R(u) is none other than our friend <Ku,u>/<u,u> which we have been using all along in this chapter, and whose maximum over u is what I call ||| K ||| . Here R(u) is given the official name of **The Rayleigh Quotient**.

The Rayleigh-Ritz Procedure (228)

OK, I grok this pretty well. Start with a set of n independent (but not at this point orthogonal) basis functions. Yes, this is not a complete set, but try to choose good ones. Trial solution is then $u = \Sigma c_i v_i$ with n parameters (real) which are the c_i. So regard $R = R(c_i)$ and set $dR/dc_i = 0$ to get the n equations top of page 229 where the equations are restated three times, lastly in 3.54. The object R* is just a symbol for R(u) with our trial function inserted and which we are trying to maximize so as to get close to μ_1 as shown in the opening equation. We know of course that max[R(u)] $= \mu_1$ but you only hit this value if you stumble upon the exact φ_1 for your trial function, not likely to happen in general. But we try to get close with our n-parameter "fit". So, 3.54 is a matrix equation of the form A**c** = 0 in which R* is a parameter. The quantities α_{ij} and K_{ij} are precomputed from our set of trial basis functions v_i and K. So, the only way the equation A**c** = 0 can given a non-zero solution for vector **c** is if detA = 0. This is then a polynomial in R* with n roots, so we are going to get n real values of R* which give detA = 0 and which then give non-trivial vectors **c** which make our R be stationary in all c_i. Obviously, all $R^*_i \leq \mu_1$ so we should take the largest value, call it R_1*, and that is then our best lower bound for μ_1 and we know that maybe μ_1 is just a little larger than R_1* if we chose a good trial function basis. Also, the vector **c** which goes with this R_1* gives us an approximation for φ_1 . Remember that equation A**c** = 0 has an n-dimensional null space and we are just finding the independent vectors which span this nullspace, and the one that goes with R_1* is our best fit for φ_1. So we have a finite E_n matrix operator problem here which helps us solve our infinite dimensional problem with K.

What about the other R_k* ? Stak claims without proof that you can use the Courant Minimax to show that $R_k^* \leq \mu_k$ so you also have estimates of n-1 other eigenvalues for k=2,3,4...n. It is not at all obvious to me how you use CM to prove this, but the claim somehow seems reasonable to me -- that each stationary point is associated with a μ_k and a set of $c_i^{(k)}$.

On pages 229 he has a set of 5 remarks:

1. The result we get for μ_1 will be more accurate than the result we get for φ_1 , just due to the zero slope business that goes with R(u),.

2. As you increase k from 1, accuracy of everything gets worse.

3. Making n larger improves everything.

4. You should be smart about selecting the v_i based on something physical or geometric. The right phrase I think is you should perhaps use group theory to get v_i to have the right symmetries for the problem.

5. If you choose an orthonormal starting set for v_i , all the math is simpler. I would guess that the lower results (for larger k) would be more accurate since the real solutions φ_i are orthogonal (at least when eigenvalues are non-degenerate).

6. If you could somehow take a complete set so that n = ∞, then you have an infinite matrix equation to try to mess with. In this case, it is easy to show that the R^*_k give the exact eigenvalues. But really you are just saying that if you solve the full ∞ dim problem (if you could), then you would get the exact answers.

EXAMPLE (p 231).

We have a very simple k(x,y) which is symmetric and H-S on the (0,1) box. The EV problem can be solved exactly and the largest eigenvalues are $1/\pi^2 = 1/9.87$ and $1/4\pi^2 \approx 1/39$. This was the problem we treated on page 199 as Example 5 where we associated an ODE with our integral equation, and we knew how to solve the ODE so we got the exact eigenvalues $\mu_n = 1/(n^2\pi^2)$ in that way.

(a) So for our trial set of functions we use just 1 function $v_0 = 1$. Can't get much simpler than that! We get $R_1^* = 1/12$ which is pretty close to 1/9.87 for such a crude trial function set!

(b) We then try a 2-parameter trial function (n=2) which has the right symmetry and we get the results that $R_1^* = 1/9.875$ which is almost perfect, but $R_2^* = 1/170$ which is still pretty bad. A comment that maybe this is really an estimate of μ_3 and not μ_2, where $\mu_3 = 1/9\pi^2 = 1/89$, still not very good. The symmetry used was to be even about x = 1/2 and only the odd solutions have this symmetry.

The Schwarz Iteration Procedure (p 231)

This is a highly technical section that I am not studying too hard. You start by picking a function $f_0(x)$ out of thin air. If K is positive, we know that $\langle Ku,u\rangle > 0$ since all the eigenvalues of $K\varphi_i = \mu_i\varphi_i$ are positive. Since we can in theory expand our $f_0(x)$ in terms of this complete set of eigenfunctions, we know $Kf_0 > 0$ no matter what f_0 we start with. But we want to say K is non-negative only (not sure why we want to say this), so some f_0 will cause $Kf_0 = 0$ and we are told to stop and pick another f_0 if this is found to be the case.

