Introduction to Mathematical Physics

Fall 2010: Spectral Theory

Main references:

[RS1] M. Reed, B. Simon: Methods of Modern Mathematical Physics I: Functional Analysis (1980)

[Rudin, FA] W. Rudin: Functional Analysis (1991)

[Kato] T. Kato: Perturbation Theory for Linear Operators (2008)

For background material (measure theory and complex analysis):

[Rudin, RCA] W. Rudin: Real and Complex Analysis (1987)

* Course announcements & further references from the course homepage

* Lectures: Tue & Thu 14-16
  Exercises: Fri 10-12

* The course can be taken for 10 cu.
  Grade consists of $\begin{cases} 50\% & \text{Final exam} \\ 50\% & \text{Exercises} \end{cases}$

* Please remember to register in OODI if you wish to take the course.

1. Introduction: Matrices ($\mathbb{C}^{d\times d} \cong \mathcal{B}(\mathbb{C}^d)$)

Recap of basic notions of linear algebra

* $\mathbb{C}^d$, $d \in \mathbb{N}_+$, is a complex vector space, with a scalar product $a \cdot b := \sum_{i=1}^{d} a_i^* b_i$

  ($a^* = \mathrm{Re}\, a - i \,\mathrm{Im}\, a$ = complex conjugate of a. Note that here the first argument is antilinear (physics convention).)

* The length of a vector $a \in \mathbb{C}^d$ is denoted by $|a| := \sqrt{\sum_{i=1}^{d} |a_i|^2}$.

* "Hat" will be used to denote the direction of $a \in \mathbb{C}^d$:

$$\hat{a} := \frac{a}{|a|}$$

* Consistently, we use "hat" also if we wish to emphasise unit vectors (i.e. $\hat{a} \in \mathbb{C}^d \Rightarrow |\hat{a}| = 1$)

* Same notations are used for $\mathbb{R}^d \subset \mathbb{C}^d$. (Then $a, b \in \mathbb{R}^d \Rightarrow a \cdot b = \sum_{i=1}^{d} a_i b_i$)

* Special notation $\hat{e}_\nu$, $\nu = 1, 2, \dots, d$, is reserved for the Cartesian basis of $\mathbb{C}^d$:

$$(\hat{e}_\nu)_i := \mathbb{1}(i = \nu) \; ; \quad i, \nu = 1, 2, \dots, d$$

* Here $\mathbb{1}(P) = \begin{cases} 1, & \text{if } P \text{ is true} \\ 0, & \text{otherwise} \end{cases}$

* Kronecker delta: $\delta_{ij} = \mathbb{1}(i = j)$ for $i, j \in \mathbb{Z}^d$.

* $\{\hat{e}_\nu\}_{\nu=1}^{d}$ forms an orthonormal basis for $\mathbb{C}^d$:

  a) $\mathbb{C}^d = \mathrm{span}\{\hat{e}_\nu \mid \nu = 1, .., d\}$
  b) $\hat{e}_{\nu'} \cdot \hat{e}_\nu = \delta_{\nu'\nu}$, $\forall \nu', \nu = 1, .., d$

* A matrix is a "vector" in $\mathbb{C}^{d_1 \times d_2} := (\mathbb{C}^{d_2})^{d_1}$
  $(\cong \mathbb{C}^{d_1 d_2})$

(Notational reminder:

* $X^I := \{ f : I \to X \}$, for any sets I, X; I is called the index set, $x_i := f(i)$.
* $X^n := X^{\{1,2,\dots,n\}}$, for $n \in \mathbb{N}_+$.
* If $M \in (X^J)^I$, $M_{ij} := (M_i)_j \in X$)
  $\Rightarrow (X^J)^I \cong X^{I \times J}$

* Row-column shorthand notation of matrices: $M = \begin{pmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \end{pmatrix} \in \mathbb{C}^{2\times 3}$

