Perturbation Theory

Martin Ravn Christiansen
tvl589@alumni.ku.dk

June 21, 2017

Introduction

Informally speaking, perturbation theory can be considered the study of how operators A and $A+B$ are related when B can be considered "small" in some way. In the course we have already seen some instances of perturbative results. One example is the Kato-Rellich theorem, which states

Theorem 1 (**Kato-Rellich**). *Let $A : \mathcal{D}(A) \subset \mathcal{H} \to \mathcal{H}$ be self-adjoint, $B : \mathcal{D}(B) \subset \mathcal{H} \to \mathcal{H}$ be symmetric and $\mathcal{D}(A) \subset \mathcal{D}(B)$. If there exists $a < 1$ and $b \in \mathbb{R}$ such that*

$$\|B\varphi\| \leq a\|A\varphi\| + b\|\varphi\|, \quad \forall \varphi \in \mathcal{D}(A)$$

then $A + B : \mathcal{D}(A) \subset \mathcal{H} \to \mathcal{H}$ is self-adjoint.

In this case, the condition that B is small is the statement that $\|B\varphi\| \leq a\|A\varphi\| + b\|\varphi\|$ holds for all $\varphi \in \mathcal{D}(A)$ and the relation this implies between A and $A+B$ is that $A+B$ is self-adjoint on $\mathcal{D}(A)$.

Another perturbative result we have seen is the stability of the essential spectrum:

Theorem 2 (**Stability of the Essential Spectrum**). *Let A and A' be self-adjoint operators on $\mathcal{H}$. If there exists a $z \in \rho(A) \cap \rho(A')$ such that $(A-z)^{-1} - (A'-z)^{-1}$ is compact then $\sigma_{ess}(A) = \sigma_{ess}(A')$. In particular, if A' is of the form $A' = A + B$ where B is compact then the conclusion holds.*

Here the "smallness" condition is that B is compact, and the relation this implies is that $\sigma_{\text{ess}}(A+B) = \sigma_{\text{ess}}(A)$.

In this summary we will present some of the main results on the behaviour of (simple) eigenvalues under a perturbation.

While we will prove a variety of results, perhaps the most interesting single result is the following:

Theorem 3. *Let $A : \mathcal{D}(A) \subset \mathcal{H} \to \mathcal{H}$ be self-adjoint, $B : \mathcal{D}(B) \subset \mathcal{H} \to \mathcal{H}$ be an operator and $\mathcal{D}(A) \subset \mathcal{D}(B)$. Let furthermore E_0 be an isolated non-degenerate eigenvalue of A, i.e. $dist(E_0, \sigma(A) \setminus \{E_0\}) = 2\varepsilon > 0$. Assume that an estimate of the form*

$$\|B\varphi\| \leq a\|A\varphi\| + b\|\varphi\|, \quad \forall \varphi \in \mathcal{D}(A)$$

holds for some $a, b \in \mathbb{R}$. Then for all $\beta \in B_0 = B\left(0, \left(a + (a(|E_0| + \varepsilon) + b)\varepsilon^{-1}\right)^{-1}\right)$ there is exactly one spectral point $E(\beta)$ of $A + \beta B : \mathcal{D}(A) \subset \mathcal{H} \to \mathcal{H}$ in the ball $B(E_0, \varepsilon)$. $E(\beta)$ is a non-degenerate eigenvalue and considered as a function of β it is analytic on some neighbourhood of 0 contained in B_0.

It can be shown that $E(\beta)$ is in fact analytic on all of B_0, see for instance Kato's classic book [1].

Results

This section is a simplified version of the presentation of the material in Reed and Simon's book [2], approximately corresponding to Theorem XII.5 through Theorem XII.9.

First we recall the definition of an analytic bounded operator-valued function:

Definition 1. An operator-valued function $L(\beta) : U \to \mathcal{L}(\mathcal{H})$ defined on an open set $U \subset \mathbb{C}$ is said to be analytic if the function

$$z \mapsto \langle \psi, L(\beta)\varphi \rangle$$

is analytic as a complex function for all φ, $\psi \in \mathcal{H}$.

Recall in particular that for an operator A the resolvent $(A-\lambda)^{-1}$ is analytic in λ for all $\lambda \in \rho(A)$.

We will study the perturbation of eigenvalues by studying the perturbation of the associated eigenvectors. Since we do not know what the eigenvector will be for the perturbed function, we will need to define an appropriate projection operator. In the self-adjoint case we can define projections using functional calculus, but since we will look at families of the form $A + \beta B$ where β is taken from an open complex set this will not suffice, so we need to extend spectral projections to general closed operators. We do this by contour integration.