Now define $f_1 = K f_0$. We of course only have $f_1(x)$ written as in integral $\int dy\, k(x,y)f_0(y)$ so I guess we are supposed to compute $f_1(x)$ numerically at each value of x, unless we are lucky and can actually do the analytic integrals. Then $f_2 = Kf_1 = K^2f_0$. And so on so $f_n = K^nf_0$. This is what is meant by the "iteration" word. The general idea is that after a while, the f_n tend to stabilize and not change much in shape from f_{n-1} so you get to a quasi limit. If there were a real limit and things really stopped changing shape, then of course you have that $f_{n+1} = Kf_n$ but $f_{n+1} = \alpha f_n$, just proportional. Then you are saying $Kf_n = \alpha f_n$ and then f_n must be an eigenfunction of K! So the idea is that after iterating, you might zero in on some eigenfunction, but he does not say which one!

The (unknown) Fourier coefficients of f_0 are c_i. He shows then that

$f_n = K^nf_0 = \Sigma c_i\mu_i^n\varphi_i$ // top page 232

Next he defines and then shows that

$a_k \equiv \langle f_o, K^kf_0\rangle = \langle f_i, f_{k-i}\rangle = \Sigma_i c_i^2\mu_i^n$ = positive = "Schwarz constants"

where all along the way we are assuming and using the fact that K is symmetric. Next, the ratio of these constants is given a name

$\theta_k \equiv a_{k+1}/a_k$ = "Schwarz quotient"

Theorem 1. In the limit, $\theta_k \rightarrow \theta \leq \mu_1$ monotonically from below.

Theorem 2. (a) If it happens that $c_1 \neq 0$, then $\theta = \mu_1$.
(b) If $c_1 = c_2 = .. = c_{m-1} = 0$, then $\theta = \mu_m$

But of course if you just pick $f_0(x)$ when you start the problem, you have no idea what the φ_i are, so you don't really know anything about which $c_i = 0$. These theorems are proved. After showing all this detail, on his REMARK on page 234 he says this whole scheme is non-practical for anything other than μ_1 . If you can pick an f_0 (perhaps by symmetry) that has $c_1 = 0$, then you can do all right for μ_1, but then as you continue the calculation in going after μ_k, rounding errors make $c_1 \neq 0$ and you end up just zeroing in on μ_1 again instead of μ_k. All in all, I would call this a low priority section of this chapter. This procedure is not in Sheid, and has a very slim web presence.

***Upper* Bounds to Eigenvalues**

So far we have been talking *lower* bounds since max[<Ku,u>] = $\mu_1$for example. That is, by varying u and doing better and better, we approach μ_1 from below, so we are making lower bounds. Obviously it would be nice if you could "enclose" μ_1 from both above and below so you know how close you are.

This section mentions two methods for getting upper bounds and doing this enclosure.

First, Stak points out that $\| Kv - \mu v \|^2 = 0$ if v is the exact eigenfunction, and to the extent you have missed with v, it is larger than 0. When you apply $d/d\mu = 0$ to this thing, you get $\mu = \langle Kv,v \rangle$, Again, if v were perfect, then you would get $\| Kv - \langle Kv,v \rangle v \|^2 = 0$ by subbing this in. Algebra apparently shows that $\| Kv - \langle Kv,v \rangle v \|^2 = \|Kv\|^2 - \langle Kv,v \rangle^2 \equiv \alpha^2$. So as $\alpha \rightarrow 0$, you zero in if you could vary v. This discussion is motivation for the first theorem below.

The WBFS Theorem. Let v(x) be an arbitrary but normalized function. Compute these two items:

$\sigma = \langle Kv,v \rangle$ // sort of a candidate eigenvalue, but wrong since v not perfect

$\alpha^2 = \|Kv\|^2 - \langle Kv,v \rangle^2$

The theorem says then that your eigenvalue is bounded between $\sigma-\alpha$ and $\sigma+\alpha$, a range of width 2α. Obviously you are motivated to try different v(x) functions to make α as small as possible.

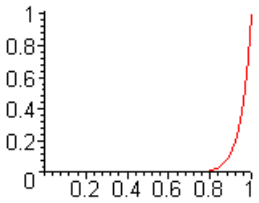
The middle section of page 235 adds some technical detail so that you know exactly which eigenvalue you are bracketing in this manner! So this involves getting a rough idea where μ_2 is and making sure you stay away from it while working on μ_1. You find some $\ell_2 > \mu_2$ somehow. He then suggests some kind of application of the Schwarz iteration thing which makes α smaller and smaller. The picture for all this is shown bottom of page 235.