* Matrices can be identified with linear maps: if $T \in \mathcal{B}(\mathbb{C}^{d_2}, \mathbb{C}^{d_1})$ and $M \in \mathbb{C}^{d_1 \times d_2}$

$$T \cong M \quad \text{if} \quad M_{ij} = \hat{e}_i \cdot T\hat{e}_j \quad \forall\, i = 1, .., d_1,\ j = 1, .., d_2$$

$$\Rightarrow a \cdot Tb = \sum_{i=1}^{d_1} \sum_{j=1}^{d_2} a_i^* M_{ij} b_j \quad \forall a \in \mathbb{C}^{d_1}, b \in \mathbb{C}^{d_2}.$$

* Here $\mathcal{B}(X, Y) := \{ T : X \to Y \mid T \text{ linear and } \|T\| < \infty \}$

where $\|T\| := \sup_{a \in X,\, |a| \leq 1} |Ta|$, X, Y are normed spaces.

* This induces the matrix norm on $\mathbb{C}^{d_1 \times d_2}$

$$\|M\| := \sup_{\substack{\hat{a} \in \mathbb{C}^{d_1}, \\ \hat{b} \in \mathbb{C}^{d_2};\ |a| = 1 = |b|}} \Big| \sum_{i,j} a_i^* M_{ij} b_j \Big| = \max_{b \in \mathbb{C}^{d_2},\, b \neq 0} \Big| \frac{1}{|b|} \sum_{j=1}^{d_2} M_{\bullet j} b_j \Big|$$

* Matrix product: $A \in \mathbb{C}^{d_1 \times d_2}$, $B \in \mathbb{C}^{d_2 \times d_3}$ then $AB \in \mathbb{C}^{d_1 \times d_3}$ defined by $(AB)_{ij} := \sum_{\nu=1}^{d_2} A_{i\nu} B_{\nu j}$.

$\Rightarrow AB \cong A \circ B$. Also $\|AB\| \leq \|A\| \|B\|$.

* The matrix norm is (numerically) different from the natural norm of $\mathbb{C}^{d_1 d_2}$;

$$\|M\|_{HS} := \sqrt{\sum_{i,j} |M_{ij}|^2} = \sqrt{\mathrm{Tr}(M^+M)} = \sqrt{\mathrm{Tr}(MM^+)}$$

where $(M^+)_{ij} := (M_{ji})^*$ defines the

<u>adjoint matrix</u> of M. ($M^+ \in \mathbb{C}^{d_2 \times d_1}$)

and $\mathrm{Tr}\, B := \sum_{i=1}^{d} B_{ii}$ is the <u>trace</u>

of a <u>square matrix</u> $B \in \mathbb{C}^{d \times d}$.

$\triangleright$ (*)

* Notation: $\mathcal{B}(\mathfrak{X}) := \mathcal{B}(\mathfrak{X}, \mathfrak{X})$
$\Rightarrow$ $\mathcal{B}(\mathbb{C}^d) = \mathbb{C}^{d \times d}$ = set of square matrices. = set of <u>operators on $\mathbb{C}^d$</u>.

(*) Proof: $\left\|\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}\right\| = 2 < \sqrt{5} = \left\|\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}\right\|_{HS}$.

* $\|M\|_{HS}$ is called the <u>Hilbert-Schmidt norm</u> of M.

---

* The main aim of the course is to generalize the following matrix result to operators on infinite-dimensional Hilbert-spaces:

<u>Theorem</u> (diagonalization of self-adjoint matrices)

Suppose $A \in \mathcal{B}(\mathbb{C}^d)$ satisfies $A^+ = A$. Then there is a unitary matrix $U \in \mathcal{B}(\mathbb{C}^d)$ and a <u>real diagonal</u> matrix $D \in \mathcal{B}(\mathbb{C}^d)$

such that $A = U^+ D U$.

Consequence: Let $\lambda_i = R_{ii} \in \mathbb{R}$, $\Rightarrow$ $(R^n)_{ij} = \lambda_i^n \delta_{ij}$,

$(e^{zR})_{ij} = e^{z\lambda_i} \delta_{ij}$ & $A^n = U^+ R^n U$, $e^{zA} = U^+ e^{zR} U$ $\forall z \in \mathbb{C}$.

(Matrix exponential will be defined in the exercises.)

* U unitary means $U \in \mathbb{C}^{d \times d}$ and $U^{-1} = U^+$; they are important since they generate Hilbert space isomorphisms: $Ua \cdot Ub = a \cdot U^+ U b = a \cdot b$.

