Math 212b Lecture 6

Rayleigh-Ritz

The discrete spectrum and the continuous spectrum.

Let H be a self-adjoint operator on a Hilbert space $\mathcal{H}$ and let $\sigma = \sigma(H) \subset \mathbb{R}$ denote is spectrum.

The **discrete spectrum** of H is defined to be those eigenvalues λ of H which are of finite multiplicity and are also isolated points of the spectrum. This latter condition says that there is some $\epsilon > 0$ such that the intersection of the interval $(\lambda - \epsilon, \lambda + \epsilon)$ with σ consists of the single point $\{\lambda\}$. The discrete spectrum of H will be denoted by $\sigma_d(H)$ or simply by σ_d when H is fixed in the discussion.

The complement in $\sigma_d(H)$ in $\sigma(H)$ is called the **essential spectrum** of H and is denoted by $\sigma_{\text{ess}}(H)$ or simply by σ_{ess} when H is fixed in the discussion.

Characterizing the discrete spectrum.

If $\lambda \in \sigma_d(H)$ then for sufficiently small $\epsilon > 0$ the spectral projection $P = P((\lambda - \epsilon, \lambda + \epsilon))$ has the property that it is invariant under H and the restriction of H to the image of P has only λ in its spectrum and hence $P(\mathcal{H})$ is finite dimensional, since the multiplicity of λ is finite by assumption.

Conversely, suppose that $\lambda \in \sigma(H)$ and that $P(\lambda - \epsilon, \lambda + \epsilon)$ is finite dimensional. This means that in the spectral representation of H, the subset $E_{(\lambda-\epsilon,\lambda+\epsilon)}$ of

$$S = \sigma \times \mathbb{N}$$

consisting of all

$$(s, n)|\lambda - \epsilon < s < l\lambda + \epsilon\}$$

has the property that

$$L_2(E_{(\lambda-\epsilon,\lambda+\epsilon)}, d\mu)$$

is finite dimensional. If we write

$$E_{(\lambda-\epsilon,\lambda+\epsilon)} = \bigcup_{n\in\mathbb{N}} (\lambda - \epsilon, \lambda + \epsilon) \times \{n\}$$

$$L_2(E_{(\lambda-\epsilon,\lambda+\epsilon)}, d\mu)$$

is finite dimensional. If we write

$$E_{(\lambda-\epsilon,\lambda+\epsilon)} = \bigcup_{n\in\mathbb{N}} (\lambda-\epsilon, \lambda+\epsilon) \times \{n\}$$

then since

$$L_2(E_{(\lambda-\epsilon,\lambda+\epsilon)}) = \hat{\bigoplus_n} L_2((\lambda-\epsilon, \lambda+\epsilon), \{n\})$$

we conclude that all but finitely many of the summands on the right are zero, which implies that for all but finitely many n we have

$$\mu\left((\lambda-\epsilon, \lambda+\epsilon) \times \{n\}\right) = 0.$$

For each of the finite non-zero summands, we can apply the case $N = 1$ of the following lemma:

Lemma 1 *Let ν be a measure on $\mathbb{R}^N$ such that $L_2(\mathbb{R}^N, \nu)$ is finite dimensional. Then ν is supported on a finite set in the sense that there is some finite set of m distinct points $x_1, \ldots, x_m$ each of positive measure and such that the complement of the union of these points has ν measure zero.*

Proof. Partition $\mathbb{R}^N$ into cubes whose vertices have all coordinates of the form $t/2^r$ for an integer r and so that this is a disjoint union. The corresponding decomposition of the L_2 spaces shows that only finite many of these cubes have positive measure, and as we increase r the cubes with positive measure are nested downward, and can not increase in number beyond $n = \dim L_2(\mathbb{R}^N, \nu)$. Hence they converge in measure to at most n distinct points each of positive ν measure and the complement of their union has measure zero. $\square$

