Vector and Function Spaces, example questions solutions

Tom Burgelman, Phileas Massün, Arthur Vanhaverbeke et Al.

January 2022 - August 2022

Please note, though this document has been made with a good amount of care, discussion and revision. It is more than possible that some mistakes slipped through the cracks, we are only students after all. If at any point you believe to have found a mistake, please do report it in #vector-en-functieruimten in the Fysica bois discord: `https://discord.gg/5TqJDXkeEp`

1 Linear algebra basics

Theorem 1.1. *Let $\hat{P} \in End(V)$ be a linear operator on a vector space V that satisfies $\hat{P}^2 = \hat{P}$. Show that $V = im(\hat{P}) \bigoplus ker(\hat{P})$.*

Proof. Choose a $v \in V$. Since P = P^2, we have $Pv = P^2v$ and thus $P(v - Pv) = 0$. Therefore $v - Pv = x$ for some $x \in ker(P)$. Thus $v = Pv + x$. This shows that V = im(P) + ker(P). Now take a $w \in im(P) \cap ker(P)$. Since $w \in im(P)$ we have $w = Pz$ for some $z \in V$. Applying P to both sides we get $Pw = P^2z$. But since $w \in ker(P)$, we have $0 = Pw = P^2z = Pz = w$. Thus $im(P) \cap ker(P) = 0$ and thus V = $im(P) \bigoplus ker(P)$. □

Theorem 1.2. *Verify the block LDU decomposition by working out the block matrix multiplication; then use this decomposition to give an expression for det(A).*

Proof. For this section look at the LDU decompostion in the notes (p49 in alt course) (you will be given the expression on the exam). Verifying the LDU decomposition is easy (just do the block matrix multiplication). We introduce the Shur complement of $A_{11} = A_{22} - A_{21}A_{11}^{-1}A_{12}$, sometimes denoted as A/A_{11}. The Shur complement can now be used to calculate det(A). We have det(A) = det(L)det(D)det(U) (note that: det(L) = det(U) = 1) = det(A_{11}) det(A/A_{11}) [= det(A_{22}) det(A/A_{22}). This is not necessary, but is obtained by permuting the blocks 1 and 2. You are asked to give AN expression for the determinant, so the first expression is perfectly valid] . □

Theorem 1.3. *Use the block LDU decomposition from the previous question to find a block matrix expression for A^{-1}.*

Proof. To find A^{-1} we calculate $A^{-1} = U^{-1}D^{-1}L^{-1}$. (page 46 in the course notes version 9).

$$
\begin{aligned}
A^{-1} &= U^{-1}D^{-1}L^{-1} \\
&= \begin{bmatrix} I_{n_1} & -A_{11}^{-1}A_{12} \\ 0 & I_{n_2} \end{bmatrix} \begin{bmatrix} A_{11}^{-1} & 0 \\ 0 & (A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1} \end{bmatrix} \begin{bmatrix} I_{n_2} & 0 \\ -A_{21}A_{11}^{-1} & I_{n_2} \end{bmatrix} \\
&= \begin{bmatrix} A_{11}^{-1} + A_{11}^{-1}A_{12}(A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1}A_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}(A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1} \\ -(A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1}A_{21}A_{11}^{-1} & (A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1} \end{bmatrix}
\end{aligned}
$$

□

Theorem 1.4. *Prove Woodbury's inversion lemma* $(A + UCV)^{-1} = A^{-1} - A^{-1}U(C^{-1} + VA^{-1}U)^{-1}VA^{-1}$ *by multiplying both sides with* $(A + UCV)$ *and show that you indeed obtain an equality. Ensure that the steps and the arithmetic you use on the right hand side are clear.*

Proof. It's easy to see that the LHS is the identity matrix, so we will prove that the RHS is also the identity matrix.

$$
\begin{aligned}
&(A + UCV)\left[A^{-1} - A^{-1}U\left(C^{-1} + VA^{-1}U\right)^{-1}VA^{-1}\right] \\
=&\left\{I - U\left(C^{-1} + VA^{-1}U\right)^{-1}VA^{-1}\right\} + \left\{UCVA^{-1} - UCVA^{-1}U\left(C^{-1} + VA^{-1}U\right)^{-1}VA^{-1}\right\} \\
=&\left\{I + UCVA^{-1}\right\} - \left\{U\left(C^{-1} + VA^{-1}U\right)^{-1}VA^{-1} + UCVA^{-1}U\left(C^{-1} + VA^{-1}U\right)^{-1}VA^{-1}\right\} \\
=&I + UCVA^{-1} - \left(U + UCVA^{-1}U\right)\left(C^{-1} + VA^{-1}U\right)^{-1}VA^{-1} \\
=&I + UCVA^{-1} - UC\left(C^{-1} + VA^{-1}U\right)\left(C^{-1} + VA^{-1}U\right)^{-1}VA^{-1} \\
=&I + UCVA^{-1} - UCVA^{-1} \\
=&I
\end{aligned}
$$

□

Theorem 1.5. *Use Woodbury's inversion lemma from the previous question to prove* $(A + B)^{-1} = \sum_{k=0}^{\infty}(-A^{-1}B)^k A^{-1}$.