* Linear isomorphism is called a change of basis and they are in 1-1 correspondence with invertible square matrices: If $\{e_i\}$ are the old and $\{e_i'\}$ the new basis vectors, then $\exists L \in \mathbb{C}^{d \times d}$ s.t. $e_i' = L e_i$ $\forall i$ and $\exists L^{-1}$.

* A change of basis L induces the following change to an operator $B \in \mathbb{C}^{d \times d}$: $B \mapsto L B L^{-1}$.

* Thus the theorem can also be summarized by saying that to every self-adjoint matrix, there is a unitary change of basis such that it becomes a real multiplication operator. ($D = UAU^{-1}$)

* $B \in \mathcal{B}(\mathbb{C}^d)$ is a multiplication operator, if there is $\vartheta : \mathbb{C}^d \to \mathbb{C}$ such that $(Ba)_i = \vartheta_i a_i$ $\forall a \in \mathbb{C}^d$ and $i = 1, .., d$. Thus multiplication operators are in 1-1 correspondence with diagonal matrices.

* In quantum mechanics, time-evolution of a wavevector $\psi_0 \in \mathcal{H}$ = Hilbert-space is defined by giving a self-adjoint operator H and defining time-evolved wavevectors by

$$\psi_t := e^{-itH} \psi_0 , \quad t \in \mathbb{R}.$$

"Diagonalization" of H will vastly simplify the analysis of properties of ψ_t.

Before going any further, let us recall another useful property of matrices which does <u>not</u> generalize to operators on infinite-dim. Hilbert spaces:

1.1. Jordan decomposition of finite-dim. matrices

a) Differentiation and integration of matrices

Consider a map $t \mapsto M(t)$ which maps an interval $(t_1, t_2) \subset \mathbb{R}$ to $\mathbb{C}^{d_1 \times d_2}$. Is there any natural way to define differentiation and integration of such maps?

* One method is to define it "component-wise" that is define $(M'(t_0))_{ij} = \frac{d}{dt} M(t)_{ij} \big|_{t=t_0}$

  and $\left(\int_{t_1}^{t_2} dt\, M(t) \right)_{ij} = \int_{t_1}^{t_2} dt\, M(t)_{ij}$

  whenever $M(t)_{ij}$ is differentiable (integrable) for all i,j.

* For integrals, we use the component-wise definition. This will generalize to what is often called "vector valued integration" in topological vector spaces.

* For derivatives, it is more convenient to think of approximations by linear maps, i.e., of <u>linearizations</u>.

<u>Definition</u> $t \mapsto M(t)$ is said to be <u>differentiable</u> at $t_0 \in (t_1, t_2)$ if $\exists M_0 \in \mathbb{C}^{d_1 \times d_2}$ such that

$$\lim_{h \to 0} \frac{1}{|h|} \| M(t_0 + h) - M(t_0) - M_0 h \| = 0.$$

Then M_0 is called the <u>derivate at t_0</u>, which we denote by $M'(t_0)$ or $\frac{d}{dt} M(t)\big|_{t=t_0}$.

* If $M'(t_0)$ exists, it is obviously unique.

b) Block matrices
= convenient shorthand notation

$M = (B_{ij})_{\substack{i=1,\dots,N_1 \\ j=1,\dots,N_2}} \in \mathbb{C}^{\bar{d}_1 \times \bar{d}_2}$ and each

B_{ij} is a matrix, $\in \mathbb{C}^{d_{1i} \times d_{2j}}$;

where the dimensions have to "match", i.e., satisfy

$$\sum_{i=1}^{N_1} d_{1i} = \bar{d}_1 \quad \text{and} \quad \sum_{j=1}^{N_2} d_{2j} = \bar{d}_2 .$$

