Stakgold Chapter 3 Meta Meta Notes PhL 4.22.09

This chapter is p 191-258 or 67 pages of the book. My raw notes are 33 pages, meta notes are 14 pages, and these meta-meta notes are about 10 pages including the Fredholm stuff which Stak omits.

Chapter 3: Linear Integral Equations

3.0 The Fredholm Determinant and Related Subjects (PhL) 1

3.1 Introduction with Six Examples (191) 3

3.2 The Neumann Series Method (206) 4

3.3 The Spectrum of a self-adjoint H-S operator. (p 212) 4

3.4 The solution of Ku = μu + f for K symmetric and H-S. (220) 6

3.5 Extremal Principles (223) 6

3.6 Approximation Methods Based on Extremal Principles (226) 7

3.7 Continuity and Uniform Convergence: bilinear series and iterated kernels 8

3.8 Approximation Methods for Solving Integral Equations 10

3.9 Non-symmetric HS operators (250) 10

3.0 The Fredholm Determinant and Related Subjects (PhL)

Stakgold *does not even mention* this whole area, but I reviewed it and have lots of notes. The method is only really useful if the λ in our Fred 2 integral equation φ = f + λKφ is "small". In his Chapter 3, Stak does not address the issue of this free parameter λ being small, apart from mentioning the Neumann series. But in QED and scattering perturbation theory it is of great interest, and it showed up a lot in my Geoff Chew hadronic work

The basic situation is that we want to solve the integral equation φ = f + λKφ as

$$\varphi = (1 - \lambda K)^{-1} f = (1 + \lambda K + \lambda^2 K^2 + ...)\, f = f + \Sigma_{m=1} \lambda^m K^m f = (1 + \Gamma)\, f$$

where Γ/λ is called **the resolvent**. This is just the sum of the Neumann series Stak mentions below. If you can find the resolvent, you have basically solved the integral equation. But it is easy to show that, if we put our problem on an N x N lattice, then

$$(1 - \lambda K)^{-1} = [\ \mathrm{cof}\,(1-\lambda K)\ ]^T / \det(1 - \lambda K) = 1 + \Gamma$$

Now if λ is in some sense "small" (perhaps λ = 1/137), then we can derive power series expansions (in λ) for the two objects $[\ \mathrm{cof}\,(1-\lambda K)\ ]^T$ and det(1 - λK) appearing in the above operator. If we keep only the first few terms in these expansions, then we have an approximate solution for the resolvent. We first proceed as if K were an N x N finite matrix where $K_{ij} = k(x_i, x_j)$ and we show that

$$\det(1-\lambda K) = 1 - (\lambda)\,\Sigma_{\mathbf{i}}\, K_{\mathbf{ii}} + (\lambda^{\mathbf{2}}/2)\Sigma_{\mathbf{ij}} \begin{vmatrix} Kii & Kij \\ Kji & Kjj \end{vmatrix} - (\lambda^{\mathbf{3}}/3!)\Sigma_{\mathbf{ijk}} \begin{vmatrix} Kii & Kij & Kik \\ Kji & Kjj & Kjk \\ Kki & Kkj & Kkk \end{vmatrix} + \ldots$$

and then we take the limit $N\to\infty$ to get

$$\det(1-\lambda K) = 1 - \lambda \int dx\, k(x,x) + (\lambda^{\mathbf{2}}/2) \int dx\, dy \begin{vmatrix} k(x,x) & k(x,y) \\ k(y,x) & k(y,y) \end{vmatrix}$$

$$- (\lambda^{\mathbf{3}}/3!) \int dx dy dz \begin{vmatrix} k(x,x) & k(x,y) & k(x,z) \\ k(y,x) & k(y,y) & k(y,z) \\ k(z,x) & k(z,y) & k(z,z) \end{vmatrix} + \ldots. \text{ (goes on forever)}$$

and this thing is then called "**the Fredholm Determinant**". Interestingly, it can also be written entirely in terms of traces of powers of the operator K.

$$\det(1-\lambda K) = \exp\left[-\Sigma_{\mathbf{n=1}}^{\infty}\ \lambda^{\mathbf{n}}/n\ \mathrm{tr}(K^{\mathbf{n}})\right]$$

And there are various other odd theorems floating around.