We conclude from this lemma that there are at most finitely many points (s_r, k) with $s_r \in (\lambda - \epsilon, \lambda + \epsilon)$ which have finite measure in the spectral representation of H, each giving rise to an eigenvector of H with eigenvalue s_r, and the complement of these points has measure zero. This shows that $\lambda \in \sigma_d(H)$. We have proved

Proposition 1 *$\lambda \in \sigma(H)$ belongs to $\sigma_d(H)$ if and only if there is some $\epsilon > 0$ such that $P((\lambda - \epsilon, \lambda + \epsilon))(\mathcal{H})$ is finite dimensional.*

Characterizing the essential spectrum

This is simply the contrapositive of Prop. 1:

Proposition 2 *$\lambda \in \sigma(H)$ belongs to $\sigma_{\text{ess}}(H)$ if and only if for every $\epsilon > 0$ the space*

$$P((\lambda - \epsilon, \lambda + \epsilon))(\mathcal{H})$$

is infinite dimensional.

Operators with empty essential spectrum.

Theorem 1 *The essential spectrum of a self adjoint operator is empty if and only if there is a complete set of eigenvectors of H such that the corresponding eigenvalues λ_n have the property that $|\lambda_n| \to \infty$ as $n \to \infty$.*

Proof. If the essential spectrum is empty, then the spectrum consists of eigenvalues of finite multiplicity which have no accumulation finite point, and so must converge in absolute value to ∞. Enumerate the eigenvalues according to increasing absolute value. Each has finite multiplicity and so we can find an orthonormal basis of the finite dimensional eigenspace corresponding to each eigenvalue. The eigenvectors corresponding to distinct eigenvalues are orthogonal. So what we must show is that the space spanned by all these eigenvectors is dense. Suppose not. The space L orthogonal to all the eigenvectors is invariant under H. If this space is non-zero, the spectrum of H restricted to this subspace is not empty, and is a subset of the spectrum of H. So there will be eigenvectors in L, contrary to the definition of L.

Conversely, suppose that the conditions hold. Let f_n be the complete set of eigenvectors. Since the set of eigenvalues is isolated, we will be done if we show that they constitute the entire spectrum of H. Suppose that z does not coincide with any of the λ_n. We must show that the operator $zI - H$ has a bounded inverse on the domain of H, which consists of all $f = \sum_n a_n f_n$ such that $\sum |a_n|^2 < \infty$ and $\sum \lambda_n^2 |a_n|^2 < \infty$. But for these f

$$\|(zI - H)f\|^2 = \sum_n |z - \lambda_n|^2 |a_n|^2 \geq c^2 \|f\|^2$$

where $c = \min_n |\lambda_n - z| > 0$. $\square$

Non-negative operators with empty essential spectrum.

Corollary 1 *Let H be a non-negative self adjoint operator on a Hilbert space $\mathcal{H}$. The following conditions on H are equivalent:*

1. *The essential spectrum of H is empty.*
2. *There exists an orthonormal basis of $\mathcal{H}$ consisting of eigenvectors f_n of H, each with finite multiplicity with eigenvalues $\lambda_n \to \infty$.*
3. *The operator $(I + H)^{-1}$ is compact.*

Since there are no negative eigenvalues, we know that 1) and 2) are equivalent. We must show that 2) and 3) are equivalent. If $(I+H)^{-1}$ is compact, we know that there is an orthonormal basis $\{f_n\}$ of $\mathcal{H}$ consisting of eigenvectors with eigenvalues $\mu_n \to 0$. (We know from the spectral theorem that $(I+H)^{-1}$ is unitarily equivalent to multiplication by the positive function $1/(1+h)$ and so $(I+H)^{-1}$ has no kernel.) Then the $\{f_n\}$ constitute an orthonormal basis of $\mathcal{H}$ consisting of eigenvectors with eigenvalues $\lambda_n = \mu_n^{-1} \to \infty$.