Theorem 4. *Let A be a closed operator and let λ_0 be an isolated point of $\sigma(A)$, i.e. there exists an $\varepsilon > 0$ such that $B(\lambda, \varepsilon) \cap \sigma(A) = \{\lambda_0\}$. Let $\mathcal{C}(r) = \{\lambda \in \mathbb{C} \mid |\lambda_0 - \lambda| = r\}$ for some $0 < r < \varepsilon$ and define the operator $P \in \mathcal{L}(\mathcal{H})$ by*

$$P = -\frac{1}{2\pi i} \oint_{\mathcal{C}(r)} (A-\lambda)^{-1} \, d\lambda.$$

Then P has the following properties:

1. *P is well-defined and is independent of the r chosen.*
2. *P is a projection, i.e. $P^2 = P$.*
3. *P is the projection onto the λ_0-eigenspace $\{\varphi \in \mathcal{D}(A) \mid A\varphi = \lambda_0 \varphi\}$ of A.*

Proof. Since $\mathcal{C}(r)$ is compact and $(A-\lambda)^{-1}$ analytic we have $\sup_{\lambda \in \mathcal{C}(r)} \left((A-\lambda)^{-1} \right) < \infty$ so the integral exists as a Banach space-valued Riemann integral. That it is independent of r follows from the fact that for φ, $\psi \in \mathcal{H}$ we have

$$\langle \psi, P\varphi \rangle = -\frac{1}{2\pi i} \oint_{\mathcal{C}(r)} \left\langle \psi, (A-\lambda)^{-1} \varphi \right\rangle d\lambda.$$

By analyticity of $(A-\lambda)^{-1}$ we have that $\lambda \mapsto \left\langle \psi, (A-\lambda)^{-1} \varphi \right\rangle$ is a holomorphic function defined on $\rho(A)$, hence if it has a pole in $B(\lambda_0, \varepsilon)$ it must be at $\lambda = \lambda_0$, so in fact

$$\langle \psi, P\varphi \rangle = \frac{i}{2\pi} \mathrm{Res}\left(\left\langle \psi, (A-\lambda)^{-1} \varphi \right\rangle, \lambda_0 \right). \tag{1}$$

That the integral is r-independent then follows from the usual Cauchy integral theorem. Since φ, ψ were arbitrary we conclude that r-independence holds in general. This shows (1).

To show that P is a projection, note that if $0 < r < R < \varepsilon$ then by what we have just shown

$$P = -\frac{1}{2\pi i} \oint_{\mathcal{C}(r)} (A-\lambda)^{-1} \, d\lambda = -\frac{1}{2\pi i} \oint_{\mathcal{C}(R)} (A-\lambda)^{-1} \, d\lambda,$$

hence using the first resolvent identity $(A-\lambda)^{-1} - (A-\mu)^{-1} = (\lambda - \mu)(A-\lambda)^{-1}(A-\mu)^{-1}$ we can write

$$\begin{aligned}
(-2\pi i)^2 P^2 &= \oint_{\mathcal{C}(r)} \oint_{\mathcal{C}(R)} (A-\lambda)^{-1} (A-\mu)^{-1} \, d\lambda d\mu \\
&= \oint_{\mathcal{C}(r)} \oint_{\mathcal{C}(R)} \frac{1}{\lambda - \mu} (A-\lambda)^{-1} \, d\lambda d\mu - \oint_{\mathcal{C}(r)} \oint_{\mathcal{C}(R)} \frac{1}{\lambda - \mu} (A-\mu)^{-1} \, d\lambda d\mu \\
&= \oint_{\mathcal{C}(R)} \left(\oint_{\mathcal{C}(r)} \frac{1}{\lambda - \mu} \, d\mu \right) (A-\lambda)^{-1} \, d\lambda - \oint_{\mathcal{C}(r)} \left(\oint_{\mathcal{C}(R)} \frac{1}{\lambda - \mu} \, d\lambda \right) (A-\mu)^{-1} \, d\mu.
\end{aligned}$$

Since $r < R$ the first (nested) integral is 0 while the second one equals $2\pi i$, hence

$$(-2\pi i)^2 P^2 = -2\pi i \oint_{\mathcal{C}(r)} (A-\mu)^{-1} \, d\mu = (-2\pi i)^2 P.$$