It is also easy to show, going back to the N x N matrix world, that the following is true:

$$[\Gamma_{\mathbf{ij}}/\lambda] = [N_{\mathbf{ij}}/\lambda] / \det(1-\lambda K) \qquad (*)$$

where $N_{\mathbf{ij}} \equiv [\mathrm{cof}\,(1-\lambda K)]_{\mathbf{ji}} - \det(1-\lambda K)\,\delta_{\mathbf{i,j}}$

which is basically our result shown above where $N_{\mathbf{ij}}$ is a "minor" (more or less) which appears when you write down the cofactor matrix.

$$[\mathrm{cof}\,(1+B)]_{\mathbf{ji}} = (-1)^{\mathbf{i+j}}\,\mathrm{minor}(1+B)_{\mathbf{ji}} = (-1)^{\mathbf{i+j}}\,\det_{\mathbf{rows\neq j;\,cols\neq i}}$$

When we again take $N\to\infty$ we get that

$$[\Gamma(x,y;\lambda)/\lambda] = [N(x,y;\lambda)/\lambda]/\det(1-\lambda K) = \textbf{the resolvent}$$

where the fact that Γ and N depend on λ is shown explicitly. A standard notation people use is this:

$D_{\mathbf{1}}(x,y;\lambda) \equiv [N(x,y;\lambda)/\lambda]$ **Fredholm's First Minor** (some people erroneously drop $/\lambda$)

$D(\lambda) \equiv \det(1-\lambda K)$ **The Fredholm Determinant**

Thus, if we know the small λ expansion for the determinant *and* the first minor, we have at once an expansion in λ for the resolvent, and this in turn solves our integral equation!! The expansion for the minor is this, for finite N and then as $N\to\infty$:

$$[N_{ij}/\lambda] = K_{ij} - \lambda\Sigma_k \begin{vmatrix} Kij & Kik \\ Kkj & Kkk \end{vmatrix} + (\lambda^2/2!)\,\Sigma_{km} \begin{vmatrix} Kij & Kik & Kim \\ Kkj & Kkk & Kkm \\ Kmj & Kmk & Kmm \end{vmatrix} - (\lambda^3/3!)\,\Sigma_{kmn} \begin{vmatrix} Kij & Kik & Kim & Kin \\ Kkj & Kkk & Kkm & Kkn \\ Kmj & Kmk & Kmm & Kmn \\ Knj & Knk & Knm & Knn \end{vmatrix} +.$$

$$D_1(x,y;\lambda) == k(x,y) - \lambda\int ds_1 \begin{vmatrix} k(x,y) & k(x,s1) \\ k(s1,y) & k(s1,s1) \end{vmatrix} + (\lambda^2/2!)\int ds_1 ds_2 \begin{vmatrix} k(x,y) & k(x,s1) & k(x,s2) \\ k(s1,y) & k(s1,s1) & k(s1,s2) \\ k(s2,y) & k(s2,s1) & k(s1,s2) \end{vmatrix} -$$

...

3.1 Introduction with Six Examples (191)

Whereas d/dx is unbounded, many useful integral equation operators are bounded. A subset of such bounded operators are the Hilbert-Schmidt (HS) operators where the double integral of the kernel is finite on the square. Such HS operators are in fact compact = completely continuous and map bounded sets into compact sets, though the significance of this fact is not clear at this point.

Stakgold first considers a separable kernel having n terms in the sum with L_2 functions. He shows such a kernel is bounded and compact. Taking the limit $n\to\infty$ he shows that this compact property is maintained, and he shows that in fact any HS integral operator can be written as such an infinite sum. This then shows that any HS operator is compact.

