Lecture notes:
How do eigenvalues depend on a parameter?

Ben Schweizer

8. July 2025

1 Introduction

> Let $A = A(t)$ be a family of matrices, depending on a parameter t. Question: How do the eigenvalues of $A(t)$ depend on t?

This is the underlying question of what is called "perturbation theory". It is a fundamental question and it is of importance in many applications. F. Rellich made fundamental contributions in the 1940ies, T. Kato then extended the theory and published 1966 the first version of his 600 pages book, see [1]. Until today, this book is *the* reference in this field of research. Kato's theory is developed for the spectrum of linear operators on general Hilbert spaces, but the most interesting properties can already been explained in finite dimensional spaces.

Setting 1.1. *Let $(-\delta, \delta) \subset \mathbb{R}$ with $\delta > 0$ be a real interval, let $d \in \mathbb{N}$ be the dimension. Let $(-\delta, \delta) \ni t \mapsto A(t) \in \mathbb{C}^{d\times d}$ be a family of matrices. Let $\lambda_0 \in \mathbb{C}$ be an eigenvalue of $A(0)$ with algebraic multiplicity $m \in \mathbb{N}$.*

With coefficients, a matrix can be written as $A(t) = (a_{ij}(t))_{ij}$, we therefore consider the situation that the entries of the matrix depend on t. The guiding questions are: When the matrices A (that is: all coefficients a_{ij}) depend in an analytic way on t, do there exist m branches $t \mapsto \lambda_k(t)$ of eigenvalues, $k = 1, ..., m$, that continue λ_0? Are these branches analytic in t?

The first question is not too hard: When A depends differentiable on t, then there are m Lipschitz branches $\lambda_k = \lambda_k(t)$ with $\lambda_k(0) = \lambda_0$ and such that eigenfunctions span an m-dimensional space. This is a well-known fact and we recall it below. The second question is hard: Are the branches smooth?

We collect a number of examples that show: One has to be very careful! Interesting things can happen. Nevertheless, when everything is done right, there is a positive answer also to the second question.

We always assume that $\lambda_0 \in \mathbb{C}$ is an eigenvalue of $A(0)$ with the *algebraic* multiplicity $m \in \mathbb{N}$. Except for an introductory example, we will restrict ourselfs to selfadjoint matrices. We recall that, when a matrix is selfadjoint, all eigenvalues are real, eigenvectors to different eigenvectors are orhogonal (both facts follow with the elementary calculation $\lambda\langle v, w\rangle = \langle \lambda v, w\rangle = \langle Av, w\rangle = \langle v, Aw\rangle = \langle v, \mu w\rangle = \bar{\mu}\langle v, w\rangle$. Furthermore, the geometric multiplicity coincides with the algebraic multiplicity, the matrix is diagonizable (as can be shown by reducing the problem to the orthogonal complement of the eigenvectors that are already found; one eigenvector can always be found by maximizing $x \mapsto \langle Ax, x\rangle$ on spheres in $\mathbb{R}^d$).

1.1 Operators on Hilbert spaces

When one is interested in operators on Hilbert spaces, then one reduces to a finite-dimensional subspace of interest and applies the results for matrices (which are discussed in this text). Along this path one can show that the results presented here remain valid in the following situation, where C^ω stands either for C^1 or for "analytic".

Setting 1.2. *Let X be a Hilbert space over $\mathbb{C}$, let $A = A(t)$ be a family of operators $A \in C^\omega((-\delta,\delta), \mathcal{L}(X))$ with the following two properties for every t: (i) $A(t) : X \to X$ is selfadjoint. (ii) For every $\lambda \neq 0$, the operator $A(t) - \lambda\,\mathrm{id}$ is a Fredholm operator with index 0. One considers $\lambda_0 \neq 0$, an isolated eigenvalue of $A(0)$.*

Indeed, in the situation of Setting 1.2, there is a family of projections $\Pi(t)$ of class C^ω, where $\Pi(t)$ projects to the relevant m-dimensional subspace spanned by eigenvectors corresponding to continuations of the eigenvalue λ_0. The proof of closely ralated facts can be found in [1]. A proof is presented on the 4 pages of a (hopefully) well-accessible appendix in [3] (the appendix uses the ideas of Kato and presents the proof for C^1 dependence $t \mapsto A(t)$).