The last paragraph on page 235 just quotes some alternative upper bound due to my old friend **Kato** (I bet it is the same Kato as in Messiah degenerate perturbation theory, but don't know). This other upper

bound has pros and cons compared to the WBFS one. The Kato bound is developed in Exercise 3.16b. It seems a bit unfriendly that Stakgold does not give exact references for these things in his book.

3.7 Continuity and Uniform Convergence: bilinear series and iterated kernels

First, a reminder on uniform convergence vs other convergences. Consider the set of functions x^n on (0,1) as n gets large. Here is n=20



The sequence approaches some narrow peak just to the left of x = 1. The sequence converges *pointwise* to the function f = 0 in this interval (ie, at any x), but it does not converge *uniformly* over the interval to f = 0 due to the problem at the right end. You can always get close enough to 1 so that $x^n > 1/2$, say. This same example has $\| x^n - 0\| \rightarrow 0$ as an L_2 norm (convergence *in the mean*) because the area of x^{2n} gets arbitrarily small as n increases. So this example shows how a sequence of functions can converge in the mean but not converge uniformly over the interval.

pointwise convergence of a function to the function 0: achieved by $f_n(x) = x^n$

For any ε, $\exists$ n $\ni$ $|f_n(x) - 0| < \varepsilon$ for any particular x you choose. That is, n = n(x)

uniform convergence of a function to the function 0: not achieved by $f_n(x) = x^n$

For any ε, $\exists$ n $\ni$ $|f_n(x) - 0| < \varepsilon$ for all x $\in$ (0,1). That is, n must work for all x

convergence in the mean of a function to the function 0: achieved by $f_n(x) = x^n$

For any ε, $\exists$ n $\ni$ $\| f_n(x) - 0\|_{L2} < \varepsilon$

Here is another fact we need to throw in here:

> **Uniform convergence theorem**. If $(f_n)_n$ is a sequence of continuous functions which converges *uniformly* towards the function f, then f is continuous as well.

Perhaps there is a typo and they just mean "if $\{f_n\}$ is a sequence". But the main point is this: if $f_n \rightarrow f$ uniformly over an interval, then if the f_n are continuous functions, so is the limit function f. So this concept invokes the phrase "uniform continuity".

Now, our old Theorem 4 of page 218 said that the φ_i for non-zero eigenvalues μ_i form a complete set for functions in the range of K. Then on page 219 went on to say that an *arbitrary* (L_2) function f can be expanded on this limited set φ_i provided you add a piece which is in the nullspace of K. So we wrote

$$f = h + \Sigma_i c_i \varphi_i \qquad c_i = \langle f,\varphi_i \rangle \qquad Kh = 0$$

$$\Rightarrow\ Kf = K(h + \Sigma_i c_i \varphi_i) = K\,\Sigma_i c_i \varphi_i = \Sigma_i c_i K\varphi_i = \Sigma_i c_i \mu_i \varphi_i = \Sigma_i \langle f,\varphi_i \rangle \mu_i \varphi_i$$

$$= \Sigma_i \langle f, \mu_i \varphi_i \rangle \varphi_i = \Sigma_i \langle f, K\varphi_i \rangle \varphi_i = \Sigma_i \langle Kf, \varphi_i \rangle \varphi_i$$

where we have just written Kf in various ways where we assume K is symmetric and maybe HS too. The last result shows that φ_i is complete for a function in the range of K.

In all these series Σ_i, we have always talked in terms of *convergence in the mean*. In all our earlier work, when we say a series converges, we always meant "in the mean" which is the L_2 norm of the difference between a finite series and its ∞ series limit. This is the case for all the sums shown above.

It would be a stronger statement if we could say that the various sums converge *uniformly* over the interval. This is what **Theorem 2** on page 237 claims. The concluding statement in the proof is page 238 B which is our last result above. It says that the series $\Sigma_i \langle Kf, \varphi_i \rangle \varphi_i(x)$ converges uniformly to f(x) for all x in (a,b).

How is this proved? The second expression in 3.63 is for a specific x in the interval (though x is not shown explicitly) . In the last expression, we have the known convergence in the mean || ...|| thing which we know can be made arbitrarily small. We then have to make an added assumption that the *single* integral of the k^2 kernel is finite (HS says *double* integral is finite). This is what appears on the last line left side of 3.63. If this thing is finite, or < M, then the LHS $\rightarrow$ 0 for any x, and we have shown uniform convergence.

As a corollary, he claims that our Section 3.4 solution to Ku = μu + f, which was this

$$u = -f/\mu + \Sigma_i\, b_i\, \mu_i\, /\mu(\mu_i - \mu)\varphi_i\ , \qquad \textbf{// agrees with 3.45}$$

also have a uniformly convergence series. I am not sure why, but I imagine that you show this based on the idea that our Ku series are uniformly convergent as in page 238 A and B. I seem to recall that the terms u + f/μ = g are in the range of K and so can be expanded on the φ_i , and Theorem 2 just showed that such an expansion is uniformly convergent.

