Lecture 10 One says that the projection P on M_P and along N_P. Note that $Q = 1 - P$ is also a projection and

$$N_P = M_Q \quad \text{and} \quad N_Q = M_P .$$

We now prove an interesting and useful lemma.

Lemma 135.1

Suppose P_1 and P_2 are two projections in a vector space V with $\dim V = n$, and suppose that $\|P_1 - P_2\| < 1$. Then

$$\dim(\operatorname{Ran} P_1) = \dim(\operatorname{Ran} P_2)$$

Proof: Suppose $n_1 = \dim(\operatorname{Ran} P_1) > n_2 = \dim(\operatorname{Ran} P_2)$. We will obtain a contradiction. Let $u_1, \dots, u_{n_1}$ be a basis for $\operatorname{Ran} P_1$, $v_1, \dots, v_{n_2}$ a basis for $\operatorname{Ran} P_2$, and $v_{n_2+1}, \dots, v_n$ be a basis for $\operatorname{Nul}(P_2)$. Then $v_1, \dots, v_{n_2}, v_{n_2+1}, \dots, v_n$ is a basis for V. This is because $\operatorname{Ran} P_2 \oplus \operatorname{Nul} P_2 = V$ is a direct decomposition of V. Then for each u_i, $1 \le i \le n_1$, $\exists$

$\{a_{ji}\}_{j=1}^{n}$, such that $u_i = \sum_{j=1}^{n_2} a_{ji} v_j + \sum_{j=n_2+1}^{n} a_{ji} v_j$.

Now let $(\gamma_1, \ldots, \gamma_{n_1})^T$ be a non-zero solution of $\sum_{i=1}^{n_1} a_{ji} \gamma_i = 0$, $1 \le j \le n_2$. As $n_2 < n_1$, such $\{\gamma_i\}_1^{n_1}$ $\exists$. Set $u = \sum_{i=1}^{n_1} u_i \gamma_i$. Then as $(\gamma_1, \ldots, \gamma_{n_1})^T \neq 0$ and the u_i's are independent, $u \neq 0$, and we have

$$u = \sum_{i=1}^{n_1} \gamma_i \left(\sum_{j=1}^{n_2} a_{ji} v_j + \sum_{j=n_2+1}^{n} a_{ji} v_j \right)$$

$$= \sum_{j=1}^{n_2} \left(\sum_{i=1}^{n_1} a_{ji} \gamma_i \right) v_j + \sum_{j=n_2+1}^{n} \left(\sum_{i=1}^{n_1} a_{ji} \gamma_i \right) v_j$$

$$= 0 + \sum_{j=n_2+1}^{n} \left(\sum_{i=1}^{n_1} a_{ji} \gamma_i \right) v_j .$$

We conclude that $u \in \text{Nul}(P_2) \cap \text{Ran}(P_1)$.

But then $\|u\| = \|P_1 u\| = \|(P_1 - P_2) u\| \le \|P_1 - P_2\| \, \|u\|$

$< \|u\|$

and so $u = 0$, which is a contradiction. We are done!

<u>Remark:</u> Another very interesting proof of this lemma is in T. Kato, I §4.6, pp 32-34.

<u>Corollary 136.1</u> Suppose $t \mapsto P(t)$ is a continuous map from $\alpha \le t \le \beta$ into the projections, $P(t) = P^2(t)$, $\forall\, t \in [\alpha, \beta]$. Then $\dim \operatorname{Ran} P(t) = \text{constant}$ $\forall\, t \in [\alpha, \beta]$.