This shows (2). Finally we show that $\operatorname{ran}(P) = \{\varphi \in \mathcal{D}(A) \mid A\varphi = \lambda_0 \varphi\}$. First, let $A\varphi = \lambda\varphi$. Then

$$P(\varphi) = -\frac{1}{2\pi i} \oint_{\mathcal{C}(r)} (A-\lambda)^{-1} \varphi \, d\lambda = -\frac{1}{2\pi i} \oint_{\mathcal{C}(r)} (\lambda_0 - \lambda)^{-1} \varphi \, d\lambda = \left(\frac{1}{2\pi i} \oint_{\mathcal{C}(r)} \frac{1}{\lambda - \lambda_0} \, d\lambda \right) \varphi = \varphi,$$

so $\{\varphi \in \mathcal{D}(A) \mid A\varphi = \lambda_0 \varphi\} \subset \operatorname{ran}(P)$. On the other hand, by the identity $A(A-\lambda)^{-1} = 1 + \lambda (A-\lambda)^{-1}$, we have for any $\varphi \in \mathcal{H}$ that

$$AP\varphi = -\frac{1}{2\pi i} \oint_{\mathcal{C}(r)} A(A-\lambda)^{-1} \varphi \, d\lambda = -\frac{1}{2\pi i} \oint_{\mathcal{C}(r)} \left(1 + \lambda (A-\lambda)^{-1}\right) \varphi \, d\lambda = -\frac{1}{2\pi i} \oint_{\mathcal{C}(r)} \lambda (A-\lambda)^{-1} \varphi \, d\lambda,$$

where we used that $\oint_{\mathcal{C}(r)} 1 \, d\lambda = 0$. Using equation (1) we see that for any $\psi \in \mathcal{H}$

$$\langle \psi, AP\varphi \rangle = \frac{i}{2\pi} \operatorname{Res}\left(\lambda \left\langle \psi, (A-\lambda)^{-1} \varphi \right\rangle, \lambda_0 \right) = \lambda_0 \frac{i}{2\pi} \operatorname{Res}\left(\left\langle \psi, (A-\lambda)^{-1} \varphi \right\rangle, \lambda_0 \right) = \lambda_0 \langle \psi, P\varphi \rangle = \langle \psi, \lambda_0 P\varphi \rangle,$$

so by Riesz lemma

$$AP\varphi = \lambda_0 P\varphi,$$

hence $\operatorname{ran}(P) \subset \{\varphi \in \mathcal{D}(A) \mid A\varphi = \lambda_0 \varphi\}$. This concludes the proof. □

Remark. In the above theorem we assume that $B(\lambda, \varepsilon) \cap \sigma(A) = \{\lambda_0\}$, but it is clear that everything still holds if $B(\lambda, \varepsilon) \cap \sigma(A) = \emptyset$ with the conclusion that $P = 0$. In addition, even if we do not know that $\mathcal{C}(r)$ surrounds a spectral point then if $\dim(\operatorname{Ran}(P)) = 1$ it must in fact surround an eigenvalue by (2) and (3).

Before we define our main object of study, we introduce a new definition:

Definition 2. A vector-valued function $\varphi(\beta) : U \to \mathcal{H}$ defined on an open set $U \subset \mathbb{C}$ is said to be analytic if

$$\beta \mapsto \langle \psi, \varphi(\beta) \rangle$$

is analytic as a complex function for all $\psi \in \mathcal{H}$.

Note in particular that if $L(\beta) : U \to \mathcal{L}(\mathcal{H})$ is analytic then $L(\beta)\varphi : U \to \mathcal{L}(\mathcal{H})$ is analytic for any $\varphi \in \mathcal{H}$.

Now we can extend a notion of analyticity to unbounded operators:

Definition 3. Let $U \subset \mathbb{C}$ be a connected, open set and for all $\beta \in U$ let $T(\beta)$ be a closed operator with $\rho(T(\beta)) \neq \emptyset$. We say that $T(\beta)$ is an analytic family (of type A) if

1. $\mathcal{D}(T(\beta)) = D$ is independent of β.
2. $T(\beta)\varphi : U \to \mathcal{H}$ is analytic for all $\varphi \in D$.

We will be focusing on linear analytic families of type A, i.e. those that are of the form $A + \beta B$ for some operators A and B. We have a result characterizing such families:

Theorem 5. *Let $A : \mathcal{D}(A) \subset \mathcal{H} \to \mathcal{H}$ be a closed operator with $\rho(A) \neq \emptyset$. Let $B : \mathcal{D}(B) \subset \mathcal{H} \to \mathcal{H}$ be another operator and define $A + \beta B$ on $\mathcal{D}(A) \cap \mathcal{D}(B)$. Then $A + \beta B$ is an analytic family of type A on an open set $U \subset \mathbb{C}$ containing 0 if and only if $\mathcal{D}(A) \subset \mathcal{D}(B)$ and there exists $a, b > 0$ such that*

$$\|B\varphi\| \le a \|A\varphi\| + b \|\varphi\|, \qquad \forall \varphi \in \mathcal{D}(A). \qquad (*)$$