Now, what does **Theorem 1** have to do with the price of eggs? This theorem basically shows that, if we assume that $k^2(x,y)$ is continuous, then each eigenfunction of K $\varphi_i(x)$ is continuous. If this fact was used in proving Theorem 2, I don't know where. Maybe Theorem 1 and Theorem 2 are completely unrelated at this point. If we invoke the uniform continuity idea above, then it would be true that if $k^2(x,y)$ is continuous, then the φ_i are continuous by Theorem 1, and then any of our series will yield a continuous function as well, so this would tell us for example that u(x) shown above would be continuous.

Now we come to developing **Theorem 3**. Let w(x) be twice differentiable and let w"(x) be piecewise continuous (there can be some points where it is not continuous). Let w(x) satisfy our usual ODE boundary conditions, and imagine that Mw = f, where M is our "adjusted" version of operator L we dealt with earlier on page 219. Our assumptions do make f be piecewise continuous. So w is a solution of an

ODE (and we know such a solution is unique from earlier work). We know the integral equation version says w = Gf where K = G is the Green's function of the ODE, so obviously w is in the range of K. Thus, we can write w as Fourier $\Sigma_i c_i y_i$ and we showed in Theorem 2 that this converges uniformly. I guess we know somehow that k^2 for a Green's function is singly integrable to a finite value. He claims that continuity of k is involved here, but that seems different from the assumed continuity of k_2. Perhaps it is easy to show that continuous k implies continuous k^2 and therefore the φ_i are continuous in $My_i = \theta_i y_i$ from Theorem 1, and therefore our Fourier expansion $w = \Sigma_i c_i y_i$ makes w be continuous. But all that Theorem 3 says is that $\Sigma_i c_i y_i$ converges uniformly to w. The idea that w comes out being continuous seems at least consistent with assuming that w"(x) is piecewise continuous. Stakgold gets a low grade from me on his clarity here!

Bilinear Expansions

If you consider both k and $k_2(x,\xi)$ just as functions of x, you can write $k_2 = Kk$. This means that k_2, as a function of x, is in the range of K, and -- from Theorem 4 page 218 (K symmetric and HS) -- can be expanded in the φ_i EF's of K with no "extra pieces", so $k_2(x,\xi) = \Sigma_i a_i(\xi)\varphi_i(x)$. I think k_2 like k is symmetric, so it must also be true that $k_2(x,\xi) = \Sigma_i \bar{a}_i(x)\bar{\varphi}_i(\xi)$, though Stak does not make this argument, but I like it, It suggests to me that we must have $a_i(\xi) = k_i\bar{\varphi}_i(\xi)$ for some k_i so that then $k_2(x,\xi) = \Sigma_i\, k_i\, \varphi_i(x)\, \bar{\varphi}_i(\xi)$. This form is in fact correct and it turns out that $k_i = \mu_i$. This is all shown directly in the algebra on top of page 239 (including my many pencil steps). This sum over $\varphi_i(x)\, \bar{\varphi}_i(\xi)$ with a weighting function is called a "bilinear expansion". Moreover, from Theorem 2 on page 237, we know that this bilinear expansion for k_2 (in either variable) must converge uniformly on the interval. Stak then claims that in fact, k_2 converges uniformly over the entire square, not too surprising to me.

I agree that with a small amount of work I could show the generalization of what we just showed, and this gives us Theorem 4 on page 239 which says:

Theorem 4: $k_n(x,\xi) = \Sigma_i \mu_i^n\, \varphi_i(x)\, \bar{\varphi}_i(\xi)$, and converges uniformly on the square, $n \geq 2$.

Note that the weight function is now μ_i^n . If we then integrate $k_n(x,x)$ over x, we can use the completeness of the φ_i to say $\int dx\, \varphi_i(x)\, \bar{\varphi}_j(x) = \delta_{ij}$ so $\int dx\, \varphi_i(x)\, \bar{\varphi}_i(x) = 1$, to get

$\int dx\, k_n(x,x) = \Sigma_i \mu_i^n \geq \mu_1^n$ // **sum rule** on eigenvalues, and inequality

so that, if K is non-negative,

$[\int dx\, k_n(x,x)]^{1/n} \geq \mu_1$ // which is 3.66 $n \geq 2$

which gives an UPPER bound for μ_1 , to complement our usual variational lower bounds.

Another fact is noted in passing:

Theorem None: $K^n\varphi_i = \mu_i^n\varphi_i$ // which is trivial to show. $n \geq 1$

Notice that we have claimed nothing regarding the case n = 1 in Theorem 4. Under certain conditions, we can make claims regarding n = 1, and this is the content of:

Mercer's Theorem: If k has some extra properties:

(1) all but a finite number of the μ_i of K have the same sign ("quasi definite")
(2) k is continuous
(I suspect k still has to be symmetric and HS)
(3) the second result below also requires k to be non-negative.

then the n=1 results apply and we have:

$k(x,\xi) = \Sigma_i \mu_i \varphi_i(x)\overline{\varphi}_i(\xi)$ converges uniformly on the square

$\int dx\, k_n(x,x) \geq \mu_1$ provides an upper bound for μ_1

Even if we don't have the Mercer extra conditions, the k series above converges in the mean.