Conversely, suppose that 2) holds. Consider the finite rank operators A_n defined by

$$A_n f := \sum_{j=1}^{n} \frac{1}{1+\lambda_j}(f, f_j) f_j.$$

Then

$$\|(I+H)^{-1} f - A_n f\| = \| \sum_{j=n+1}^{\infty} \frac{1}{1+\lambda_j}(f, f_j) f_j)\| \leq \frac{1}{1+\lambda_n}\|f\|.$$

This shows that n$(I+H)^{-1}$ can be approximated in operator norm by finite rank operators.

The variational method.

Let H be a non-negative self-adjoint operator on a Hilbert space $\mathcal{H}$. For any finite dimensional subspace L of H with $L \subset \mathcal{D} = \mathrm{Dom}(H)$ define

$$\lambda(L) := \sup\{(Hf, f) | f \in L \text{ and } \|f\| = 1\}.$$

Define

$$\lambda_n = \inf\{\lambda(L), \quad | L \subset \mathcal{D}, \text{ and } \dim L = n\}. \qquad (1)$$

The λ_n are an increasing family of numbers. We shall show that they constitute that part of the discrete spectrum of H which lies below the essential spectrum:

Theorem 3 *Let H be a non-negative self-adjoint operator on a Hilbert space $\mathcal{H}$. Define the numbers $\lambda_n = \lambda_n(H)$ by (1). Then one of the following three alternatives holds:*

1. *H has empty essential spectrum. In this case the $\lambda_n \to \infty$ and coincide with the eigenvalues of H repeated according to multiplicity and listed in increasing order, or else $\mathcal{H}$ is finite dimensional and the λ_n coincide with the eigenvalues of H repeated according to multiplicity and listed in increasing or-*

2. *There exists an $a < \infty$ such that $\lambda_n < a$ for all n, and $\lim_{n\to\infty} \lambda_n = a$. In this case a is the smallest number in the essential spectrum of H and $\sigma(H) \cap [0, a)$ consists of the λ_n which are eigenvalues of H repeated according to multiplicity and listed in increasing order.*

3. *There exists an* $a < \infty$ *and an* N *such that* $\lambda_n < a$ *for* $n \leq N$ *and* $\lambda_m = a$ *for all* $m > N$. *Then* a *is the smallest number in the essential spectrum of* H *and* $\sigma(H) \cap [0, a)$ *consists of the* $\lambda_1, \ldots, \lambda_N$ *which are eigenvalues of* H *repeated according to multiplicity and listed in increasing order.*

Proof. Let b be the smallest point in the essential spectrum of H (so $b = \infty$ in case 1.). So H has only isolated eigenvalues of finite multiplicity in $[0, b)$ and these constitute the entire spectrum of H in this interval. Let $\{f_k\}$ be an orthonormal set of these eigenvectors corresponding to these eigenvalues μ_k listed (with multiplicity) in increasing order.

Let M_n denote the space spanned by the first n of these eigenvectors, and let $f \in M_n$. Then $f = \sum_{j=1}^{n} (f, f_j) f_j$ so

$$Hf = \sum_{j=1}^{n} \mu_j (f, f_j) f_j$$

and so

$$(Hf, f) = \sum_{j=1}^{n} \mu_j |(f, f_j)|^2 \leq \mu_n \sum_{j=1}^{n} |(f, f_j)|^2 = \mu_n \|f\|^2$$

$$(Hf, f) = \sum_{j=1}^{n} \mu_j |(f, f_j)|^2 \leq \mu_n \sum_{j=1}^{n} |(f, f_j)|^2 = \mu_n \|f\|^2$$

so

$$\lambda_n \leq \mu_n.$$

In the other direction, let L be an n-dimensional subspace of $\mathrm{Dom}(H)$ and let P denote orthogonal projection of $\mathcal{H}$ onto M_{n-1} so that