Idea is that M acts as "matrix of matrices",

"$(Ma)_i = \sum_{j=1}^{N_2} B_{ij} a_j$", which means

explicitly that $\forall\, k' = 1, \dots, \bar{d}_1;\ k = 1, \dots, \bar{d}_2$

$$M_{k'k} = (B_{ij})_{l'l} \quad \text{for the } \underline{\text{unique}}$$

$i \in \{1, \dots, N_1\}$, $j \in \{1, \dots, N_2\}$, $l' \in \{1, \dots, d_{1i}\}$
$l \in \{1, \dots, d_{2j}\}$ such that

$$k' = l' + \sum_{i'=1}^{i-1} d_{1i'} \quad \text{and} \quad k = l + \sum_{j'=1}^{j-1} d_{2j'}$$

Example: If $A \in \mathbb{C}^{d_1 \times d_1}$, $B \in \mathbb{C}^{d_2 \times d_2}$, and $C \in \mathbb{C}^{d_1 \times d_2}$ then

$$M = \begin{pmatrix} A & C \\ C^+ & B \end{pmatrix} \in \mathbb{C}^{(d_1+d_2) \times (d_1+d_2)}$$

c) Nilpotent matrices

* If M is diagonalizable $\Rightarrow \forall n \geq 1: M^n = L R^n L^{-1}$.
  Thus $M^n = 0 \Rightarrow R^n = 0 \Rightarrow R = 0 \Rightarrow M = 0$.

Definition If $M \in \mathbb{C}^{d\times d}$ and there is $n \in \mathbb{N}_+$ such that $M^n = 0$, M is called nilpotent.

* Since, for instance, $M = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ satisfies $M^2 = 0$, there are non-zero nilpotent matrices $\Rightarrow$
  Not every matrix is diagonalizable!

* The following theorem yields a characterization of nilpotent matrices:

Theorem: Let $M \in \mathbb{C}^{d\times d}$ be a nilpotent matrix. Then there is a change of basis $L \in \mathbb{C}^{d\times d}$ which brings it into a block-diagonal form, where each diagonal block is a simple Jordan block matrix.

Explicitly: $\exists L$ s.t. $M = L^{-1} J L$

where
$$J = \begin{pmatrix} J_1 & 0 & 0 & \cdots & 0 \\ 0 & J_2 & 0 & \cdots & 0 \\ 0 & 0 & J_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & J_N \end{pmatrix}, \quad N \geq 1,$$

and for each $n = 1, \dots, N$, $\exists d_n \geq 1$, so that

a) $\displaystyle\sum_{n=1}^{N} d_n = d$

b) $(J_n a)_i = a_{i+1}, \quad i = 1, \dots, d_n - 1$
$(J_n a)_{d_n} = 0$

that is,
$$J_n = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \vdots & & \ddots & 0 \\ & & & & \ddots & 1 \\ 0 & 0 & 0 & \cdots & \cdots & 0 \end{pmatrix}$$

* Note that $(0) \in \mathbb{C}^{1\times 1}$ is also a simple Jordan block matrix.

* A convenient formula for the matrix elements of a simple Jordan block matrix $J \in \mathbb{C}^{d\times d}$ is given by

$$J_{ij} = \mathbb{1}(j = i+1) \quad \forall i, j = 1, \dots, d.$$

* <u>Notations</u>: For $M \in \mathbb{C}^{d_1 \times d_2} \equiv \mathcal{B}(\mathbb{C}^{d_2}, \mathbb{C}^{d_1})$

we denote its <u>range</u> by $R(M) := \{Ma \mid a \in \mathbb{C}^{d_2}\}$

and <u>kernel or null space</u> by

$\ker(M) := \{a \in \mathbb{C}^{d_2} \mid Ma = 0\}$.

<u>Reminders</u>: $R(M)$ is a subspace of $\mathbb{C}^{d_1}$
$\ker(M)$ is a subspace of $\mathbb{C}^{d_2}$

<u>Proof of the Theorem</u>:

Since M is nilpotent there is

$$m = \min\{n \geq 0 \mid M^{n+1} = 0\}.$$

If $m = 0 \Rightarrow M^1 = M = 0$, and we can choose $L = 1$ (= identity matrix), $N = d$, $d_n = 1 \ \forall n$ and $J_n = (0) \ \forall n$. $(\Rightarrow J = 0)$.