Comment: Recall that $k(x,\xi)$ is the Green's function of the associated ODE problem. I am used to seeing Green's written as the above bilinear expansion, I think. Here we have proven that such an expansion is valid given our various conditions on K. Stak does not mention this, nor do I see it mentioned in Chapter 1. However, an example of such a Green's sum might appear in Exercise 3.19.

Exercises p 240-241. I read through them, 3.19 looks especially interesting, all are good. If I were doing a real course, I would do all these things. Maybe in some future "pass" I will do that.

3.8 Approximation Methods for Solving Integral Equations

We are trying to solve $Ku - \mu u = f$ in general when $\mu \neq$ EV. We shall also be looking at $K\varphi_i = \mu_i \varphi_i$ in some of the following sections. No error estimates are given for the various methods, but we are referred to some 1958 tome by two Russians, reference bottom of page 242. **Kantorovich and Krylov**.

1. Lattice Integration (242).

Replace the integral with an n-fold sum, that is, put the integral equation "on a lattice". The integral is then the sum of n terms. We now talk about the n numbers u_i or f_i. The integral equation becomes an n x n matrix equation 3.70 and we can solve for the u_i in the usual way, such as Cramer's rule. We can then somehow interpolate between the u_i to get u(x). One easy method of interpolation is provided by the integral equation itself, written as p 242 A.

Here we are just dividing our "area" into thin vertical rectangles where the height is the average of adjacent function values. Recall from Scheid that you get better results if you use thin trapezoids (doing a linear fit within each strip, Trapezoidal Rule) or of you use little quadratic fits (doing a quadratic fit within each strip's region, Simpson's Rule).

Stak comments that Fredholm originally (perhaps 1903) did things in this way when he developed his theory of integral equations. He did the n x n matrix approach on the lattice, then took the limit $n \to \infty$ to

get the theory. This approach gives the right results, but has been replaced by H-S theory (perhaps 1912), and so Fredholm's approach is only of historical interest.

Notice that this section is still describing an analytic method. We are not just blindly doing numeric integration of anything as we might do in a variational method.

2. Approximating the kernel as a finite separable sum.

We have some kernel $k(x,\xi)$ and we expand it on some arbitrary finite set of n orthonormal functions $q_i(\xi)$ so that we then have $k(x,\xi) = \Sigma_{i=1}^{n} p_i(x)\, q_i(\xi)$ where $p_i(x)$ are the "coefficients", but you can then compute $p_i(x) = \int d\xi\, k(x,\xi)\, q_i(\xi)$, so now you know both the $q_i(\xi)$ and $p_i(x)$, and now we have a finite separable representation of k. Since we did not use an *infinite* set, this is of course just an *approximation* to $k(x,\xi)$. We then write (3.71) our candidate solutions as $\mu u(x) = \Sigma_i c_i p_i(x) - f(x)$ where we then seek to find the c_i coefficients. We shuffle things and are not surprised to obtain in p343 B an n x n matrix equation for the c_i of the form $A\mathbf{c} = \mathbf{d}$ where the d_i are integrals of f(x) against the q_i. We solve this for the c_i and have our approximate solution for u(x) ! I would add that it would be good to use the symmetry of the underlying physical problem when selecting a good finite set of basis functions. Obviously larger n gives better results. [Caveat: this method requires that the $p_i(x)$ be independent, not clear how one can cause this to be true for an arbitrary k. In general, each set p_i and q_i needs to be an independent set.]

3. Approximating u(x) in terms of a finite set of independent functions $v_i(x)$.

In 2 above we approximated the ξ-part of the $k(x,\xi)$ with such a set. Here we directly approximate the solution u(x). To the extent that $u(x) = \Sigma_i c_i v_i(x)$ is good, 3.72 is good (just $Ku - \mu u = f$ with u replaced twice by this sum). In 3.72, we combine the two sums by defining $g_i(x)$ as shown, and these are n functions that we can compute from k and v_i. The question is now: how should we pick the c_i to make 3.73 ($\Sigma c_i g_i = f$) most true?

(1) Best in the Mean. One obvious thing to do is pick the c_i to minimize the L_2 difference in the two sides, and that gives the n x n matrix problem $A\mathbf{c} = \mathbf{d}$ shown in 3.74.

(2) Galerkin's Method (Russian ~1915). His method is to require that both sides of $\Sigma c_i g_i = f$ have the same projections onto some *new* set of n functions $w_i(x)$, as shown in 3.75. That is to say, you pick a subspace and make both sides equal there. These n equations are called Galerkin's equations and they form a matrix equation $A\mathbf{c} = \mathbf{d}$ which you then solve for **c**.