$$Pf = \sum_{j=1}^{n-1} (f, f_j) f_j.$$

$$Pf = \sum_{j=1}^{n-1} (f, f_j) f_j.$$

The image of P restricted to L has dimension $n-1$ while L has dimension n. So there must be some $f \in L$ with $Pf = 0$. By the spectral theorem, the function $\tilde{f} = Uf$ corresponding to f is supported in the set where $h \geq \mu_n$ and hence $(Hf, f) \geq \mu_n \|f\|^2$ so

$$\lambda_n \geq \mu_n.$$

There are now three cases to consider: If $b = +\infty$ (i.e. the essential spectrum of H is empty) the $\lambda_n = \mu_n$ can have no finite accumulation point so we are in case 1). If there are infinitely many μ_n in $[0, b)$ they must have finite accumulation point $a \leq b$, and by definition, a is in the essential spectrum. Then we must have $a = b$ and we are in case 2). The remaining possibility is that there are only finitely many $\mu_1, \ldots \mu_M < b$. Then for $k \leq M$ we have $\lambda_k = \mu_k$ as above, and also $\lambda_m \geq b$ for $m > M$. Since $b \in \sigma_{\text{ess}}(H)$, the space

$$K := P(b - \epsilon, b + \epsilon)\mathcal{H}$$

is infinite dimensional for all $\epsilon > 0$. Let $\{f_1, f_2, \ldots, \}$ be an orthonormal basis of K, and let L be the space spanned by the first m of these basis elements. By the spectral theorem, $(Hf, f) \leq (b + \epsilon)\|f\|^2$ for any $f \in L$. so for all m we have $\lambda_m \leq b + \epsilon$. So we are in case 3). $\square$

In applications (say to chemistry) on deals with self-adjoint operators which are bounded from below, rather than being non-negative. But this requires just a trivial shift in stating and applying the preceding theorem. In some of these applications the bottom of the essential spectrum is at 0, and one is interested in the lowest eigenvalue λ_1 which is negative.

Variations on the variational formula.

An alternative formulation of the formula.

Instead of (1) we can determine the λ_n as follows: We define λ_1 as before:

$$\lambda_1 = \min_{f \neq 0} \frac{(Hf, f)}{(f, f)}.$$

Suppose that f_1 is an f which attains this minimum. We then know that f_1 is an eigenvector of H with eigenvalue λ_1. Now define

$$\lambda_2 := \min_{f \neq 0, f \perp f_1} \frac{(Hf, f)}{(f, f)}.$$

$$\lambda_2 := \min_{f \neq 0, f \perp f_1} \frac{(Hf, f)}{(f, f)}.$$

This λ_2 coincides with the λ_2 given by (1) and an f_2 which achieves the minimum is an eigenvector of H with eigenvalue λ_2. Proceeding this way, after finding the first n eigenvalues $\lambda_1, \ldots, \lambda_n$ and corresponding eigenvectors $f_1, \ldots, f_n$ we define

$$\lambda_{n+1} = \min_{f \neq 0, f \perp f_1, f \perp f_2, \ldots, f \perp f_n} \frac{(Hf, f)}{(f, f)}.$$

This gives the same λ_k as (1).

Variations on the condition $L \subset \mathrm{Dom}(H)$.

In some applications, the condition $L \subset \mathrm{Dom}(H)$ is unduly restrictive, especially when we want to compare eigenvalues of different self adjoint operators. In these applications, one can frequently find a common **core** $\mathcal{D}$ for the quadratic forms Q associated to the operators. That is,

$$\mathcal{D} \subset \mathrm{Dom}(H^{\frac{1}{2}})$$

and $\mathcal{D}$ is dense in $\mathrm{Dom}(H^{\frac{1}{2}})$ for the metric $\|\cdot\|_1$ given by

$$\|\|f\|_1^2 = Q(f,f) + \|f\|^2$$

where

$$Q(f,f) = (Hf, f).$$

Theorem 4 *Define*

$$\begin{aligned}
\lambda_n &= \inf\{\lambda(L), \quad |L \subset \mathrm{Dom}(H) \\
\lambda'_n &= \inf\{\lambda(L), \quad |L \subset \mathcal{D} \\
\lambda''_n &= \inf\{\lambda(L), \quad |L \subset \mathrm{Dom}(H^{\frac{1}{2}}).
\end{aligned}$$

Then

$$\lambda_n = \lambda'_n = \lambda''_n.$$

I will leave the proof of this important but technical theorem to the notes.