Proof: By uniform continuity, $\exists\, n$ st $|t-t'| \le \frac{\beta-\alpha}{n}$ implies $\|P(t)-P(t')\| < 1$, $\alpha \le t, t' \le \beta$. Now for $t \in [\alpha, \beta]$, set $t_k = \alpha + \frac{k}{n}(t-\alpha)$, $0 \le k \le n$. Clearly $\|P(t_k) - P(t_{k+1})\| < 1$, $0 \le k \le n-1$, and so

$$\dim \operatorname{Ran} P(t_k) = \dim \operatorname{Ran} P(t_{k+1})$$

and we conclude that

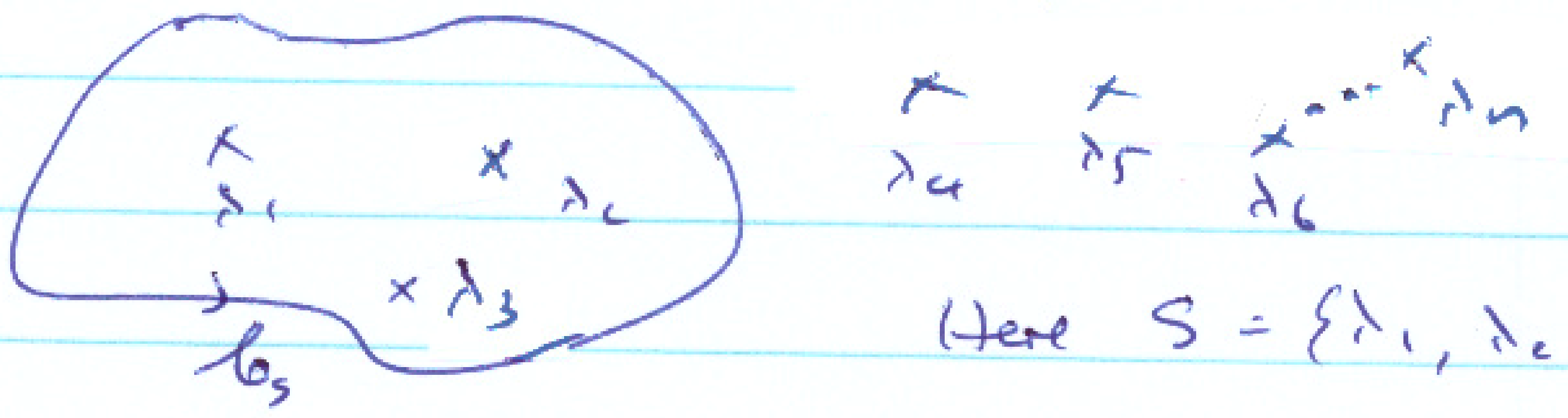
$$\dim \operatorname{Ran} P(\alpha) = \dim \operatorname{Ran} P(t_0) = \dim \operatorname{Ran} P(t_n)$$

$= \dim \operatorname{Ran} P(t)$, and we are done. $\square$

Now let A be an $n \times n$ matrix. Suppose $\{\lambda_1, \dots, \lambda_n\} = \sigma(A)$ and let S be any subset of $\sigma(A)$. and define (cf 66.1)

$$P_S(A) \equiv \frac{1}{2\pi i} \int_{\mathscr{C}_S} \frac{dz}{z - A}$$

where $\mathscr{C}_S$ is a smooth simple anti-clockwise curve enclosing S, but not enclosing $\sigma(A) \setminus S$