Example 1: try $w_i(x) = g_i(x)$. This replicates method 1, best in the mean.

Example 2: try $w_i(x) = v_i(x)$. This gives 3.77 and 3.78.

Example 3: same as Example 2, but set f = 0 so we have the EV problem. This gives p 245 A which is the same as the set of equations from Rayleigh-Ritz, 3.54 on page 229, with μ now in place of R*. Recall in R-R how the determinant of this thing gave R* values which were "close to the eigenvalues μ_i" and we make exactly that claim for the Galerkin Method. It is interesting that we arrive at the R-R result without having to set any variation derivatives to 0.

Stak claims that the best in the mean method is optical. Web says Galerkin had a tough life, died in 1945, his method also applies to ODE's and other problems, and is widely used in the world.

4. A Variational Principle (245)

Equation of interest here is Au= f which is inhomo1 ($\mu = 0$ in inhomo2). Define this functional:

$$I(v) \equiv 2\langle f,v\rangle - \langle Av,v\rangle$$

Stak states and then gives a very simple proof of this two part theorem,

Theorem (where A is real, symmetric, and positive:)
(a) If Au = f has a solution u, that solution u maximizes I(u).
(b) If some function v maximizes I(v), then that function is u, the solution of Au = f.
(c) (my own part). If Au=f, then I(u) = <f,u> (trivial to show)

The proof of (a) is this: I(u)-I(v) = <A(u-v),u-v> if we assume u solves Au = f and v arbitrary. Since A is positive, we know that I(u) > I(v) in general, but if v = u, then I(u) = I(v) . Thus, I(v) is maximized when you select v = u.

The proof of (b) is to show that $dI/d\varepsilon = 0$ for a variation ε about a max of I(u) at u, and that then yields the fact that Au = f.

Remark 1: Could use (b) so show a solution exists (by finding something that max's I). If we already know a solution exists, could use (a) as a variation deal to find the solution.

Remark 2: Replace I with -I and can talk a minimum principle, and this is the language of minimal energy in physics problems, which will appear somewhere in Stak Vol II he says.

Remark 3a. Since A is positive and symmetric, you can regard $[f,g] \equiv \langle Af,g\rangle$ as a new inner product. It is the old inner product but with a "weight function" which is operator A. Then the natural norm induced by this scalar product would be $|f| = \langle Af,f\rangle$. You can then show that I(u) - I(v) = [u-v, u-v] when u is a solution of Au=f and v arbitrary, and then this makes it very obvious that I(v) only equals I(u) when u = v.

Remark 3b. Using this new norm with the CSI, we get p 247 A and B which we just rewrite as 3.82. The idea here is that v = u will max the RHS and make it equal to the LHS. But this will also be true if v = cu for any real constant c. We can write this then as: $\langle f,u\rangle = \langle Au,u\rangle = \max_{\mathbf{v}} J(v)$ as in 3.82a.

The idea here is that you can not worry about the scale of your trial function v and concentrate on getting its shape right, and this can make the variational problem easier. He shows that if you find some u_0 that does the maxing of J, you can find the c in $u = cu_0$ as shown, so that is no problem.

Remark 3c. Suppose we want to solve Ku - μu = f (3.69, inhomo2) which we can write as Au = f where A = K – μI. If it happens that $\mu < 0$, then this operator A is symmetric and positive, just like K. Now suppose we apply our variational idea using part c of our theorem above so that $\langle f,u\rangle = I(u) = \max_{\mathbf{v}} I(v)$. We then write out the fact that $I(v) = 2\langle f,v\rangle - \langle Av,v\rangle$ with A = K – μI so $I(v) = 2\langle f,v\rangle - \langle Kv,v\rangle + \mu\langle v,v\rangle$. So

$$\langle f,u\rangle = \max_{\mathbf{v}} I(v) = \max_{\mathbf{n}} \{ 2\langle f,v\rangle - \langle Kv,v\rangle + \mu\langle v,v\rangle \}$$

and this is what equation I marked 3.80' says on page 247. But using the J(v) alternate method, we get the equation I marked as 3.82'. As noted, 3.80' is maxed when v = u, but 3.82' has the nice property that it is maxed when v = cu.

5. The Rayleigh-Ritz Equations (248)

We want to solve Au=f and now we take a trial function $v = \Sigma_{i=1}^{n} c_i v_i$ with n selected v_i. If we jam this into our variation principle $I(u) = \max_v I(v)$ we set $dI/dc_i = 0$ and get p 248 A,

$\langle f, v_i \rangle = \Sigma_j^n \langle Av_i, v_j \rangle c_j$ // The Rayleigh-Ritz Equations

This should not be confused with the Rayleigh-Ritz Procedure we used on pp 228-231 in our approximate method to find the EVs and EF solutions of Ku = μu.