1.2 Lipschitz extensions

We indicated that it is easy to find Lipschitz extensions of the eigenvalue. We will use this fact which can be proved with complex analysis methods. This result is classical and (of course) contained in [1]; it is actually also obtained on the 4 pages of the appendix in [3].

Proposition 1.3 (Lipschitz branches for selfadjoint matrices)**.** *We consider Setting 1.1 with selfadjoint matrices $A(t)$ of class C^1, or the Setting 1.2. Then, for a possibly reduced parameter $\delta > 0$, there exist m Lipschitz continuous branches $\mu_1, ..., \mu_m : (-\delta,\delta) \to \mathbb{C}$ with $\mu_k(0) = \lambda_0$ (we allow $\mu_k(t) = \mu_\ell(t)$ for $k \neq \ell$) such that $\mu_k(t)$ is an eigenvalue for $A(t)$ for every $t \in (-\delta,\delta)$ and every $k \leq m$. Furthermore, corresponding eigenvectors $u_k(t)$ can be chosen such that $u_1(t), ..., u_m(t)$ spans an m-dimensional subspace. The projection $\Pi(t)$ to this subspace is of class C^1 in t.*

When all matrices $A(t)$ are selfadjoint, then all μ_k are real and we can order the m branches such that, for every t, there holds $\mu_1(t) \leq ... \leq \mu_m(t)$. We will see below that this is not necessarily a good ordering.

1.3 Simple eigenvalues

There is no problem at all when $\lambda_0 \in \mathbb{C}$ is an algebraically simple eigenvalue of $A(0)$. Then the eigenvalue can be extended with $(-\delta,\delta) \ni t \mapsto \lambda(t) \in \mathbb{C}$ with $\lambda(0) = \lambda_0$ and this map has the same smoothness as the coefficient functions. This can be shown with the implicit function theorem.

2 Examples with a lack of smooth dependence

In this section, we provide examples that show that one has to be very careful.

2.1 Non-selfadjoint matrix

Example 2.1 (Non-selfadjoint). *Matrix and eigenvalues are given by*

$$A(t) = \begin{bmatrix} 0 & t \\ 1 & 0 \end{bmatrix}, \qquad \lambda_1(t) = \sqrt{t}\,, \quad \lambda_2(t) = -\sqrt{t}\,. \tag{2.1}$$

We have chosen here the standard complex root such that $\sqrt{t}$ is the real square root for $t \geq 0$ and $\sqrt{t} = i\sqrt{-t}$ for $t < 0$.

The coefficients are analytic in t, the eigenvalues are continuous in t, but they are not Lipschitz continuous.

The example shows, in particular, that one cannot simply drop the assumption of $A(t)$ being self-adjoint in Proposition 1.3.

Kato's example II, 5.9 adds another aspect and improves the finding: The example there is a C^∞ family of matrices (again, not selfadjoint), depending on a single real parameter, such that the eigenvalues are not C^1 (in any ordering). So far, this is as in our example above. The interesting improvement of II, 5.9 is: For every t, the matrix is diagonizable. We see from the example: The failure of smooth dependence cannot be remedied by imposing, additionally, that the matrices are diagonizable.

We conclude: In the non-selfadjoint case, the behavior of eigenvalues is quite bad. They are only Lipschitz or, under additional assumptions, differentiable; in any case, the derivatives can be discontinuous. From now on, we therefore restrict ourselfs to selfadjoint families of matrices.

2.2 Selfadjoint matrix with bad natural ordering

Example 2.2 (Selfadjoint, bad natural ordering). *We consider*

$$A(t) = \begin{bmatrix} 0 & 0 \\ 0 & t \end{bmatrix}, \qquad \lambda_1 = 0\,, \quad \lambda_2 = t\,. \tag{2.2}$$

Let $\mu_1(t)$ and $\mu_2(t)$ be the two ordered eigenvalues, which means here

$$\mu_1(t) = \begin{cases} t & \text{for } t < 0 \\ 0 & \text{for } t \geq 0 \end{cases} \qquad \mu_2(t) = \begin{cases} 0 & \text{for } t < 0 \\ t & \text{for } t \geq 0 \end{cases} \tag{2.3}$$

Then $t \mapsto \mu_1(t)$ and $t \mapsto \mu_2(t)$ are both Lipschitz continuous, but not differentiable.