Proof. Using Woodbury's inversion lemma with U = V = I_n, and defining C = B $\in \mathbb{F}^{n\times n}$:

$$
\begin{aligned}
(A + B)^{-1} &= A^{-1} - A^{-1}(B^{-1} + A^{-1})^{-1}A^{-1} \\
&= A^{-1} - A^{-1}(I + AB^{-1})^{-1} \\
&= A^{-1} - A^{-1}B(A + B)^{-1}
\end{aligned}
$$

We see that we again have $(A + B)^{-1}$. Repeating the process of filling in the expression for $(A + B)^{-1}$ we have an infinite sequence. Our end result will be: $(A + B)^{-1} = \sum_{k=0}^{\infty}(-A^{-1}B)^k A^{-1}$ □

Theorem 1.6. *Given a matrix A(x), the entries of which are functions of a (scalar) variable x. Find an expression for the derivative of the inverse matrix,* $\frac{dA^{-1}}{dx}(x)$ *in terms of* A^{-1} *(x) and* $\frac{dA}{dx}(x)$*. You may use the expansion from the previous question.*

Proof. We begin by considering $AA^{-1} = I$. Taking the derivative of this expression with respect to x we get: $\frac{d(AA^{-1})}{dx} = 0$. Applying the product rule we get that $\frac{d(AA^{-1})}{dx} = \frac{dA}{dx}A^{-1} + A\frac{dA^{-1}}{dx} = 0$. From this expression we find that $\frac{dA^{-1}}{dx} = -A^{-1}\frac{dA}{dx}A^{-1}$. Note that we have to apply the ordering of the product rule very strictly since matrix derivation generally does not commute. □

Theorem 1.7. *Let W be an invariant subspace of a linear operator* $\hat{A} \in End(V)$*; show that it is also an invariant subspace of* $\hat{A} + a\hat{1}$ *with* $a \in \mathbb{F}$ *some scalar.*

Proof. Pick a random $w \in W$. We know that $\hat{A}w \in W$. Next we apply $\hat{A} + a\hat{1}$ to w so that we have: $(\hat{A} + a\hat{1})w = \hat{A}w + aw \in W$. $\hat{A}w$ is an an element of W and aw is an element of W since it is closed under multiplication with a scalar so that W must be a invariant subspace of $(\hat{A} + a\hat{1})$. □

Theorem 1.8. *Show that* $ker(\hat{A} - \lambda\hat{1})^k$ *for some non-negative integer* $k \in \mathbb{N}$ *is an invariant subspace of* $\hat{A}$*.*

Proof. p.55
let $U_\lambda^{(k)} = ker((\hat{A} - \lambda)^k)$ with thus $U_\lambda^{(0)} = 0 \preceq U_\lambda^{(1)} = V_\lambda \preceq U_\lambda^{(2)} \preceq \cdots \preceq U_\lambda^{(k)} \preceq \cdots$
which we know will saturate to a fixed subspace $U_\lambda = U_\lambda^{(k \geq s_\lambda)}$ for all values of k above a minimal value s_λ. The space U_λ is known as the generalized eigenspace. Shifting $\hat{A}$ with a constant times the identity operator does not affect the results from the above construction: each of the spaces $U_\lambda^{(k)}$ are invariant subspaces of $\hat{A}$, and so is the natural complement of the generalized eigenspace, i.e. $U_\lambda^{(c)} = im((\hat{A} - \lambda)^{s_\lambda})$ □

Theorem 1.9. *Show that* $im(\hat{A} - \lambda\hat{1})^k$ *for some non-negative integer* $k \in \mathbb{N}$ *is an invariant subspace of* $\hat{A}$*.*

Proof. See theorem 1.8 for proof. □

Theorem 1.10. *Given two operators* $\hat{A}, \hat{B} \in End(V)$ *that admit a spectral decomposition (i.e., are diagonalisable) as* $\hat{A} = \sum_{\lambda \in \sigma_{\hat{A}}} \lambda \hat{P}_\lambda$ *and* $\hat{B} = \sum_{\lambda \in \sigma_{\hat{B}}} \mu \hat{Q}_\mu$*. If* $\hat{A}$ *and* $\hat{B}$ *commute,* $[\hat{A}, \hat{B}] = 0$*, show that they admit a common spectral decomposition (or thus, a basis in which they are simultaneously diagonal)*