Here $S = \{\lambda_1, \lambda_2, \lambda_3\}$

Proposition 138.1

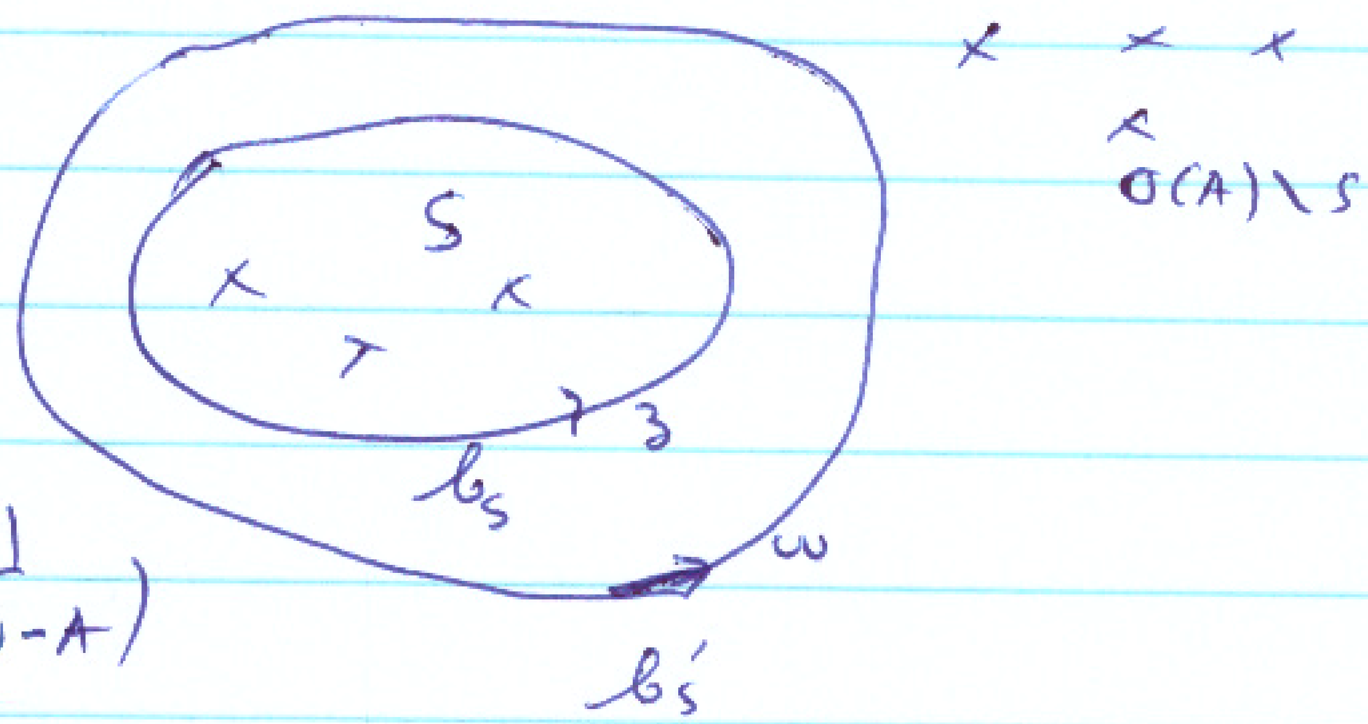
1) $P_S(A)$ is a projection commuting with A,
$P_S(A)A = AP_S(A)$. In particular, A
is reduced by $M_{P_S(A)} = \operatorname{Ran} P_S(A)$

2) $\sigma(A \restriction M_{P_S(A)}) = S$

Proof: 1) Similar to a previous calculation for the functional calculus (see p67)
Clearly $[P_S(A), A] = 0$.

Now $P_S^2(A) = \left(\frac{1}{2\pi i}\int_{\mathcal{C}_s} \frac{dz}{z-A}\right)\left(\frac{1}{2\pi i}\int_{\mathcal{C}_s'} \frac{dw}{w-A}\right)$

$= \left(\frac{1}{2\pi i}\right)^2 \iint_{\mathcal{C}_s \mathcal{C}_s'} dz\, dw \left(\frac{1}{z-A}\frac{1}{w-A}\right)$



$= \left(\frac{1}{2\pi i}\right)^2 \int_{\mathcal{C}_s}\int_{\mathcal{C}_s'} \frac{dz\, dw}{w-z}\left(\frac{1}{z-A} - \frac{1}{w-A}\right)$

$= \frac{1}{(2\pi i)^2}\left[\int_{\mathcal{C}_s} \frac{dz}{z-A}\int_{\mathcal{C}_s'} \frac{dw}{w-z} - \int_{\mathcal{C}_s'} \frac{dw}{w-A}\cancel{\int_{\mathcal{C}_s} \frac{dz}{w-z}}\right]$
$= 0$

$= \frac{1}{2\pi i}\int_{\mathcal{C}_s} \frac{dz}{z-A} = P_S(A)$

2) Suppose $\lambda_j \in S$. Then $\exists u_j \neq 0$ such that
$Au_j = \lambda_j u_j$. Now
$P_S(A)u_j = \frac{1}{2\pi i}\int_{\mathcal{C}_s}\left(\frac{1}{z-A}u_j\right)dz = \frac{1}{2\pi i}\left(\int_{\mathcal{C}_s} \frac{1}{z-\lambda_j}dz\right)u_j$