2.3 Excursion: Two parameters

In the next example, we use two real parameters, t and s. The characteristic polynomial is $(\lambda - s)(\lambda + s) - t^2 = \lambda^2 - s^2 - t^2$, the roots are $\lambda_{1,2} = \pm\sqrt{s^2 + t^2} = \pm|(s,t)|$, we interpret here $(s,t) \in \mathbb{R}^2$ as a point in the two-dimensional space.

Example 2.3 (Two parameters)**.** *Matrix and eigenvalues are given by*

$$A(s,t) = \begin{bmatrix} s & t \\ t & -s \end{bmatrix}, \qquad \lambda_{1,2} = \pm|(s,t)|\,. \tag{2.4}$$

When we try to achieve a smooth dependence on the parameters, we must choose the eigenvalues constant on circles, e.g., $\lambda_1(s,t) = |(s,t)|$ *and* $\lambda_2(s,t) = -|(s,t)|$*. With this choice, the map* $t \mapsto \lambda_1(0,t) = |t|$ *is Lipschitz continuous, but not differentiable.*

In Example 2.3, the coefficients are analytic in s and t, the eigenvalues are Lipschitz continuous, but not C^1. The example shows that no smooth dependence can be expected when two (or more) parameters are considered. Accordingly, in the following, we consider only the case with only one real parameter $t \in \mathbb{R}$.

2.4 Selfadjoint matrix with non-smooth eigenvectors

We now present the most interesting and most famous examples. The first is going back to Rellich, it appears on page 52 in [5] and as II, 5.3 in [1].

Example 2.4 (Selfadjoint, non-smooth eigenvectors)**.** *The matrices are*

$$A(t) = e^{-1/t^2} \begin{bmatrix} \cos(2/t) & \sin(2/t) \\ \sin(2/t) & -\cos(2/t) \end{bmatrix} \tag{2.5}$$

for $t \neq 0$*, extended with* $A(0) = 0$*. Two eigenvectors for* $t \neq 0$ *are*

$$\begin{pmatrix} \cos(1/t) \\ \sin(1/t) \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} \sin(1/t) \\ \cos(1/t) \end{pmatrix}, \tag{2.6}$$

the corresponding eigenvalues are $\pm e^{-1/t^2}$*. One easily checks this fact with the formulas* $\cos(2x) = \cos^2(x) - \sin^2(x)$ *and* $\sin(2x) = 2\sin(x)\cos(x)$*.*

The map $t \mapsto A(t)$ *is of class* C^∞*, the matrices are selfadjoint, but the eigen*vectors *do not depend continuously on* t*.*

In the above example, the eigenvectors are somehow "wildly spinning around" in the plane as $t \searrow 0$. The example shows that one cannot expect smoothness of the eigenvectors, not even Lipschitz-continuity.

We learned the following example from [2], it is a bit more accessible and shows the same point. The eigenvectors are not behaving quite as wild.

Example 2.5 (Selfadjoint, non-smooth eigenvectors)**.** *We choose the two matrices*

$$M_+ = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad M_- = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}. \tag{2.7}$$

These matrices are symmetric and have the property that the eigenvectors are different; one may choose $(1,1)$ *and* $(1,-1)$ *for* M_+ *and* $(1,0)$ *and* $(0,1)$ *for* M_-*.*

We consider a C^∞-function $g : \mathbb{R} \to \mathbb{R}$ with $g \geq 0$ and $g(0) = 0$, with all derivatives of g vanishing in $t = 0$, and with $g(t) \neq 0$ for $t \neq 0$. The matrix family of this example is given by

$$A(t) = \begin{cases} g(t)M_+ & \text{for } t \geq 0 \\ g(t)M_- & \text{for } t < 0\,. \end{cases} \tag{2.8}$$

*Since the eigenvectors are those of M_+ for $t > 0$ and those of M_- for $t < 0$, the eigen*vectors *do not depend continuously on t.*

Also in Example 2.5, the coefficients of A are C^∞ in t, but the eigenvectors are not even continuous.