He then mentions the basic kinds of Fredholm Equations: $(L-\mu)u = 0$, homogeneous=EV, and then the inhomo $(L-\mu)u = f$. If $\mu=0$ this latter is called first kind inhomo, otherwise second kind inhomo. The Volterra form with x as an upper integration endpoint can be absorbed into the general fixed-endpoints theory by adjusting the kernel. I directly quote the following note:

Note added 4.22.09. I have now read Chapter 1. If we have $Lu = \mu u$ with some BC's, and if $Lu=0$ has no non-trivial solutions, we know that $Lg=\delta$ has a unique solution g that meets the BC's and then $Lu=f$ with the same BC's has the solution $u = Gf$ where G is the integral operator of kernel g. If L is self-adjoint, then g is symmetric. If we take $f = \mu 1$, then $Lu = \mu u$ becomes $u = \mu Gu$ which is then an integral EV equation for u. The solutions u of this equation respect the BC's. This is why one says that the ODE eigenvalue with BC's can be cast as an integral equation eigenvalue problem which "builds in" the same BC's. Since the integral operator is "nicer" (bounded), maybe best to study the EV problem there instead of as an ODE, at least for certain purposes such as obtaining the general nature of the EV's.

Stak then gives six examples with simple kernels, and studies the eigenvalue situation in each case.

It is probably mentioned later, but here is an important point. If the kernel k is symmetric, then the EF's of the integral EV problem form a complete set! If the kernel k is a Green's Function g of an ODE where L was self-adjoint, then such a kernel is symmetric. Therefore, the EF's of a self-adjoint L form a complete set!

3.2 The Neumann Series Method (206)

We want to solve our inhomo Fred 2 $Ku = \mu u + F$. Divide through by μ to get $(1/\mu)\,Ku = u + F/\mu$. Redefine the last term to be -f, so have $(1/\mu)\,Ku = u - f$. Set $\lambda \equiv 1/\mu$ so have $\lambda Ku = u - f$. This then says that $f = u - \lambda Ku = (1-\lambda K)u$. If $(1-\lambda K)^{-1}$ exists, then we can solve to get $u = (1-\lambda K)^{-1} f$. This type of solution is often used in scattering and is the basis of all of QED, and really all scattering theory, it is pretty major.

We can write a formal solution as

$$u = (1 - \lambda K)^{-1} f = (1 + \lambda K + \lambda^2 K^2 + \ldots.)\, f \qquad 3.25 \qquad k_2(x,z) = \int dz'\, k(x,z')k(z',z) \text{ etc}$$

typical term in series: $u(x) = u_x = \lambda^2 [K^2 f]_x = \lambda^2 \int dz\, [K^2]_{xz} f_z = \lambda^2 \int dz\, k_2(x,z) f(z)$

So, to get a solution using this series, you first have to compute all the objects $[K^n]_{xz} = k_n(x,z)$ which means doing a bunch of integrals only using the kernel k(x,z), and second you have to integrate each of these iterated kernels against the driving function f(z), then add up all the terms. It is a prescription and it gives a definite and unique solution u.

Notice that if λ is an inverse eigenvalue, then $Ku = \mu u \Rightarrow \lambda Ku = u \Rightarrow (1-\lambda K)u = 0$, where u is an eigenfunction. Now suppose the series for $(1-\lambda K)^{-1}$ shown above converges. Then if λ is an inverse eigenvalue, we get $u = (1-\lambda K)^{-1} 0 = 0$. But this is a contradiction because u is supposed to be an eigenfunction. Thus, the series must not converge if λ is an inverse eigenvalue, or λ EV => not converge. Therefore series converges $\Rightarrow$ λ not EV.

The series above *is* the Neumann Series and, when it converges, it is the unique solution. In fact, it always converges when $|\lambda| < 1/\|K\|$, a disc in the λ plane (again, K is bounded we assume). According to the last sentence of the last paragraph, there can be no eigenvalues inside this circle! This translates into saying that $Ku = \mu u$ must have its eigenvalues inside the disc $|\mu| \le \|K\|$. This is consistent with our idea later that EV's are bounded by ||K||. Neumann downside: often not easy to compute the K^n.

Note that the formal operator $(1 - \lambda K)^{-1}$ is called the **resolvent**. (Stak no mention).

Example 1 is a good workingman's example, but Example 2 has a special payoff. If your integral equation has the Volterra form, then the Neumann series converges for $|\lambda| < \infty$, whereas a normal Fredholm case has $|\lambda| < 1/\|K\|$, so the Volterra Form is very happy with a Neumann series solution.

As an application, Stak considers a general initial-value n=2 EV ODE and converts it to a Volterra form integral equation for u"(x). The Neumann series gives a guaranteed solution for u"(x) since Volterra, regardless of value of the EV. Integrate twice meeting BC's and you have u(x). This then proves that the solution to such an ODE exists and its unique. I think this idea could be extended to arbitrary order n, and then this would be a proof of the theorem quoted in Chapter 1 for which we had no proof there.