Assume thus that $m > 0 \Rightarrow M^m \neq 0$. Each $R(M^n)$ is a subspace, and let $r_n = \dim R(M^n)$. Since obviously $R(M^{n+1}) \subset R(M^n)$, we have now $d = r_0 \geq r_1 \geq \dots \geq r_m > 0 = r_n \quad \forall n > m$.

Then $R(M^m)$ has a basis $\{v_j^{(m)}\}_{j=1}^{r_m}$.
For each $v_j^{(m)}$ we choose $e_j^{(m)} \in \mathbb{C}^d$ s.t.
$v_j^{(m)} = M^m e_j^{(m)}$.

Define then $N_m := r_m$ and set

$$e^{(m)}_{j,n} := M^n e^{(m)}_j \quad \text{for } j = 1, \dots, N_m \,;\ n \geq 0.$$

By definition, $e^{(m)}_{j,n} = 0 \ \forall n > m$ and

$(e^{(m)}_{j,m})$ is a basis for $R(M^m)$.

(As an aside: the vectors $e^{(m)}_{j,n}$; $j = 1, \dots, N_m$; $n = 0, \dots, m$

are all linearly independent.

(<u>Pf.</u> if $\sum_{j,n} \alpha_{j,n} e^{(m)}_{j,n} = 0 \Rightarrow \forall n' \geq 0$:

$$0 = \sum_{j,n} \alpha_{j,n} M^{n'+n} e^{(m)}_j \overset{n''=n'+n}{=} \sum_{j=1}^{N_m} \sum_{n''=n'}^{m} \alpha_{j,n''-n'} M^{n''} e^{(m)}_j$$

Thus $n' = m \Rightarrow \alpha_{j,0} = 0 \ \forall j$. Then an induction in n shows that $\alpha_{j,n} = 0 \ \forall j, n$ $\square$))

(use $n' = m - n$)

We now claim that $\forall n = 0, \dots, m$ $\exists N_n \geq 0$ and vectors $e^{(n)}_j \in \mathbb{C}^d$, $j = 1, \dots, N_n$, s.t. with $e^{(n)}_{j,n'} := M^{n'} e^{(n)}_j$, $n' \geq 0$, $e^{(n)}_{j,n'} = 0$ for $n' > n$

and $(e^{(k)}_{j,k'})_{\substack{k=n,\dots,m \\ k'=n,\dots,k \\ j=1,\dots,N_k}}$ is a basis for $R(M^n)$.

The construction can be done iteratively starting from $n = m$ down until $n = 0$ is reached. Assume the claim is true for values up to $n+1$. If

$$\sum_{k=n+1}^{m} \sum_{k'=n}^{k} \sum_{j=1}^{N_k} \alpha_{j,k,k'} M^{k'} e^{(k)}_j = 0$$

By applying M, and using the induction assumption, $\Rightarrow \alpha_{j,k,k'} = 0 \ \forall k' < k$.

$\Rightarrow$

$$0 = \sum_{k=n+1}^{m} \sum_{j=1}^{N_k} \alpha_{j,k,k} \underbrace{M^k e^{(k)}_j}_{\in R(M^{n+1})} \Rightarrow \alpha_{j,k,k} = 0.$$

Thus $e^{(k)}_{j,k'}$, $k = n+1, \dots, m$; $k' = n, \dots, k$; $j = 1, \dots, N_k$ form a lin. indep. subset of $R(M^n)$

Let $N_n = r_n - \sum_{k=n+1}^{m} (k+1-n) N_k \geq 0$.

If $N_n = 0$, the previous vectors already span $R(M^n)$, and the statement holds for n.

If $N_n > 0$, we can add new vectors $\tilde{v}_j^{(n)}$, $j = 1, \dots, N_n$, to complete the set of vectors into a basis for $R(M^n)$. Then $M\tilde{v}_j^{(n)} \in R(M^{n+1})$ and thus $\exists \alpha_j(j', k', k)$ s.t.

$$M\tilde{v}_j^{(n)} = \sum_{k=n+1}^{m} \sum_{k'=n+1}^{k} \sum_{j'=1}^{N_k} \alpha_j(j', k', k) \underbrace{M^{k'}}_{k''=k'-1} e_{j'}^{(k)}$$