3 Positive results: Smooth dependence

We have seen that we can expect a positive result only for eigenvalues and not for eigenvectors (see Example 2.4 or Example 2.5), only for a single real parameter (see Example 2.3), only for self-adjoint matrices (see Example 2.1), and only up to ordering (see Example 2.2). But — when we take all these restrictions into account — we find a positive answer. Some of the early formulations are: Page 39 of [5] and Theorem II, 6.1 of [1].

Theorem 3.1 (Rellich)**.** *Let $(-\delta, \delta) \ni t \mapsto A(t) \in \mathbb{C}^{d\times d}$ be an analytical family of selfadjoint matrices, let $\lambda_0 \in \mathbb{C}$ be an eigenvalue of $A(0)$ with multiplicity $m \in \mathbb{N}$.*

Then, for a possibly reduced parameter $\delta > 0$, there exist m analytic *branches $\lambda_1, ..., \lambda_m : (-\delta, \delta) \to \mathbb{C}$ (we allow $\lambda_k(t) = \lambda_\ell(t)$ for $k \neq \ell$) such that $\lambda_k(t)$ is an eigenvalue for $A(t)$ for every $t \in (-\delta, \delta)$ and every $k \leq m$. Furthermore, corresponding eigenvectors $u_k(t)$ can be chosen such that $u_1(t), ..., u_m(t)$ spans an m-dimensional subspace of $\mathbb{C}^d$.*

The following remark describes another theorem of Rellich. It allows that the dependence is only of class C^1.

Remark 3.2 (Variation: C^1-dependence)**.** *Theorem 3.1 remains valid with the following change: When the map $t \mapsto A(t)$ is only of class C^1, there are m branches $t \mapsto \lambda_k(t)$ of class C^1. See Theorem II, 6.8 of [1].*

Can one also have a smooth family of the corresponding eigenvectors? Astonishingly, the answer is positive in the case of analytic dependence.

Remark 3.3 (Smooth dependence of eigenvectors)**.** *In the setting of Theorem 3.1, one can also choose orhonormal eigenvectors $u_k(t)$ that depend analytically on t. See Theorem II, 6.1 of [1] and the comments in the subsequent section (starting on page 121).*

Remarks 3.2 and 3.3 show that one really has to be very careful: Regarding the eigenvalues, one can replace everywhere "analytic" by C^1 and gets the same result. For eigenvectors this is not the case, see Examples 2.4 and 2.5. In the light of these two examples, Remark 3.3 is astonishing. It can be read as: Demanding analyticity instead of C^∞ makes constructions as in the two examples impossible.

3.1 Help from algebra: Puiseux's theorem

The proof of Theorem 3.1 is based on a result from algebra.

Theorem 3.4 (Newton-Puiseux)**.** *Let $\delta > 0$ and $n \in \mathbb{N}$ be fixed and let $U \subset \mathbb{C}$ be an open set containing the real interval $[-\delta, \delta]$. Let $\Phi : U \times \mathbb{C} \ni (t, \lambda) \mapsto \Phi(t, \lambda) \in \mathbb{C}$ be of the form*

$$\Phi(t, \lambda) = \lambda^n + \varphi_{n-1}(t)\lambda^{n-1} + ... + \varphi_0(t)\,, \tag{3.1}$$

where, for every $k < n$, the function $\varphi_k : U \to \mathbb{C}$ is an analytic function. Then, for some number $r \in \mathbb{N}$, on a possibly smaller open set $0 \in U' \subset U$, there exist n analytic functions $g_k : U' \to \mathbb{C}$ such that

$$\Phi(t, \lambda) = \Pi_{k=1}^{n}(\lambda - g_k(t^{1/r}))\,. \tag{3.2}$$

In particular, the n roots of Φ can be written as analytic functions in $t^{1/r}$. Equation (3.2) *holds for every $t \in U \setminus \{z \in \mathbb{C} \,|\, \mathrm{Re}(z) \leq 0\,, \mathrm{Im}(z) = 0\}$ and for every $\lambda \in \mathbb{C}$.*

3.2 Rellich's proof of analytic dependence

We can now give the proof for smooth dependence of eigenvalues. We present here the proof that Rellich gives in his lecture notes [5], starting on page 37.