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

Something missing: Stak never comments on how a H-S operator fits into the general operator theory of Chapter 2. There the spectrum could have 3 types: point, continuous, and that extra thing. I think for self-adjoint, the "extra thing" does not exist. Probably when a and b are both finite (the endpoints), there is no

continuous spectrum. In any event, Stak procedes along as if there were only a discrete spectrum for integral equation operators of the type he considers in this Chapter 3.

If the kernel k(x,y) is symmetric, then the operator K is self-adjoint (as well as symmetric). In this case the following facts are true: (but the last two items also require that K be HS.)

(1) The eigenvalues of K are all real
(2) the EF's of different EV's are orthogonal
(3) the multiplicity of any *non-zero* EV μ is finite.
(4) if number of EV's is ∞, then $\mu=0$ is the only possible limit point.

I think that part of the magic of symmetric K being HS and therefore compact is this idea that only $\mu=0$ can be a limit point of eigenvalues or, if an eigenvalue, only $\mu=0$ can have an infinite multiplicity. ($\mu=0$ means $\lambda = \infty$).

At this point we enter a long saga which concerns upper bounds on eigenvalues. It is trivial to show for a bounded K that $|\mu_i| \le ||K||$ (see raw notes). But since the HS K is compact, we get this extra fact:

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

So compactness of K says we actually *hit* ||K|| with the largest eigenvalue (in abs value). This then is why we are interested in K being H-S, since H-S implies K is compact.

Now, recall that in my notation,

$$|| K ||^2 = \max_u \langle Ku,Ku\rangle/\langle u,u\rangle \quad \Rightarrow \quad || K || = \max_u ||Ku|| / ||u||$$

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

When the dust settles in my raw notes, we learn using Theorem 6 that for a symmetric HS operator,

$$|| K || = ||| K ||| = |\mu_1| \qquad \text{// this is } \textbf{Theorem 1} \text{ on page 215}$$

The rest of this section talks about variational methods to "find" (at least approximately) the eigenvalues and their eigenfunctions.

Theorem 3A: Subject to $\langle u,\varphi_i\rangle=0$ for i=1..n, $\max_u [\langle Ku,u\rangle/\langle u,u\rangle] = |\mu_{n+1}|$ and gives φ_{n+1}.

This suggests an iterative process using an idealized variation method. You first vary u with no conditions to find $|\mu_1|$ and then, in a perfect world, you know μ_1 and φ_1. Then you vary u subject to $\langle u,\varphi_1\rangle = 0$ and this gives you μ_2 and φ_2 , and you just keep going. You need some kind of "perfect varying machine" to make this go, and I suspect the computer now provides such a machine!

If our varying machine is imperfect, we get some value $a_k \le |\mu_k|$ because we did not reach the perfect limit. In this case, the thing we get a_k is a **"lower bound"** for $|\mu_k|$, the thing we are seeking.

We now arrive at this strange intermediate theorem (always talking symmetric and HS)

Theorem 4: If you take the EF's you get from the above procedure but *only include* the EF's from non-zero EV's $\mu \neq 0$, the set of EF's you then have is a complete basis for any function g that can be written g = Kf. In other words, this set of EF's spans the range of K [f of course is L_2 as usual]. So this explains how you could have a "basis" that might only have a finite number of EF's -- if the range of K is finite dimensional. // This is pretty obvious since the zero-eigenvalue is associated with the nullspace.

We then come to the main result:

Theorem 5: The full set of EF's of a symmetric HS kernel is complete on L_2.

Theorem 5A: Obviously, if $\mu = 0$ is not an EV, then the Theorem 4 set of EF's is complete on L_2.

Theorem 6: The eigenfunctions of any self-adjoint ODE operator L form a complete basis on L_2.

This is then the justification for writing things like closure/completeness for, say, Legendre functions. But as noted in the raw and meta notes, Theorem 6 depends on a Green's function kernel being H-S, and Stak has never shown that, and it is not obvious to me. This is a loose bolt never resolved.