Now, if we set A = K – μI, we can imagine we are solving Ku - μu = f, and it turns out that the Rayleigh-Ritz Equations shown above are the same as the Galerkin equations for the Ku - μu = f problem as shown in 3.78. (where the Galerkin weight functions are taken the same as the v_i).

Now just suppose you could magically select your v_i functions such that $[v_i, v_j] = \delta_{ij}$ using that "new" scalar product. Then the Rayleigh-Ritz Equations become just

$\langle f, v_i \rangle = c_i$

and you are done, and your solution trial is then $v = \Sigma_{i=1}^{n} c_i v_i = \Sigma_{i=1}^{n} \langle f, v_i \rangle v_i$. In this case, you see that your approximation for the solution of Au=f is just the first n terms of an infinite series which is the exact solution, assuming the v_i are a complete set in terms of the [a,b] inner product.

Exercise 3.23. Amazingly, this treats the "extended eigenvalue problem" which I ran into while doing the degenerate perturbation theory stuff, where it was called "the generalized eigenvalue problem" and terms like "matrix pencil" occurred. See Kato Round One notes (in Messiah file) for more detail. Here Stak shows that this generalized EV problem has features very similar to the regular EV problem. He does suggest as I did in those notes that M positive means M^{-1} exists.

3.9 Non-symmetric HS operators (250)

The problem here is that the EV problem EF's might not be a complete set for solution expansions of the inhomo problem. A simple example is given of a non-symmetric H-S kernel which just yields the integral operator on (0,1). The inhomo IE has a fine solution, but the EV problem has no EV's or EF's !!!

We are going to "salvage" some of the symmetric theory benefits by making use of the **left and right iterates** of K, where K is H-S but not symmetric:

$K_L = KK^*$
$K_R = K^*K$

Stak shows that (1) these are both symmetric and H-S. (2) both are non-negative so write EV's as μ^2 instead of the usual μ. (3) they have the same eigenvalues and the same multiplicities. (4) the eigenfunctions are simply related to each other.

Suppose $K_L\, v = \mu^2 v$. Then you can show that $u = Kv$ satisfies $K_R\, u = \mu^2 u$.

Suppose $K_R\, u = \mu^2 u$. Then you can show that $v = K^* u$ satisfies $K_L\, v = \mu^2 v$.

Suppose u_n is orthonormal set for K_R with μ_n^2. Then $v_n = (1/\mu_n)\, K^* u_n$ is ON set for K_L with μ_n^2.

Suppose v_n is orthonormal set for K_L with μ_n^2. Then $u_n = (1/\mu_n)\, K v_n$ is ON set for K_R with μ_n^2.

With this as warm-up, we now have a few theorems.

Theorem 1: $Ky = 0 \;\Leftrightarrow\; \langle y, v_n \rangle = 0$. That is, if y is orthogonal to *all* the EF's v_n of K_L, then y is in the nullspace of K, and vice versa.

Proofs are short and simple.

Corollary 1: The v_n span the perp space of K.

Assume true, then $x \in N_K^{\perp}$ would mean that you can write $x = \Sigma_i c_i v_i$. To show claim is true, we would have to show that x.y = 0 for any y in N_K (Ky=0). This would prove that our x really was in $N_K^{\perp}$. But y.x = y. $\Sigma_i c_i v_i = \Sigma_i c_i$ y. v_i. But theorem 1 says if y in nullspace, then y. v_i= 0, so we have shown y.x = 0 and verified the corollary's claim.

Corollary 2 just says the usual thing that a function can be decomposed into its part in the nullspace and its part in the perp space, where f_0 is in the nullspace of K and where the perp space part can be expanded on the v_i. So he writes $f = f_0 + \Sigma_i\, f.v_i\; v_i$

Theorem 2 then says if g = Kf (g in the range of K), then g can be expanded on the u_i which, recall, are the EF's of K_R. Two forms are given: $g = \Sigma\; (g.u_n)u_n$ and $g = \Sigma\, \mu_n\, (f.v_n)\, u_n$. The proof is trivial.

Both theorems and corollaries above are true if you replace: $K \leftrightarrow K^*$ and $u_n \leftrightarrow v_n$ everywhere. He gives these four items "1" suffices. Here they are:

Theorem 1A: $K^* y = 0 \;\Leftrightarrow\; \langle y, u_n \rangle = 0$.

Corollary 1A: The u_n span the perp space of K*.

Corollary 2A: $f = f_0' + \Sigma_i\, f.u_i\; u_i$ where f_0' is in the nullspace of K*.

Theorem 2A: If $g = K^* f$, then $g = \Sigma\; (g.v_n)v_n = \Sigma\, \mu_n\, (f.u_n)\, v_n$ // now bottom page 253.

Now we consider trying to solve Ky=h where K is HS but non-symmetric. Stak shows that the range function h cannot be completely general for a solution to exist, but is restricted by two conditions:

(1) (3.94) h must be in the perp space of K*

So if K*z = 0, then you must have h.z = 0, called "consistency conditions". (I would add that for Ky=h to have a solution, y has to be in the perp space of K, but that tells us nothing about h.)