The secular equation.

The definition (1) makes sense in a real finite dimensional vector space. If Q is a real quadratic form on a finite dimensional real Hilbert space V, then we can write $Q(f) = (Hf, f)$ where H is a self-adjoint (=symmetric) operator, and then find an orthonormal basis according to (1). In terms of such a basis $f_1, \ldots, f_n$, we have

$$Q(f) = \sum_k \lambda_k r_k^2 \quad \text{where} \quad f = \sum r_k f_k.$$

If we consider the problem of finding an extreme point of $Q(f)$ subject to the constraint that $(f, f) = 1$, this becomes (by Lagrange multipliers), the problem of finding λ and f such that

$$dQ_f = \lambda dS_f, \text{where} \quad S(f) = (f, f).$$

$$dQ_f = \lambda dS_f, \text{where} \;\; S(f) = (f, f).$$

In terms of the coordinates $(r_1, \ldots, r_n)$ we have

$$\frac{1}{2} dQ_f = (\mu_1 r_1, \ldots, \mu_n r_n) \;\; \text{while} \;\; \frac{1}{2} dS_f = (r_1, \ldots, r_n).$$

So the only possible values of λ are $\lambda = \mu_i$ for some i and the corresponding f is given by $r_j = 0, \; j \neq i$ and $r_i \neq 0$. This is a watered down version version of Theorem 3. In applications, one is frequently given a basis of V which is *not* orthonormal. Thus (in terms of the given basis)

$$Q(f) = \sum H_{ij} r_i r_j, \quad \text{and} \quad S(f) = \sum_{ij} S_{ij} r_i r_j$$

where

$$f = \sum r_i f_i.$$

$$dQ_f = \lambda dS_f$$

i.e.

$$\begin{pmatrix} H_{11} - \lambda S_{11} & H_{12} - \lambda S_{12} & \cdots & H_{1n} - \lambda S_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ H_{n1} - \lambda S_{n1} & H_{n2} - \lambda S_{n2} & \cdots & H_{nn} - \lambda S_{nn} \end{pmatrix} \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix} = 0$$

As a condition on λ this becomes the algebraic equation

$$\det \begin{pmatrix} H_{11} - \lambda S_{11} & H_{12} - \lambda S_{12} & \cdots & H_{1n} - \lambda S_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ H_{n1} - \lambda S_{n1} & H_{n2} - \lambda S_{n2} & \cdots & H_{nn} - \lambda S_{nn} \end{pmatrix} = 0$$

which is known as the secular equation due to its previous use in astronomy to determine the periods of orbits.

The Dirichlet problem for bounded domains.

Let Ω be an open subset of $\mathbb{R}^n$. The Sobolev space $W^{1,2}(\Omega)$ is defined as the set of all $f \in L_2(\Omega, dx)$ (where dx is Lebesgue measure) such that all first order partial derivatives $\partial_i f$ in the sense of generalized functions belong to $L_2(\Omega, dx)$. On this space we have the Sobolev scalar product

$$(f, g)_1 := \int_\omega (f(x)\overline{g}(x) + \nabla f(x) \cdot \nabla \overline{g}(x))dx.$$

It is not hard to check (and we will do so within the next three lectures) that $W^{1,2}(\Omega)$ with this scalar product is a Hilbert space. We let $C_0^\infty(\Omega)$ denote the space of smooth functions of compact support whose support is contained in Ω, and let $W_0^{1,2}(\Omega)$ denote the completion of $C_0^\infty(\Omega)$ with respect to the norm $\| \cdot \|_1$ coming from the scalar product $(\cdot, \cdot)_1$.