3.4 The solution of Ku = μu + f for K symmetric and H-S. (220)

Assume that you have found the EV's μ_i and EF's φ_i of K. You then want to consider the inhomo2 equation $(K-\mu)u = f$. You know the φ_i are a complete set, so expand $u = \Sigma_i a_i \varphi_i$ so $Ku = \Sigma_i a_i K\varphi_i = \Sigma_i a_i \mu_i \varphi_i$. Also expand $f = \Sigma_i b_i \varphi_i$. Then equation becomes $a_i(\mu_i - \mu) = b_i$. Stak then systematically considers the four possible cases : $\mu \neq 0$ and $\mu \neq$ EV; $\mu \neq 0$ and $\mu =$ EV; $\mu = 0$ and $\mu \neq$ EV; $\mu = 0$ and $\mu =$ EV; In any case, if $\mu \neq$ EV, you cannot add homo solutions to your particular solution. As usual, the alternative theorem is lurking in all these cases. So basically we have series solutions on the φ_i for our inhomo2 problem in any of these cases. Convergence is discussed in each case.

3.5 Extremal Principles (223)

Back in Section 3.3 we developed the following two ideas:

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

Theorem 3A: Subject to $\langle u,\varphi_i \rangle = 0$ for i=1..n, $\max_u [\langle Ku,u \rangle / \langle u,u \rangle] = |\mu_{n+1}|$ and gives φ_{n+1}.

In this section, we are going to see what these things look like if we get rid of the absolute values around the eigenvalues. We start by listing the eigenvalues in increasing order, negatives on the left:

μ_1^- μ_2^- ..... (0) ..... μ_2^+ μ_1^+

The above two ideas then translate into the following:

||| K ||| = μ_1^+ // 3.47

Subject to <u,φ^+_i>=0 for i=1..n, $\max_u$ [<Ku,u>/<u,u>] = μ_{n+1}^+ and gives φ^+_{n+1}. ` // 3.48

Subject to <u,φ^-_i>=0 for i=1..n, **min**$_u$ [<Ku,u>/<u,u>] = μ_{n+1}^- and gives φ^-_{n+1}. // 3.50

When you work on the negative side, the max turns into min as you see.

Comment: You can see generally how the above works. If you have a 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.

We then have some definitions for **qualities of K**: +, non-, -, non+, indefinite, real

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**.
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: Since I don't see the significance of this thing, I won't requote it here, see the meta notes.

3.6 Approximation Methods Based on Extremal Principles (226)

The Rayleigh-Ritz Procedure (228)

The goal is to try to solve the EV problem $K\varphi_i = \mu_i\varphi_i$. We don't even know how many eigenvalues there are. But we pick n arbitrary but independent functions and call them v_i and we make trial eigensolutions $u = \Sigma c_i v_i$ with some real coefficients **c**. R(u) is called R*. We then vary the c_i so as to make stationary R(u) which we know should get us close to μ_1. The procedure yields a little nxn matrix equation A**c** = 0 where **c** are our coefficients. There can be a solution if detA = 0, and this is a polynomial of degree n in R*. We then use the largest root R^*_1 as our resulting estimate (or lower bound) for μ_1, knowing $R^*_1 \le \mu_1$, and the resulting **c** as our estimate for $\varphi_1 = \Sigma c_i v_i$. It is also claimed that $R^*_2 \le \mu_2$ and so on, but these

lower solutions are far less accurate. Stak then has several remarks: (1) accuracy of μ_1 will be much better than accuracy of φ_1, due to the stationary idea. (2) accuracy goes down rapidly for lower solutions like μ_2, μ_3 ...; (3) larger n gives more accuracy; (4) good to select v_i based on symmetry. (5) Things are simpler if you use orthonormal v_i, but you don't have to. (6) if you could use n = ∞, you would get exact results since your v_i would be a complete basis, but of course this is not practical.

Stak then gives a single good example of using the R-R procedure on page 231.

In the nuts and bolts of R-R, you first "precompute" the integrals $K_{ik} = \langle Kv_i, v_k \rangle$ and $\alpha_{ik} = \langle v_i, v_k \rangle$. For some complicated K, this can only mean numerical integration, so we are talking computer work here, not analytic stuff. Then you compute the n x n determinant shown in 3.55, and then you have to find the roots of your polynomial. If you pick orthonormals, then $\alpha_{ik} = \delta_{ik}$ and the determinant is the much simpler one shown in 3.57. Matrix K will be symmetric.