$= u_j$ as $\lambda_j \in$ interior of $\mathcal{C}_S$. Hence $u_j \in M_{P_S(A)}$.

and so $S \subset \sigma(A \restriction M_{P_S(A)})$. Conversely, suppose $\lambda \in \sigma(A \restriction M_{P_S(A)})$, then $\exists\, u = P_S(A)$, $u \neq 0$, such that $Au = \lambda u$. We have

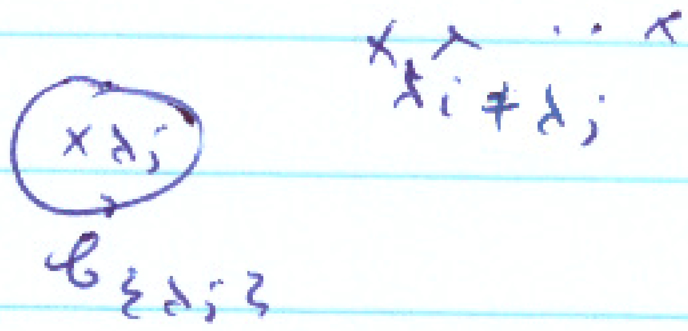
$$u = P_S(A)u = \frac{1}{2\pi i}\int_{\mathcal{C}_S} \frac{dz}{z-A}\, u = \frac{1}{2\pi i}\left(\int_{\mathcal{C}_S} \frac{dz}{z-\lambda}\right) u$$

Hence as $u \neq 0$, we must have $\frac{1}{2\pi i}\int_{\mathcal{C}_S} \frac{dz}{z-\lambda} = 1$

This is only possible if $\lambda \in$ interior of $\mathcal{C}_S$. Hence $\lambda \in S$ and we are done. $\square$

<u>Note:</u> If $\lambda_j \in \sigma(A)$, consider $P_{\{\lambda_j\}}(A)$,

Then clearly



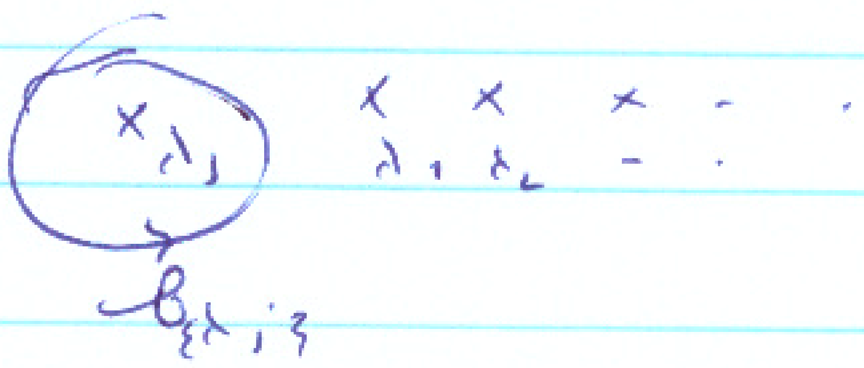
(139.1) $$P_S(A) = \sum_{\lambda_j \in S} P_{\{\lambda_j\}}(A)$$

<u>Proposition 139.1</u>

$$\dim M_{P_{\{\lambda_j\}}(A)} = \dim \operatorname{Ran}(P_{\{\lambda_j\}}(A)) = m_j = \text{algebraic multiplicity of } \lambda_j \in \sigma(A).$$

<u>Proof</u>: Let A_k be a sequence of matrices with simple spectrum converging to A as $k \to \infty$. Fix $\lambda_j \in \sigma(A)$,

algebraic mult. $= m_j$, and fix a suitable $\mathcal{C}_{\{\lambda_j\}}$



Then by previous results, for k large enough, A_k has m_j (distinct) eigenvalues inside $\mathcal{C}_{\{\lambda_j\}}$. Set

$$P_{(k)} = \frac{1}{2\pi i} \int_{\mathcal{C}_{\{\lambda_j\}}} \frac{dz}{z - A_k} .$$