We will show that Δ defines a non-negative self-adjoint operator with domain $W_0^{1,2}(\Omega)$ known as the Dirichlet operator associated with Ω. I want to postpone the proofs of these general facts and concentrate on what Rayleigh-Ritz tells us when Ω is a bounded open subset which we will assume from now on.

We are going apply Rayleigh-Ritz to the domain $\mathcal{D}(\Omega)$ and the quadratic form $Q(f) = Q(f, f)$ where

$$Q(f, g) := \int_\Omega \nabla f(x) \cdot \nabla \overline{g}(x) dx.$$

Define

$$\lambda_n(\Omega) := \inf\{\lambda(L) | L \subset C_0^\infty(\Omega), \quad \dim(L) = n\}$$

where

$$\lambda(L) = \sup Q(f), \ f \in L, \quad \|f\| = 1$$

as before.

$$\lambda_n(\Omega) := \inf\{\lambda(L) | L \subset C_0^\infty(\Omega), \quad \dim(L) = n\}$$

where

$$\lambda(L) = \sup Q(f), \ f \in L, \quad \|f\| = 1$$

as before.

Here is the crucial observation: If $\Omega \subset \Omega'$ are two bounded open regions then

$$\lambda_n(\Omega) \geq \lambda_n(\Omega')$$

since the infimum for Ω is taken over a smaller collection of subspaces than for Ω'.

Suppose that Ω is an interval $(0, a)$ on the real line. Fourier series tells us that the functions $f_k = \sin(\pi k x/a)$ form a basis of $L_2(\Omega, dx)$ and are eigenvectors of Δ with eigenvalues k^2a^2. By Fubini we get a corresponding formula for any cube in $\mathbb{R}^n$ which shows that $(I + \Delta)^{-1}$ is a compact operator for the case of a cube. Since any Ω contains a cube and is contained in a cube, we conclude that the $\lambda_n(\Omega)$ tend to ∞ and so $H_o = \Delta$ (with the Dirichlet boundary conditions) have empty essential spectra and $(I + H_0)^{-1}$ are compact.

Furthermore the $\lambda_n(\Omega)$ aer the eigenvalues of H_0 arranged in increasing order.

Continuity of the eigenvalues from the inside.

Proposition 3 *If Ω_m is an increasing sequence of open sets contained in Ω with*

$$\Omega = \bigcup_m \Omega_m$$

then

$$\lim_{m \to \infty} \lambda_n(\Omega_m) = \lambda_n(\Omega)$$

for all n.

Proof. For any $\epsilon > 0$ there exists an n-dimensional subspace L of $C_0^\infty(\Omega)$ such that $\lambda(L) \leq \lambda_n(\Omega) + \epsilon$. There will be a compact subset $K \subset \Omega$ such that all the elements of L have support in K. We can then choose m sufficiently large so that $K \subset \Omega_m$. Then

$$\lambda_n(\Omega) \leq \lambda_n(\Omega_m) \leq \lambda_n(\Omega) + \epsilon. \quad \square$$

Approximation methods.

The minimum eigenvalue λ_1 is determined according to (1) by

$$\lambda_1 = \inf_{\psi \neq 0} \frac{(H\psi, \psi)}{(\psi, \psi)}.$$

Unless one has a clever way of computing λ_1 by some other means, minimizing the expression on the right over all of $\mathcal{H}$ is a hopeless task. What is done in practice is to choose a finite dimensional subspace and apply the above minimization over all ψ in that subspace (and similarly to apply (1) to subspaces of that subspace for the higher eigenvalues). The hope is that this yield good approximations to the true eigenvalues.

If M is a finite dimensional subspace of $\mathcal{H}$, and P denotes projection onto M, then applying (1) to subspaces of M amounts to finding the eigenvalues of PHP, which is an algebraic problem as we have seen. A **chemical theory** (when H is the Schrödinger operator) then amounts to cleverly choosing such a subspace.