Proof. First assume that $A+\beta B$ is an analytic family of type A. Since the domain is independent of β we have that

$$\mathcal{D}(A)=\mathcal{D}(A+\beta B)=\mathcal{D}(A)\cap\mathcal{D}(B)$$

hence $\mathcal{D}(A)\subset\mathcal{D}(B)$. Now, since A is closed $D=\mathcal{D}(A)$ is a Banach space with the graph norm

$$\|\varphi\|_A=\|\varphi\|+\|A\varphi\|,\quad \varphi\in\mathcal{D}(A).$$

Let β be so small that both β and $-\beta$ are in U. Then the graphs of $A+\beta B$ and $A-\beta B$ are closed in $\mathcal{H}\times\mathcal{H}$ by assumption, hence also in $D\times\mathcal{H}$ (which has a stronger topology). Then, by the closed graph theorem,

$$\|(A+\beta B)\varphi\|\le a\|\varphi\|_A,\quad \|(A-\beta B)\varphi\|\le a'\|\varphi\|_A,\quad \forall\varphi\in\mathcal{D}(A)$$

hence for all $\varphi\in\mathcal{D}(A)$

$$\begin{aligned}\|B\varphi\| &= \left\|\frac{1}{2\beta}\left((A+\beta B)\varphi+(A-\beta B)\varphi\right)\right\|\le\frac{1}{2|\beta|}\left(\|(A+\beta B)\varphi\|+\|(A-\beta B)\varphi\|\right)\\ &\le \frac{a+a'}{2|\beta|}\|\varphi\|_A=\frac{a+a'}{2|\beta|}\left(\|A\varphi\|+\|\varphi\|\right)\end{aligned}$$

so we have an inequality of the form $(*)$.

Now assume that $\mathcal{D}(A)\subset\mathcal{D}(B)$ and that $(*)$ holds. Then $\mathcal{D}(A+\beta B)=\mathcal{D}(A)\cap\mathcal{D}(B)=\mathcal{D}(A)$ is constant by definition and that $(A+\beta B)\varphi$ is analytic for all $\varphi\in\mathcal{D}(A)$ is obvious, since $\langle\psi,(A+\beta B)\varphi\rangle=\langle\psi,A\varphi\rangle+\langle\psi,B\varphi\rangle\beta$ is linear for all $\psi\in\mathcal{D}(A)$, hence certainly analytic, so all that remains to be shown is closedness and the resolvent condition. For the closedness, observe that for all $\varphi\in\mathcal{D}(A)$

$$\|A\varphi\|\le\|(A+\beta B)\varphi\|+|\beta|\,\|B\varphi\|\le\|(A+\beta B)\varphi\|+|\beta|\,a\,\|A\varphi\|+|\beta|\,b\,\|\varphi\|,$$

hence for $|\beta|<a^{-1}$

$$\|A\varphi\|\le\frac{1}{1-a|\beta|}\|(A+\beta B)\varphi\|+\frac{|\beta|\,b}{1-a|\beta|}\|\varphi\|,\quad\forall\varphi\in\mathcal{D}(A).$$

This implies that $\|\cdot\|_A$ is weaker than $\|\cdot\|_{A+\beta B}$ for $|\beta|<a^{-1}$.

Now, let $(\varphi_n)_{n\in\mathbb{N}}\subset\mathcal{D}(A)$ be such that $\varphi_n\to\varphi\in\mathcal{H}$ as $n\to\infty$ (in $\|\cdot\|$) and $((A+\beta B)\varphi_n)_{n\in\mathbb{N}}$ also converges, hence is Cauchy. By the above inequality the sequence $(A\varphi_n)_{n\in\mathbb{N}}$ is then also Cauchy, hence by closedness of A we conclude that $\varphi\in\mathcal{D}(A)$ and $A\varphi_n\to A\varphi$ as $n\to\infty$. Then also $(A+\beta B)\varphi_n\to(A+\beta B)\varphi$ as $n\to\infty$, since $(*)$ assures that B is continuous in the graph norm of A, so we conclude that $A+\beta B$ is closed for all $|\beta|<a^{-1}$.