Now in the basis which diagonalizes A_k, we have

$$(z - A_k)^{-1} = \begin{pmatrix} (z-\lambda_{k1})^{-1} & & & 0 \\ & (z-\lambda_{k2})^{-1} & & \\ & & \ddots & \\ 0 & & & (z-\lambda_{kn})^{-1} \end{pmatrix}$$

where without loss $\lambda_{k1}, \dots, \lambda_{km_j}$ are the eigenvalues of A_k inside $\mathcal{C}_{\{\lambda_j\}}$. Thus, in that basis,

$$P_{(k)} = \begin{pmatrix} \frac{1}{2\pi i}\int_{\mathcal{C}_{\{\lambda_j\}}} \frac{dz}{z-\lambda_{k1}} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{2\pi i}\int_{\mathcal{C}_{\{\lambda_j\}}} \frac{dz}{z-\lambda_{kn}} \end{pmatrix}$$

$$\begin{array}{cc} & \overset{m_j}{\longleftrightarrow} \\ m_j \updownarrow & \begin{pmatrix} 1 & & & & \\ & \ddots & & 0 & \\ & & 1 & & \\ 0 & & & 0 & \\ & & & & \ddots \end{pmatrix} \end{array}$$

Thus dim Ran $P_{(k)} = m_j$. But as $A_k \to A$, clearly $P_{(k)} \to P_{\{\lambda_j\}}(A)$. In particular $\| P_{(k)} - P_{\{\lambda_j\}}(A) \| < 1$, for k large enough. Thus dim Ran $P_{\{\lambda_j\}}(A) = \dim P_{(k)} = m_j$, by Lemma 135.1. We are done. $\square$.

<u>Remark</u> By calculations that are now familiar, for any $\ell = 0, 1, 2, \dots$,

$$(A - \lambda_i)^\ell P_{\{\lambda_i\}}(A) = \frac{1}{2\pi i} \int_{C_{\{\lambda_i\}}} (z - \lambda_i)^\ell \frac{dz}{z - A}.$$

If $\ell = m_i$, then as $\det(z - A) = (z - \lambda_i)^{m_i} \prod_{\lambda_j \neq \lambda_i} (z - \lambda_j)$ we see that $(A - \lambda_i)^\ell P_{\{\lambda_i\}}(A) = 0$, by Cramer's rule. Thus $(A - \lambda_i)$ is <u>nilpotent</u> in Ran $P_{\{\lambda_i\}}(A)$. By choosing the basis appropriately in each $P_{\{\lambda_i\}}(A)$, $\lambda_i \in \sigma(A)$, it is not difficult to prove the existence of the

Jordan form for A (Exercise: Reference Kato Chap. I).

First Perturbation Theorem (142.1):
Suppose A_0 is an $n \times n$ matrix and suppose that $\lambda \in \sigma(A_0)$ is an algebraically simple eigenvalue of A_0, $A_0 v = \lambda v$, $v \neq 0$. Then for A in a neighborhood $S_\delta = \{A : \|A - A_0\| < \delta\}$ for some $\delta > 0$, A has an algebraically simple eigenvalue $\lambda(A)$ which is analytic in (the entries $\{a_{ij}\}$ of) A in S_δ. Furthermore, $v(A)$, the eigenvector associated with $\lambda(A)$, $A v(A) = \lambda(A) v(A)$, may be chosen as an analytic function of A in S_δ, and $v(A_0) = v$.