Consider the case where M is two dimensional with a basis ψ_1 and ψ_2. The idea is that we have some grounds for believing that that the true eigenfunction has characteristics typical of these two elements and is likely to be some linear combination of them. If we set

$$H_{11} := (H\psi_1, \psi_1), \quad H_{12} := (H\psi_1, \psi_2) = \overline{H_{21}},$$
$$H_{22} := (H\psi_2, \psi_2)$$

and

$$S_{11} := (S\psi_1, \psi_1), \quad S_{12} := (\psi_1, \psi_2) = \overline{S_{21}}, \quad S_{22} := (\psi_2, \psi_2)$$

then if these quantities are real we can apply the secular equation

$$\det \begin{pmatrix} H_{11} - \lambda S_{11} & H_{12} - \lambda S_{12} \\ H_{21} - \lambda S_{21} & H_{22} - \lambda S_{22} \end{pmatrix} = 0$$

to determine λ.

Suppose that $S_{11} = S_{22} = 1$, i.e. that ψ_1 and ψ_2 are separately normalized. Also assume that ψ_1 and ψ_2 are linearly independent. Let

$$\beta := S_{12} = S_{21}.$$

This β is sometimes called the "overlap integral" since if our Hilbert space is $L_2(\mathbb{R}^3)$ then $\beta = \int_{\mathbb{R}^3} \psi_1 \overline{\psi}_2 dx$. Now

$$H_{11} = (H\psi_1, \psi_1)$$

is the guess that we would make for the lowest eigenvalue (= the lowest "energy level") if we took L to be the one dimensional space spanned by ψ_1. So let us call this value E_1. So $E_1 := H_{11}$ and similarly define $E_2 = H_{22}$. The secular equation becomes

$$(\lambda - E_1)(\lambda - E_2) - (H_{12} - \lambda\beta)^2 = 0.$$

$$(\lambda - E_1)(\lambda - E_2) - (H_{12} - \lambda\beta)^2 = 0.$$

If we define $F(\lambda) := (\lambda - E_1)(\lambda - E_2) - (H_{12} - \lambda\beta)^2$ then F is positive for large values of $|\lambda|$ since $|\beta| < 1$ by Cauchy-Schwarz. $F(\lambda)$ is non-positive at $\lambda = E_1$ or E_2 and in fact generically will be strictly negative at these points. So the lower solution of the secular equations will generically lie strictly below $\min(E_1, E_2)$ and the upper solution will generically lie strictly above $\max(E_1, E_2)$. This is known as the **no crossing rule** and is of great importance in chemistry. I hope to explain the higher dimensional version of this rule (due to Teller-von Neumann and Wigner) later.

Hückel theory of hydrocarbons.

In this theory the space M is the n-dimensional space where each carbon atom contributes one electron. (The other electrons being occupied with the hydrogen atoms.) It is assumed that the S in the secular equation is the identity matrix. This amounts to the assumption that the basis given by the electrons associated with each carbon atom is an orthonormal basis. It is also assumed that $(Hf, f) = \alpha$ is the same for each basis element. In a crude sense this measures the electron-attracting power of each carbon atom and hence is assumed to be the same for all basis elements. If $(Hf_r, f_s) \neq 0$, the atoms r and s are said to be "bonded". It is assumed that only "nearest neighbor" atoms are bonded, in which case it is assumed that $(Hf_r, f_s) = \beta$ is independent of r and s.

So PHP

has the form

$$\alpha I + \beta A$$

where A is the adjacency matrix of the graph whose vertices correspond to the carbon atoms and whose edges correspond to the bonded pairs of atoms. If we set

$$x := \frac{E - \alpha}{\beta}$$

then finding the energy levels is the same as finding the eigenvalues x of the adjacency matrix A. In particular this is so if we assume that the values of α and β are independent of the particular molecule.