Finally we must show that $\rho(A+\beta B)$ is non-empty for $|\beta|$ small enough. Let $\lambda\in\rho(A)$ and write

$$A+\beta B-\lambda=\left(1+\beta B(A-\lambda)^{-1}\right)(A-\lambda).$$

To show that $\lambda\in\rho(A+\beta B)$ it is then sufficient to show that we can choose β such that $\left\|\beta B(A-\lambda)^{-1}\right\|<1$, by Neumann series. Using $(*)$ we estimate for $\varphi\in\mathcal{H}$

$$\begin{aligned}\left\|B(A-\lambda)^{-1}\varphi\right\| &\le a\left\|A(A-\lambda)^{-1}\varphi\right\|+b\left\|(A-\lambda)^{-1}\varphi\right\|\le a\|\varphi\|+a|\lambda|\left\|(A-\lambda)^{-1}\varphi\right\|+b\left\|(A-\lambda)^{-1}\varphi\right\|\\ &\le \left(a+(a|\lambda|+b)\left\|(A-\lambda)^{-1}\right\|\right)\|\varphi\|\end{aligned}$$

where we used the identity $A(A-\lambda)^{-1}=1+\lambda(A-\lambda)^{-1}$. We see that demanding $|\beta|<\left(a+(a|\lambda|+b)\left\|(A-\lambda)^{-1}\right\|\right)^{-1}$ does the job, so we are done. $\square$

When A is self-adjoint we have the following corollary:

Corollary 1. *Let $A:\mathcal{D}(A)\subset\mathcal{H}\to\mathcal{H}$ be self-adjoint, $B:\mathcal{D}(B)\subset\mathcal{H}\to\mathcal{H}$ be an operator and $\mathcal{D}(A)\subset\mathcal{D}(B)$. If there exists $a,\, b\in\mathbb{R}$ such that*

$$\|B\varphi\|\leq a\|A\varphi\|+b\|\varphi\|, \quad \forall\varphi\in\mathcal{D}(A)$$

then $A+\beta B:\mathcal{D}(A)\subset\mathcal{H}\to\mathcal{H}$ is an analytic family of type A for all $\beta\in B\left(0,a^{-1}\right)$.

Proof. The corollary follows from the fact that in the second half of the proof above we only had to (possibly) shrink the domain of $A+\beta B$ from $B\left(0,a^{-1}\right)$ to ensure a non-empty resolvent set. For a self-adjoint A we know that $i\lambda\in\rho(A)$ for all $\lambda\in\mathbb{R}\setminus\{0\}$, so we can do as in the proof of the Kato-Rellich theorem and write

$$(A+\beta B-i\lambda)=\left(1+\beta B(A-i\lambda)^{-1}\right)(A-i\lambda)$$

followed by the estimates (recalling that for self-adjoint A we also have the estimate $\left\|A(A-i\lambda)^{-1}\varphi\right\|\leq\|\varphi\|$)

$$\left\|\beta B(A-i\lambda)^{-1}\varphi\right\|\leq a|\beta|\left\|A(A-i\lambda)^{-1}\varphi\right\|+b|\beta|\left\|(A-i\lambda)^{-1}\varphi\right\|\leq a|\beta|\|\varphi\|+\frac{b|\beta|}{|\lambda|}\|\varphi\|, \quad \varphi\in\mathcal{D}(A),$$

hence we can establish that $A+\beta B$ has a non-empty resolvent set by enlarging $|\lambda|$ instead of shrinking $|\beta|$. □

Remark. First, note that if B is also symmetric then $A+\beta B$ is also self-adjoint for $\beta\in B\left(0,a^{-1}\right)\cap\mathbb{R}$ by Kato-Rellich. Secondly, if B is infinitesimally bounded with respect to A then $A+\beta B$ is an analytic family for all $\beta\in\mathbb{C}$.

Our strategy for studying the eigenvalues of an analytic family $A+\beta B$ will be to use our spectral projection to define a β-dependent projection $P(\beta)$ using Theorem 4, and then formulate eigenvectors and eigenvalues in terms of $P(\beta)$. In order to do this, however, we need to make sure that we can control β in such a way that the spectrum $\sigma(A+\beta B)$ stays away from the integration contour that defines $P(\beta)$. This we turn to now.

Theorem 6. *Let $A+\beta B$ be an analytic family of type A. Let $\lambda_0\in\sigma(A)$ be such that $\mathrm{dist}(\lambda_0,\sigma(A)\setminus\{\lambda_0\})=2\varepsilon>0$. Then there exists an $r>0$ such that the set*

$$\mathcal{C}(\varepsilon)=\{\lambda\in\mathbb{C}\mid|\lambda_0-\lambda|=\varepsilon\}$$

obeys $\mathrm{dist}(\mathcal{C}(\varepsilon),\sigma(A+\beta B))>0$ for all $\beta\in B(0,r)$. Furthermore, if A is self-adjoint then one can take