Basically, we are making an n x n matrix approximation to our ∞ x ∞ matrix problem.

The Schwarz Iteration Procedure (p 231) // abbreviated quote from meta notes

Again, the goal is to solve the EV problem $K\varphi_i = \mu_i\varphi_i$. Recall that we used iterated kernels K^n in the Neumann series method for solving the inhomo equation written as $u = f + \lambda Ku$ and we found that the series was $u = (1 - \lambda K)^{-1}f = (1 + \lambda K + \lambda^2K^2 + ...) f$. In the Schwarz method, we again make use of these iterated kernels K^n, but they are used differently and of course for the homo equation. The idea is to start off with some seed function f_0 and then compute $f_n = K^nf_0$. If you are lucky, this sequence will become stable after a while and you get $f_{n+1} \approx \alpha f_n$ which says $f_{n+1} = Kf_n \approx \alpha f_n$ and thus this stabilized limiting function f_n is *an* eigenfunction (or an approximation for one), eigenvalue α. That is the first idea.

The second part of Schwarz is that you define $a_n \equiv \langle f_o, K^nf_0 \rangle$ and consider the sequence for large n. You define the ratio $\theta_n = a_{n+1}/a_n$ (Schwarz quotient) . Basically the claim is that lim $\theta_n \leq \mu_1$.

***Upper* Bounds to Eigenvalues**

Our usual variation methods like $R(u) = \langle Ku,u \rangle / \langle u,u \rangle$ always give lower bounds for eigenvalues, it would be nice to have some *upper* bounds so you could bracket (enclose) the eigenvalue to know how accurate you might be. The ideas here are based on the notion that if you define

$$\sigma \equiv \langle Kv,v \rangle$$

$$\alpha^2 \equiv \| Kv - \sigma v \|^2 = \|Kv\|^2 - \sigma^2$$

if v is an exact normalized eigenfunction, then α = 0 and σ = μ. If v is only an approximate normalized eigenvalue, then you find by the WBFS Theorem (four names, no web hits so no date) that your eigenvalue is bracketed in this way, reminding me of a mean and standard deviation idea,

$$\sigma - \alpha \leq \mu \leq \sigma + \alpha$$

3.7 Continuity and Uniform Convergence: bilinear series and iterated kernels

Uniform convergence is more demanding that convergence in the mean, and this section shows that certain partial sums of functions of x and ξ converge uniformly "on the square" to the infinite sum limits.

Preamble: Consider self-adjoint ODE Lu=f with no nullspace, so solution is u = Gf. Note that u is "in the range of integral operator G" and therefore by a theorem above we know we can expand u on the complete set $\varphi_i(x)$, since G is symmetric. So can write $u(x) = \Sigma_i a_i \varphi_i(x)$. Similarly, can write $f(x,\xi) = \Sigma_i C_i \varphi_i(x)\bar{\varphi}_i(\xi)$ which is called a bilinear expansion of a function of two variables. Stak then shows:

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

Mercer's Theorem: For suitable extra conditions on K, the above is also true for n = 1.

If we assume these "extra conditions", then we obtain this somewhat impressive result:

$g(x,\xi) = \Sigma_i^{\infty} \mu_i \varphi_i(x)\bar{\varphi}_i(\xi)$ or

$RHS = \Sigma_i^{\infty} \langle x|i\rangle \mu_i \langle i|\xi\rangle = \Sigma_i^{\infty} \langle x|\, \mu_i |\, i\rangle \langle i|\xi\rangle = \Sigma_i^{\infty} \langle x|\, G \,|\, i\rangle \langle i|\xi\rangle = \langle x|G|\xi\rangle = g(x,\xi)$

which gives the Green's function of Lg=δ in terms of the φ_i of $G\varphi_i = \mu_i\varphi_i$. This result did not appear anywhere in Chapter 1 on "Green's Functions". Page 72 (1.77) looks a little like it, but not really because there are no eigenfunctions involved in (1.77).

Theorem 4 (or Mercer's) says that the sum shown equals the LHS, and the *partial* sum to some N converges as $N\to\infty$ to the LHS *uniformly* on the square.