The essence of the proof is: The characteristic polynomial of $A(t)$ is a polynomial of the form (3.1). Algebra helps and we can conclude that the zeros of the polynomial (which are the eigenvalues) depend analytically on t.

The proof is not that simple! The characteristic polynomial of (2.1) is $\lambda^2 - t$, the roots are $\pm\sqrt{t}$, they are not smooth! In fact: Theorem 3.4 only yields that the zeros of the polynomial depend analytically on complex roots of t, not necessarily analytically on t.

We must exploit that $A(t)$ is selfadjoint in order to conclude the proof.

Proof of Theorem 3.1. We study the characteristic polynomial $\Phi(t, \lambda)$ of the matrix family, it is of the form (3.1) with analytical coefficients, there holds $\Phi(0, \lambda_0) = 0$. By Theorem 3.4, we can write Φ as in (3.2). For fixed index k, we consider the family of zeros $\lambda_k(t) = g_k(t^{1/r})$. Because of $g_k(0) = \lambda_0$, we can write

$$\lambda_k(t) = \lambda_0 + b_1\, t^{1/r} + b_2\, t^{2/r} + b_3\, t^{3/r} + ... \tag{3.3}$$

where the series converges in a neighborhood of $t = 0$.

We note that λ_0 and all $\lambda_k(t)$ are real, since these are eigenvalues of a selfadoint matrix. Considering only real and positive numbers t, we can conclude that all coefficients b_ℓ must be real. This can be concluded with an induction argument. We sketch this as follows: When b_1 is real, then b_2 is the limit of the real numbers $t^{-2/r}\left(\lambda_k(t) - \lambda_0 - b_1\, t^{1/r}\right)$ for $t \searrow 0$.

Let us assume that there is any index $\ell \in \mathbb{N}$, with a non-vanishing coefficient $b_\ell \neq 0$, such that the corresponding exponent is non-integer, $\ell/r \notin \mathbb{N}$. Under this assumption, let ℓ be the smallest index with these properties. Then, for t real and small in absolute value, the number

$$\lambda_0 + \left(\sum_{j \text{ with } jr<\ell} b_{jr}\; t^j\right) + b_\ell\; t^{\ell/r} \tag{3.4}$$

is a good approximation to the real number $\lambda_k(t)$ in the sense that

$$\frac{1}{|t|^{\ell/r}} \left[\lambda_k(t) - \lambda_0 - \left(\sum_{j \text{ with } jr<\ell} b_{jr}\, t^j \right) - b_\ell\, t^{\ell/r} \right] \to 0 \tag{3.5}$$

as $|t| \to 0$. In particular, the imaginary part of the left-hand side vanishes in the limit. Only the last entry is non-real, therefore the imaginary part of the left-hand side is simply $\mathrm{Im}(b_\ell\,(t/|t|)^{\ell/r})$. When we consider $t > 0$, this expression is just $\mathrm{Im}(b_\ell)$ and (3.5) yields $\mathrm{Im}(b_\ell) = 0$. When we consider $t < 0$, then $(t/|t|)^{\ell/r}$ is not real and we conclude from (3.5) that also $\mathrm{Re}(b_\ell)$ vanishes. This is in contradiction to the choice of ℓ. □

References

[1] T. Kato. *Perturbation Theory for Linear Operators*, volume 132 of *Grundlehren der mathematischen Wissenschaften*. Springer-Verlag, Berlin, New York, 1976.

[2] J. L. Kazdan. Matrices $A(t)$ depending on a parameter t. pages 1–5, 2010+.

[3] A. Kirsch and B. Schweizer. Spectral gap properties of perturbed periodic media. Technical report, 2025.

[4] K. Nowak. Some elementary proofs of Puiseux's theorems. *European Mathematical Information Service*, 2000.

[5] F. Rellich. *Perturbation Theory of Eigenvalue Problems*. Gordon and Breach, New York, 1968. Originally published as New York University Lecture Notes, 1953.