(2) (3.95) $\Sigma_i \mid h.u_i|^2 / \mu_i^2 < \infty$

which I think just says y is L_2, $\| y\| < \infty$, since can write solution $y = y_0 + \Sigma_i [(h.u_i)/\mu_i] v_i$. If there are an infinite number of EV's μ_i , this condition can be pretty serious, remembering that $\mu_i \rightarrow 0$ for limits.

Theorem: If h satisfies the two conditions listed above, then Ky = h has the solution just shown above for y, where y_0 is any solution of $Ky_0 = 0$, ie, y_o is any nullspace solution.

$$y = y_0 + \Sigma_i [(h.u_i)/\mu_i] v_i$$

where u_i = EF's of K_R, and where v_i = EF's of K_L, and where μ_i = sqrt (EV). So this is our best result for HS but non-symmetric K. We have indeed "salvaged" quite a lot.

Question: What is the analogous solution for symmetric K ?

Our solution for symmetric HS K from Section 3.4 above Case 3 is just this: Ku = f =>

$$u = \Sigma_i(b_i/\mu_i)\varphi_i = \Sigma_i[(f.\varphi_i)/\mu_i]\varphi_i$$

where φ_i are the EF's of K. In case 3 we assumed μ=0 was not an EV of K, so there is no homo solution to add to u. But in case 4 we allow $\mu = 0$ to be an EV of K, then you would add a u_0 homo solution. Recall that μ_i^2 are eigenvalues of K^2 in our case here, so solutions are very analogous! Basically we cannot use the φ_i for the non-symmetric case, but we can make use of the u_i and the v_i .

Example:

This is the same example with which we opened the section, where Ky is simply "integration". The kernel is HS but not symmetric. K has no eigenvalues. Stak states the kernel and then computes K_L which is very simple. He then ponders the EV's and EF's of K_L which we called v_i. He shows that the integral EV equation can be converted to an ODE for v with certain BC's. It is just a trig type ODE and he just quotes the solutions for v_n and μ_n^2 . He then considers K_R without detail and shows the u_n which are just sin instead of cos. We now have all our ingredients to consider Ky=h. Recall our two conditions mentioned above. The first is a non-issue because K and K* have no nullspaces (no eigenvalues, so μ=0 not EV). The second condition $< \infty$ is stated in 3.98 for our example. If it is satisfied, then our general Ky=h solution is given in 3.99 which is just 3.96 applied to our example. We know of course that our solution has to be y = dh/dx . He then expands h on the sin series (cos not needed for some reason, he claims the u_n are a complete set, I guess this has to do with the arguments of the trig functions) and does term by term diff to get h'(x) and this replicates our claimed solution 3.99. But term by term diff is only justifiable it 3.98 $< \infty$ is true, so everything is consistent!

Exercises.

3.27. The first shows that ||K|| = $|\mu_1|$ even if K not symmetric.

The remaining two examples consider solving the inhomo2 equation Ky - λu = f which we did NOT consider anywhere in the section above.

3.28 It turns out that (1) knowing the EF's of K_L and K_R in this case does not help; (2) eigenfunctions of the corresponding iterated kernels for K-λI would help, but are too complicated.

3.29. Here we assume that K has some EF solutions μ_n and φ_n while K* has some other ones $\bar{\mu}_n$ and ψ_n. These EF sets are not the same, in general. Stak shows that $\langle \varphi_i, \psi_j \rangle = \delta_{ij}$ so this is not an orthonormal set but instead a **biorthogonal set.** The solution of Ky - λu = f can be found as 3.101, but only if φ_n forms a basis, which is really begging the question, so I think this exercise buys us little, except we do learn a new term. The idea was to try to use the separate EV sets of K and K*, such as they are, and see what can be done with them.

And so we finally come to the end of Chapter 3 on Linear Integral Equations, about 68 pages of book.

Comments on this chapter: I have at least two other books which talk about integral equations: M&M and R-N (Frigyes Riesz and Béla Szőkefalvi-Nagy). As I browse those book sections, I see that Stakgold has omitted various methods of Fredholm that involve, for example, the **Fredholm Determinant** and discussion of the **resolvent** whose expansion gives the Newman series. I feel I should know a little about these things. My guess is that the H-S methods surpass these older Fredholm methods so they are only of historical interest, I think that is what Stakgold would say.

Also, Stakgold's 2-volume set is really about ODE's, and he could argue that his integral equations section is just background material for the ODE work, so he does not have to be exhaustive on the subject of integral equations.

Notes Discovered. I just found in my Math I binder that I have lots of notes on integral equations, which are stored in a section labeled Fredholm, including notes from the book of Mikhlin which Stak mentions in his references page 258. These notes talk about the two omitted topics I mention in the above Comments.