$$r=\left(a+\frac{a(|\lambda_0|+\varepsilon)+b}{\varepsilon}\right)^{-1}.$$

Proof. As we saw in the last part of Theorem 5 we have by writing

$$A+\beta B-\lambda=\left(1+\beta B(A-\lambda)^{-1}\right)(A-\lambda)$$

the conclusion that $\lambda\in\rho(A+\beta B)$ if $\left\|\beta B(A-\lambda)^{-1}\right\|<1$. We need to show that we can ensure this for all $\lambda\in\mathcal{C}(\varepsilon)$ by a single constraint on β. In Theorem 5 we also found the estimate

$$\left\|B(A-\lambda)^{-1}\right\|\leq a+(a|\lambda|+b)\left\|(A-\lambda)^{-1}\right\|, \quad \lambda\in\rho(A).$$

In order to ensure that $\mathrm{dist}(\mathcal{C}(\varepsilon),\rho(A+\beta B))>0$ we consider instead of $\mathcal{C}(\varepsilon)$ the set

$$\overline{B}(\mathcal{C}(\varepsilon),\delta)=\{\lambda\in\mathbb{C}\mid\mathrm{dist}(\mathcal{C}(\varepsilon),\lambda)\leq\delta\}$$

for $\delta<\varepsilon$. Since $\overline{B}(\mathcal{C}(\varepsilon),\delta)$ is compact we must have $\sup_{\lambda\in\overline{B}(\mathcal{C}(\varepsilon),\delta)}\left\|(A-\lambda)^{-1}\right\|=C_\delta<\infty$, so we can estimate

$$\left\|B(A-\lambda)^{-1}\right\|\leq a+C_\delta(a|\lambda|+b)\leq a+C_\delta(a(|\lambda_0|+\varepsilon+\delta)+b), \quad \lambda\in\overline{B}(\mathcal{C}(\varepsilon),\delta)$$

independently of λ so taking $|\beta|<(a+C_\delta(a(|\lambda_0|+\varepsilon+\delta)+b))^{-1}$ ensures that $\mathrm{dist}(\mathcal{C}(\varepsilon),\rho(A+\beta B))\geq\delta$. Noting that $C_\delta\leq C_{\delta'}$ if $\delta\leq\delta'$, we can take δ to 0 and set

$$r = \left(a + C_0\left(a\left(|\lambda_0| + \varepsilon\right) + b\right)\right)^{-1}$$

which ensures that $\operatorname{dist}\left(\mathcal{C}\left(\varepsilon\right), \sigma\left(A + \beta B\right)\right) > 0$ for any $\beta \in B\left(0, r\right)$.

The claim regarding self-adjoint A is due to the fact that by functional calculus $\left\|\left(A - \lambda\right)^{-1}\right\| = \operatorname{dist}\left(\sigma\left(A\right), \lambda\right)^{-1}$ so we know that $C_0 = \varepsilon^{-1}$ by definition of ε. □

Remark. Since $a^{-1} \geq r = \left(a + C_0\left(a\left(|\lambda_0|\right) + b\right)\right)^{-1}$ we still have analyticity of $A + \beta B$ on $B\left(0, r\right)$ in the self-adjoint case.

Before we can prove our main result on analyticity of the simple, non-degenerate eigenvalues of an analytic family $A + \beta B$ we need a result which will ensure us that we actually have eigenvectors. First, a preliminary lemma:

Lemma 1. *If $P, Q \in \mathcal{L}\left(\mathcal{H}\right)$ are two finite-rank projections and $\dim\left(\operatorname{Ran}\left(P\right)\right) \neq \dim\left(\operatorname{Ran}\left(Q\right)\right)$ then $\|P - Q\| \geq 1$.*

Proof. Assume $\dim\left(\operatorname{Ran}\left(P\right)\right) < \dim\left(\operatorname{Ran}\left(Q\right)\right)$. Then $P|_{\operatorname{Ran}(Q)} : \operatorname{Ran}\left(Q\right) \to \operatorname{Ran}\left(P\right)$ must have a non-trivial kernel by dimensionality. Let $x \in \operatorname{Ran}\left(Q\right)$ be a (non-zero) element of this kernel. Then

$$Px = 0, \quad Qx = x$$

hence $\|\left(P - Q\right)x\| = \|x\|$ so $\|P - Q\| \geq 1$. □

Lemma 2. *If $P\left(\beta\right) : U \to \mathcal{L}\left(\mathcal{H}\right)$ is a continuous projection-valued function of finite rank defined on a connected set $U \subset \mathbb{C}$ then $\dim\left(\operatorname{Ran}\left(P\left(\beta\right)\right)\right)$ is constant in β.*

Proof. Observe that given any $\beta' \in U$ we can by the above lemma find an open ball $B\left(\beta', r\right) \subset U$ such that $\dim\left(\operatorname{Ran}\left(P\left(\beta\right)\right)\right)$ is constant on $B\left(\beta', r\right)$ - indeed, by continuity we can choose $r > 0$ such that $|\beta' - \beta| < r$ implies $\|P\left(\beta'\right) - P\left(\beta\right)\| < 1$, hence by the above lemma $\dim\left(\operatorname{Ran}\left(P\left(\beta\right)\right)\right) = \dim\left(\operatorname{Ran}\left(P\left(\beta'\right)\right)\right)$.