Proof Let $\varepsilon > 0$ be such that $\mathcal{C}_{\{\lambda\}} = \{z : |z - \lambda| = \varepsilon\}$

contains only λ from $\sigma(A_0)$. As λ has alg. mult. $= 1$,
$P_{\{\lambda\}}(A_0) = \frac{1}{2\pi i} \oint_{\mathcal{C}_{\{\lambda\}}} \frac{dz}{z - A_0}$ has 1-dimensional range. By

By Lemma 135.1, for δ small enough such that

$$\left\| P_{\{\lambda\}}(A_0) - \frac{1}{2\pi i} \int_{\mathcal{C}_{\{\lambda\}}} \frac{dz}{z - A} \right\| < 1$$

for $\|A - A_0\| < \delta$, $P(A) = \frac{1}{2\pi i} \int_{\mathcal{C}_{\{\lambda\}}} \frac{dz}{z - A}$

is also has 1-dimensional range, $\sigma(A \restriction \operatorname{Ran} P(A))$ is the spectrum of A inside $\mathcal{C}_{\{\lambda\}}$.

Now note first that for $\|A - A_0\| < \delta$, $P(A)$ depends analytically on the entries $\{a_{ij}\}$ of A. Indeed

$$\frac{\partial P}{\partial a_{ij}}(A) = \frac{1}{2\pi i} \int_{\mathcal{C}_{\{\lambda\}}} dz \, \frac{1}{z-A} E_{ij} \frac{1}{z-A}$$

where E_{ij} is the matrix with 1 in the (i,j) position, and zero elsewhere. Of course, as $A \to A_0$, $P(A) \to P_{\{\lambda\}}(A_0)$. Now as $\operatorname{Ran} P_{\{\lambda\}}(A_0)$ is 1-dimensional, $P_{\{\lambda\}}(A_0) = (u, \cdot) v$ for some non-zero vectors $u, v \in \mathbb{C}^n$, and $(u,v) = 1$ as $P^2_{\{\lambda\}}(A_0) = P_{\{\lambda\}}(A_0)$.

Thus $P_{\{\lambda\}}(A_0) v = v$. Set

(143.1) $$v(A) = P(A) v$$

Then for $\delta > 0$ small enough, $v(A) \neq 0$ for all $\|A - A_0\| < \delta$, by continuity. As $\dim \operatorname{Ran} P(A)$ is also $= 1$, we must have similarly

(143.2) $$P(A) = (u(A), \cdot) \, v(A)$$

for some $(u(A), v(A)) = 1$

Now $P(A)$ commutes with A. Hence

$$A v(A) = A\, P(A) v(A) = P(A) A v(A)$$
$$= (u(A), A v(A))\, v(A)$$

and we see that $v(A)$ is an eigenvector of A with eigenvalue

$$\lambda(A) = (u(A), A\, v(A))$$

But as $v(A) = P(A) v$ and $P(A)$ is analytic, so is $v(A)$. As $A \to A_0$, $v(A) \to v$. Thus $(v, v(A)) \neq 0$ and analytic for $\|A - A_0\|$ small.

We have

$$(v, A v(A)) = \lambda(A)\, (v, v(A))$$

i.e.

(144.1) $$\lambda(A) = \frac{(v, A\, v(A))}{(v, v(A))} = \frac{(v, A\, P(A) v)}{(v, P(A) v)}$$

In particular, $\lambda(A)$ depends analytically on A for $\|A - A_0\|$ small. This completes the proof. $\square$

Of course, the eigenvector of A is determined only up to a scalar multiple. In the above proof we have just displayed a vector $v(A)$ which is non-zero, analytic (for $\|A - A_0\|$ small), and which

solves $A\, v(A) = \lambda(A)\, v(A)$. But any multiple
$$\tilde{v}(A) = \gamma(A)\, v(A)$$

with $\gamma(A)$ analytic and $\neq 0$, would also do. For

example
$$\tilde{v}(A) = \frac{v(A)}{(v, v(A))}$$

is the unique eigenvector of A such that

$$(v, \tilde{v}(A)) = 1$$

We will give other normalizations below.