$= M r_j^{(n)}$ where

$$r_j^{(n)} = \sum_{k=n+1}^{m} \sum_{k''=n}^{k-1} \sum_{j'=1}^{N_k} \alpha_j(j', k''+1, k) M^{k''} e_{j'}^{(k)} \in R(M^n).$$

We define $v_j^{(n)} = \tilde{v}_j^{(n)} - r_j^{(n)} \in R(M^n) \Rightarrow M v_j^{(n)} = 0$. It is easy to check that also adding the vectors $v_j^{(n)}$, $j = 1, \dots, N_n$, yields a basis for $R(M^n)$. Then we choose $e_j^{(n)}$ s.t. $v_j^{(n)} = M^n e_j^{(n)}$. This completes the induction step.

Now $\left(e_{j,k'}^{(k)}\right)_{\substack{k=0,\dots,m \\ k'=0,\dots,k \\ j=1,\dots,N_k}}$ is a basis for $R(M^0) = \mathbb{C}^d$, and thus there is invertible $L \in \mathbb{C}^{d\times d}$ s.t. $L^{-1}\hat{e}_i = e_{j,k'}^{(k)}$ where we go through the indices

by starting from $k = m$, $j = 1$, $k' = k$, then decreasing k' until 0 is reached, after which $j = 2$ (if $N_m > 1$) and k' again runs from k to 0. After $j = N_m$ has been completed, we decrease k, but skip over values for which $N_k = 0$. Since $M e_{j,k'}^{(k)} = e_{j,k'+1}^{(k)}$ inside each block of $k' = 0, \dots, k$,

we clearly have $L M L^{-1} \hat{e}_i = J \hat{e}_i \ \forall i$, where J has N_k blocks of size $k+1$, $k = 0, \dots, m$.

$\Rightarrow N = \sum_{k=0}^{m} N_k$. $\square$

d) Jordan normal form

Consider an arbitrary $M \in \mathbb{C}^{d\times d}$.

* $\lambda \in \mathbb{C}$ is an eigenvalue of M if there is $a \in \mathbb{C}^d$ s.t.

$$Ma = \lambda a.$$

* $\sigma(M) = \{\lambda \in \mathbb{C} \mid \lambda \text{ is an eigenvalue of } M\}$

* "known": a) $1 \leq |\sigma(M)| \leq d$.
  b) $\exists R_\lambda := (\lambda 1 - M)^{-1}$ iff. $\lambda \notin \sigma(M)$.

Theorem For any $M \in \mathbb{C}^{d\times d}$ we can find a change basis $L \in \mathbb{C}^{d\times d}$ which brings it into a Jordan normal form.

Explicitly: $\exists L$ s.t. $M = L^{-1} J L$,

$$J = \text{Block-diagonal}(J_1, J_2, \dots, J_N)$$

and each J_n is a Jordan block matrix: $J_n = \lambda_n 1 + S_n$, S_n = Simple Jordan block. Here $\lambda_n \in \sigma(M)$ and each eigenvalue appears at least once.

Proof. Maybe later...

* This result does not generalize to $\mathcal{B}(\mathcal{H})$ if $\dim \mathcal{H} = \infty$.

* $$J_n = \begin{pmatrix} \lambda_n & 1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_n & 1 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_n & 1 & \ddots & \vdots \\ \vdots & \vdots & & \ddots & \ddots & 0 \\ & & & & \ddots & 1 \\ 0 & 0 & \cdots & & & \lambda_n \end{pmatrix}$$

* The different "blocks" can be expressed in terms of integrals over the resolvent:

$$M = \sum_{\lambda \in \sigma(M)} (\lambda P_\lambda + T_\lambda)$$

where all $(P_\lambda, T_\lambda)_\lambda$ commute, P_λ is a projection and T_λ is nilpotent.

In addition,

$$P_\lambda = \oint_{\Gamma_\lambda} \frac{dz}{2\pi i} R_z \quad \text{and} \quad T_\lambda = \oint_{\Gamma_\lambda} \frac{dz}{2\pi i} (z-\lambda) R_z$$

where Γ_λ denotes a path in $\mathbb{C}$ which goes once around λ and avoids all other points in the spectrum.