Now choose some $\beta_0 \in U$ and define the set V by

$$V = \left\{\beta \in U \mid \dim\left(\operatorname{Ran}\left(P\left(\beta\right)\right)\right) = \dim\left(\operatorname{Ran}\left(P\left(\beta_0\right)\right)\right)\right\}.$$

On one hand V is open since for any $\beta \in V$ we can find an open ball containing β which is contained in V. By the same argument V is also closed, however, since given $\beta' \notin V$ we can also find a ball containing β' which is disjoint from V.

Since V is both open, closed and nonempty (clearly $\beta_0 \in V$) we conclude by connectedness that $V = U$, hence that $\dim\left(\operatorname{Ran}\left(P\left(\beta\right)\right)\right)$ is constant on U. □

Now we are ready for our main result:

Theorem 7. *Let $A + \beta B$ be an analytic family. Let E_0 be an isolated non-degenerate eigenvalue of A, i.e. $\operatorname{dist}\left(E_0, \sigma\left(A\right) \setminus \left\{E_0\right\}\right) = 2\varepsilon > 0$, with eigenvector φ_0. Then there exists a ball B_0 around 0 such that for all $\beta \in B_0$ the following holds:*

1. *There is exactly one point $E\left(\beta\right) \in \sigma\left(A + \beta B\right)$ in the ball $B\left(E_0, \varepsilon\right)$ and $E\left(\beta\right)$ is an isolated non-degenerate eigenvalue of $A + \beta B$.*
2. *There exists a neighbourhood $U \subset B_0$ around 0 such that $E\left(\beta\right) : U \to \mathbb{C}$ is analytic and there exists an analytic eigenvector $\varphi\left(\beta\right) : U \to \mathcal{H}$ such that $\left(A + \beta B\right)\varphi\left(\beta\right) = E\left(\beta\right)\varphi\left(\beta\right)$ holds for all $\beta \in U$.*

Furthermore, if A is self-adjoint then one can take $B_0 = B\left(0, \left(a + \left(a\left(|E_0| + \varepsilon\right) + b\right)\varepsilon^{-1}\right)^{-1}\right)$.

Proof. By Theorem 6 we can find a ball B_0 around 0 such that the contour $\mathcal{C} = \left\{\lambda \in \mathbb{C} \mid |E_0 - \lambda| = \varepsilon\right\}$ has a positive distance to $\sigma\left(A + \beta B\right)$ for all $\beta \in B_0$. By Theorem 4 we can then define the projection

$$P\left(\beta\right) = -\frac{1}{2\pi i}\oint_{\mathcal{C}}\left(A + \beta B - \lambda\right)^{-1} d\lambda, \quad \beta \in B_0.$$

If we take B_0 to be as in Theorem 6 we have the expansion

$$\left(A + \beta B - \lambda\right)^{-1} = \left(A - \lambda\right)^{-1}\sum_{k=0}^{\infty}\left(-1\right)^k\left(B\left(A - \lambda\right)^{-1}\right)^k\beta^k, \quad \beta \in B_0,$$

which can be obtained from the second resolvent identity by writing

$$(A+\beta B-\lambda)^{-1}=(A-\lambda)^{-1}-\beta\,(A+\beta B-\lambda)^{-1}\,B\,(A-\lambda)^{-1}$$

and then recursively inserting the right-hand side into itself, as in

$$\begin{aligned}(A+\beta B-\lambda)^{-1} &= (A-\lambda)^{-1}-\beta\left((A-\lambda)^{-1}-\beta\,(A+\beta B-\lambda)^{-1}\,B\,(A-\lambda)^{-1}\right)B\,(A-\lambda)^{-1}\\ &= (A-\lambda)^{-1}\left(1-\beta B\,(A-\lambda)^{-1}\right)+\beta^2\,(A+\beta B-\lambda)^{-1}\left(B\,(A-\lambda)^{-1}\right)^2\\ &= (A-\lambda)^{-1}\left(1-\beta B\,(A-\lambda)^{-1}+\beta^2\left(B\,(A-\lambda)^{-1}\right)^2\right)-\beta^3\,(A+\beta B-\lambda)^{-1}\left(B\,(A-\lambda)^{-1}\right)^3.\end{aligned}$$