If you use "in the mean" convergence, Theorem 4 applies for $n \geq 1$.

Just integrate Theorem 4 over x and use orthog of φ_i to get

Theorem 4A: $\int dx\, k_n(x,x) = \Sigma_i \mu_i^n$ // an interesting Parseval like sum rule $n \geq 2$.

which follows from Theorem 4 since $k_n(x,x) = \Sigma_i \mu_i^n \varphi_i(x)\, \bar{\varphi}_i(x)$ and the φ_i are normalized. In our QM notation we can write this as

$RHS = \Sigma_i^{\infty} \langle i\,|\mu_i^n|i\rangle = \Sigma_i^{\infty} \langle i\,|K^n|i\rangle = \text{trace}(K^n) = \int dx \langle x|K^n|x\rangle$

with the idea that trace is independent of basis.

Theorem 4B: $[\int dx\, k_n(x,x)]^{1/n} \geq \mu_1$ if K is non-negative, an *upper* bound for μ_1 $n \geq 2$

But is obvious from 4A since 4A says: $\int dx\, k_n(x,x) = \mu_1^n + \Sigma_{i\neq 1} \mu_i^n \geq \mu_1^n$.

Note: Lg=δ goes with Lu=f which goes with u = Gf which we usually write as Gu = f as in 3.4 p 195, and this is the inhomo1 problem where parameter μ = 0. We can write our n=1 case above as: ($\mu_i = 1/\lambda_i$)

$g(x,\xi) = \Sigma_i^{\infty} \varphi_i(x)\bar{\varphi}_i(\xi)/\lambda_i$

and this Green's Function is associated with the parameter μ=0 (λ=∞) situation. Later in Chap 4 we shall learn the more general result for the inhomo2 problem with parameter μ = 1/λ which is this:

$$g(x,\xi;\lambda) = \Sigma_i{}^{\infty}\varphi_i(x)\overline{\varphi}_i(\xi)/(\lambda_i-\lambda) \qquad \text{// p 273 (4.42)}$$

where *this* g is defined by (L - λ)g = δ as in 4.1 page 259.

3.8 Approximation Methods for Solving Integral Equations

For gory details see **Kantorovich and Krylov** 1958 English translation. In this section we discuss 5 different plans for finding approximate solutions to our inhomo2 integral equation.

1. Approximate area as little strips, then do lattice integral (Trap and Simpson versions allowed).
2. Approximate the kernel k as a finite separable sum of independent functions.
3. For some $\{v_i\}$ approx the solution $u(x) = \Sigma_i c_i v_i(x)$, install in the IE, then do one of two things:
   (a) minimize the L_2 error of the two sides of the IE (Best in the Mean);
   (b) require that both sides have the same projection on some finite subspace spanned by some w_i (Galerkin's Method).
4. Two variational methods are presented, theory only: maximize either I(u) or J(u).
5. Carry out the maximize I(u) plan, get the Rayleigh-Ritz Equations, same as a Galerkin case.

All of these methods reduce the problem of solving Ku=f to an nxn linear algebra problem. See meta notes for a little summary of each of the above items.

Exercise 3.23. Discusses the **extended (generalized) eigenvalue problem** Nu = λMu.

3.9 Non-symmetric HS operators (250)

For non-symmetric K, it turns out that you can still do pretty much in terms of solving an integral equation of the form Ky = h, but you cannot make use of the φ_i EF's of K because these are not in general a complete set, as a very simple example shows. *However*, the EF's v_n and u_n of the left and right *iterated* kernels K_L = KK* and K_R= K*K can be used instead! Operators $K_{L,R}$ are both HS (if K is), are symmetric, have the same eigenvalues, and are non-negative, so they have all the properties we like! For example, either set u_n or v_n is complete. The upshot is that you can solve Ky = h as long as h meets a few restrictions.

Along the way, we learn that the v_n span the perp space of K, while the u_n the perp space of K*. These facts lead to a final solution in the usual form we expect such as $y = y_0 + \Sigma_i [(h.u_i)/\mu_i] v_i$. Notice that both the u_i and v_i functions are involved in this formula which solves Ky=h.

Everything studied in this section is illustrated in the simple Example which ends the section.