Suppose that we consider

$$A(\beta) = A_0 + \beta B$$

for some given matrices A_0 and B, $\beta \in \mathbb{C}$. Suppose

λ_0 is an algebraically simple eigenvalue of A_0 with

associated eigenvector $v_0 \neq 0$. As $A_0 v_0 = \lambda_0 v_0$,

$v_0 = P_{\{\lambda_0\}}(A_0)\, v_0$. Then by the composition rule

($A(\beta)$ clearly depends analytically on β), $A(\beta)$ has an

analytic algebraically simple eigenvalue $\lambda(\beta) = \lambda(A(\beta))$ near

λ_0, $\lambda(\beta) \to \lambda_0$ as $\beta \to 0$, and an analytic

eigenvector $v(\beta) = v(A(\beta))$, $v(\beta) \to v_0$ as $\beta \to 0$.

We now write out the first terms in the power

series

(146.1) $$\lambda(\beta) = \lambda_0 + \alpha_1 \beta + \alpha_2 \beta^2 + \dots$$

(146.2) $$v(\beta) = v_0 + v_1 \beta + v_2 \beta^2 + \dots$$

in the special case that $A_0 = A_0^*$ i.e. A_0 is self-adjoint. From (144.1) we have

$$\lambda(\beta) = \lambda(A(\beta)) = \frac{(v_0, (A_0 + \beta B) P(A_0 + \beta B) v_0)}{(v_0, P(A_0 + \beta B) v_0)}$$

Now for $|\beta|$ small, so that $\|A(\beta) - A_0\| = \|\beta B\|$
$= |\beta| \, \|B\| < \delta$

in the above Theorem (142.1), and for z

$$z \in \mathcal{C}_{\{\lambda_0\}} = \{ z : |z - \lambda_0| = \varepsilon \}$$

for some $\varepsilon > 0$,

$$\frac{1}{z - A(\beta)} = \frac{1}{z - A_0 - \beta B} = \frac{1}{z - A_0} \, \frac{1}{1 - \beta B (z - A_0)^{-1}}$$

$$= (z - A_0)^{-1} \left(1 + \beta B \frac{1}{z - A_0} + \beta^2 B \frac{1}{z - A_0} B \frac{1}{z - A_0} + \dots \right.$$

$$= (z - A_0)^{-1} + \beta (z - A_0)^{-1} B (z - A_0)^{-1} + \beta^2 (z - A_0)^{-1} B (z - A_0)^{-1} B (z - A_0)^{-1} + \dots$$

Thus

$$P(A_0+\beta B) = \frac{1}{2\pi i} \oint_{|z-\lambda_0|=\varepsilon} \frac{1}{(z-A(\beta))}\, dz$$

$$= \frac{1}{2\pi i} \oint_{|z-\lambda_0|=\varepsilon} \frac{1}{z-A_0}\, dz$$

$$+ \beta \frac{1}{2\pi i} \oint_{|z-\lambda_0|=\varepsilon} \frac{1}{z-A_0} B \frac{1}{z-A_0}\, dz$$

$$+ \beta^2 \frac{1}{2\pi i} \oint_{|z-\lambda_0|=\varepsilon} \frac{1}{z-A_0} B \frac{1}{z-A_0} B \frac{1}{z-A_0}\, dz$$

$$+ \dots$$

Thus

(147.1) $$P(A_0+\beta B) = P_0 + \beta P_1 + \beta^2 P_2 + \dots$$

where

(147.2) $$P_j = \frac{1}{2\pi i} \oint_{|z-\lambda_0|=\varepsilon} \frac{1}{z-A_0}\left(B\frac{1}{(z-A_0)}\right)^j dz, \quad j \geq 0.$$

Now

$$P_0 = \frac{1}{2\pi i} \int_{|z-\lambda_0|=\varepsilon} \frac{1}{z-A_0}\, dz = P_{\{\lambda_0\}}(A_0)$$

Thus

$$\lambda(\beta) = \frac{(v_0, (A_0+\beta B)(P_0+\beta P_1+\dots)v_0)}{(v_0, (P_0+\beta P_1+\dots)v_0)}$$

$$= \frac{(v_0, A_0 P_0 v_0) + \beta\left((v_0, B P_0 v_0) + (v_0, A_0 P_1 v_0) + \cdots\right)}{(v_0, P_0 v_0) + \beta (v_0, P_1 v_0) + \cdots}$$