We now see that $P(\beta)$ is in fact analytic in β, as we can write $P(\beta)$ as a power series in β (since the integration is over a compact set, there is no problem regarding convergence):

$$P(\beta)=-\frac{1}{2\pi i}\sum_{k=0}^{\infty}(-1)^k\left(\oint_{\mathcal{C}}(A-\lambda)^{-1}\left(B\,(A-\lambda)^{-1}\right)^k d\lambda\right)\beta^k.$$

Since E_0 is non-degenerate, $\dim(\operatorname{Ran}(P(0)))=1$, and so by Lemma 2 we in fact have $\dim(\operatorname{Ran}(P(\beta)))=1$. By Theorem 4 we conclude that there is exactly one spectral point $E(\beta)$ enclosed by $\mathcal{C}$ and that this is a an eigenvalue of $A+\beta B$ for $\beta\in B_0$. This shows (1).
Now define $\varphi(\beta)$ by

$$\varphi(\beta)=P(\beta)\,\varphi_0,\quad \beta\in B_0.$$

Since $P(\beta)$ is analytic on B_0 so is $\varphi(\beta)$. In particular then

$$\langle\varphi_0,\varphi(\beta)\rangle=\langle\varphi_0,P(\beta)\,\varphi_0\rangle$$

is an analytic function in β. Since $\langle\varphi_0,\varphi(0)\rangle=\langle\varphi_0,\varphi_0\rangle=\|\varphi_0\|^2\neq 0$ this implies that $\varphi(\beta)\neq 0$ for β on some neighbourhood U of 0.
Since $P(\beta)$ projects onto the $E(\beta)$-eigenspace of $A+\beta B$ we see that $\varphi(\beta):U\to\mathcal{H}$ is our desired analytic eigenvector.
What remains to be shown is that $E(\beta)$ is analytic on U. In the "usual" case of symmetric A and B this is particularly easy, since

$$E(\beta)=\frac{\langle\varphi_0,E(\beta)\,\varphi(\beta)\rangle}{\langle\varphi_0,\varphi(\beta)\rangle}=\frac{\langle\varphi_0,(A+\beta B)\,\varphi(\beta)\rangle}{\langle\varphi_0,\varphi(\beta)\rangle}=\frac{\langle A\varphi_0,\varphi(\beta)\rangle}{\langle\varphi_0,\varphi(\beta)\rangle}+\beta\frac{\langle B\varphi_0,\varphi(\beta)\rangle}{\langle\varphi_0,\varphi(\beta)\rangle}=E_0+\beta\frac{\langle B\varphi_0,\varphi(\beta)\rangle}{\langle\varphi_0,\varphi(\beta)\rangle},$$

and that the right-hand side is analytic is clear since $\varphi(\beta)$ is analytic and $\langle\varphi_0,\varphi(\beta)\rangle\neq 0$ for $\beta\in U$. In the general case, however, we have to formulate it in a somewhat awkward fashion by writing

$$(E(\beta)-E_0-\varepsilon)^{-1}=(E(\beta)-E_0-\varepsilon)^{-1}\,\frac{\langle\varphi_0,\varphi(\beta)\rangle}{\langle\varphi_0,\varphi(\beta)\rangle}=\frac{\left\langle\varphi_0,(A+\beta B-E_0-\varepsilon)^{-1}\,P(\beta)\,\varphi_0\right\rangle}{\langle\varphi_0,\varphi(\beta)\rangle}.$$

By considering power series it can be shown that the composition of two analytic operator-valued functions is also analytic, hence the right-hand side is analytic, and so by isolating $E(\beta)$ we conclude that this is also analytic (at least on some ball around 0 where this inversion is valid). □

Remark. As long as some care is taken regarding the root one can instead define $\varphi(\beta)=\langle\varphi_0,P(\beta)\,\varphi_0\rangle^{-\frac{1}{2}}\,P(\beta)\,\varphi_0$, which has the property that it is normalized for β such that $A+\beta B$ is self-adjoint.

References

[1] T. Kato, "Perturbation Theory for Linear Operators".

[2] M. Reed and B. Simon, "Methods of Modern Mathematical Physics IV: Analysis of Operators".

[3] N. Benedikter, "Lecture Notes for Advanced Mathematical Physics".

[4] J. Sok, "Lecture Notes for Advanced Mathematical Physics".