As $v_0 = P_0 v_0$, $(v_0, A_0 P_0 v_0) = (v_0, A_0 v_0)$

$= \lambda_0 (v_0, v_0)$, λ_0 real (as $A_0 = A_0^*$)

$(v_0, B P_0 v_0) = (v_0, B v_0)$

$(v_0, A_0 P_1 v_0) = (A_0 v_0, P_1 v_0)$

$= \lambda_0 (v_0, P_1 v_0)$, as λ_0 real

Finally

$$(v_0, P_1 v_0) = \frac{1}{2\pi i} \int_{|z-\lambda_0| = \varepsilon} \left(v_0, \frac{1}{z - A_0} B \frac{1}{z - A_0} v_0\right) dz$$

$$= \frac{1}{2\pi i} \int_{|z-\lambda_0| = \varepsilon} \left(\frac{1}{z - A_0} v_0, B \frac{1}{z - A_0} v_0\right) dz$$

$$= \frac{1}{2\pi i} \int_{|z-\lambda_0| = \varepsilon} \left(\frac{1}{z - \lambda_0} v_0, B \frac{1}{z - \lambda_0} v_0\right) dz$$

$$= \frac{1}{2\pi i} \int_{|z-\lambda_0| = \varepsilon} (v_0, B v_0) \frac{1}{(z-\lambda_0)^2} dz$$

$$= \frac{(v_0, B v_0)}{2\pi i} \int_{|z-\lambda_0| = \varepsilon} \frac{1}{(z-\lambda_0)^2} dz$$

$$= 0$$

Thus we find

$$\lambda(\beta) = \left(\lambda_0 (v_0, v_0) + \beta (v_0, B v_0) + O(\beta^2)\right) \Big/ \left((v_0, v_0) + O(\beta^2)\right)$$

(148.1) so $\lambda(\beta) = \lambda_0 + \beta (v_0, B v_0) / (v_0, v_0) + O(\beta^2)$

Also

$$v(\beta) = P_0 v_0 + \beta P_1 v_0 + O(\beta^2)$$

$$= v_0 + \frac{\beta}{2\pi i} \int_{|z-\lambda_0|=\varepsilon} \frac{1}{z-A_0} B \frac{1}{z-A_0} v_0 \, dz + O(\beta^2)$$

$$= v_0 + \frac{\beta}{2\pi i} \int_{|z-\lambda_0|=\varepsilon} \frac{1}{z-A_0} B \frac{dz}{z-\lambda_0} v_0 + O(\beta^2)$$

Now let $v_0, v_1, \ldots, v_{n-1}$ be the orthonormalized eigenvectors of A_0 corresponding to $\lambda_0, \lambda_1, \ldots, \lambda_{n-1}$, resp., $\lambda_i \neq \lambda_0$ for $i \neq 0$.

Then $\left(v_i, \int_{|z-\lambda_0|=\varepsilon} \frac{1}{z-A_0} B \frac{dz}{z-\lambda_0} v_0\right)$

$$= \frac{1}{2\pi i} \int_{|z-\lambda_0|=\varepsilon} \left(\frac{1}{\bar{z}-A_0} v_i, \frac{B v_0 \, dz}{z-\lambda_0}\right)$$

$$= \frac{1}{2\pi i} \int_{|z-\lambda_0|=\varepsilon} (v_i, B v_0) \frac{1}{(z-\lambda_i)(z-\lambda_0)} dz$$

If $i=0$, then this integral is zero, as before. If $i \neq 0$, this integral is

$$\frac{(v_i, B v_0)}{\lambda_0 - \lambda_i}$$

Hence

(149.1) $$v(\beta) = v_0 + \beta \sum_{i \neq 0} \frac{(v_i, B v_0)}{\lambda_0 - \lambda_i} + O(\beta^2)$$

(148.1) and (149.1) are the basic formulae for the perturbation $A_0 \to A_0 + \beta B$ of an algebraically simple eigenvalue of A_0 in the case that A_0 is self-adjoint.

Exercise Apply the above calculations, where relevant, to the examples (a) → (f) in the previous lecture in the neighborhood of an eigenvalue.