Proof. Given that we can construct the spectral projectors $\hat{P}_\lambda$ of a diagonalisable operator $\hat{A} = \sum \lambda \hat{P}_\lambda$ as a polynomial of $\hat{A}$, it follows that if another operator $\hat{B}$ commutes with $\hat{A}$, then also $[\hat{B}, \hat{P}_\lambda] = \hat{0}$ for all $\lambda \in \sigma_{\hat{A}}$.
If also $\hat{B}$ is diagonalisable with spectral decomposition $\hat{B} = \sum_{\mu \in \sigma_{\hat{B}}} \mu \hat{Q}_\mu$, then

it follows that $[\hat{P}_\lambda, \hat{Q}_\mu] = \hat{0}$ for all λ, μ in the corresponding spectra. As a consequence , the operators $P_\lambda Q_\mu$ are projectors with

$$\sum_{\lambda,\mu} P_\lambda Q_\mu = (\sum_\lambda P_\lambda)(\sum_\mu Q_\mu) = \hat{1}$$

and we can write the spectral decomposition of both operators as
$\hat{A} = \sum_\lambda \lambda \hat{P}_\lambda = \sum_\lambda \lambda \hat{P}_\lambda (\sum_\mu \hat{Q}_\mu) = \sum_{\lambda,\mu} \lambda (\hat{P}_\lambda \hat{Q}_\mu)$
$\hat{B} = \sum_\mu \mu \hat{Q}_\mu = \sum_\mu \mu \hat{Q}_\mu (\sum_\lambda \hat{P}_\lambda) = \sum_{\lambda,\mu} \mu (\hat{P}_\lambda \hat{Q}_\mu)$
and we obtain a common spectral decomposition. □

Theorem 1.11. *Given linear operators $\hat{A}, \hat{B} \in End(V)$ which statisfy $\hat{A}\hat{B} + \hat{B}\hat{A} = \hat{0}$. Let $\hat{A}$ have an eigenvector v with an eigenvalue λ, i.e. $\hat{A}v = \lambda v$. Show that either $\hat{B}v = 0$, or that $\hat{A}$ also has an eigenvalue $-\lambda$.*

Proof. If we multiply the equation times v, we find that:

$$\begin{aligned}(\hat{A}\hat{B} + \hat{B}\hat{A})v &= \hat{0} \\ \hat{A}\hat{B}v &= -\hat{B}\hat{A}v \\ \hat{A}\hat{B}v &= -\hat{B}\lambda v \\ \hat{A}\hat{B}v &= -\lambda \hat{B}v\end{aligned}$$

We can easily see that if $\hat{B}v = 0$ this would be correct.
It is also quite clear that $\hat{A}$ has an eigenvalue $-\lambda$ with eigenvector $\hat{B}v$. □

Theorem 1.12. *Let the linear operator $\hat{P} \in End(V)$ be a projector, $\hat{P}^2 = \hat{P}$. What is the spectrum of $\hat{P}$?*

Proof. If we denote the spectrum of $\hat{P}$ as $\sigma(\hat{P})$. To find the values we have to find all $\alpha \in \sigma(\hat{P})$ such that $\alpha\hat{I} - \hat{P}$ is not invertible. We can easily find that $(\alpha\hat{I} - \hat{P})^{-1} = \frac{1}{\alpha}(\hat{I} + \frac{\hat{P}}{\alpha - 1})$, this is only possible if $\alpha \neq 0, 1$. The spectrum is thus $\sigma(\hat{P}) = \{0, 1\}$. □

Proof. (Alternative proof) If λ is an eigenvalue of $\hat{P}$ than we have:

$$\hat{P}v = \lambda v$$
$$\hat{P}^2 v = \hat{P}v = \lambda v = \lambda \hat{P}v = \lambda^2 v$$

So $\lambda^2 = \lambda$. This is only possible if $\lambda = 0$ or 1. □

Theorem 1.13. *Let the linear operators $\hat{P}, \hat{Q} \in End(V)$ be projectors in such a way that also $\hat{P} + \hat{Q}$ is a projector. Show that this implies that $\hat{P}\hat{Q} = \hat{0}$ (hint: the result of the previous two theorems can be useful).*

Proof. proof using the previous results: If $\hat{P}+\hat{Q}$ is a projector, this means that $(\hat{P}+\hat{Q})^2 = \hat{P}+\hat{Q} \Rightarrow \hat{P}^2+\hat{Q}^2+\hat{P}\hat{Q}+\hat{Q}\hat{P} = \hat{P}+\hat{Q}$. We know that $\hat{P}^2$ and $\hat{Q}^2$ are equal to $\hat{P}$ and $\hat{Q}$ respectively. So that $\hat{P}\hat{Q}+\hat{Q}\hat{P} = 0$. from this we can combine theorems 1.11 and 1.12 to say that: $\hat{P}\hat{Q}v_\lambda = 0$ for all eigenvectors v_λ of P (because the -λ case is excluded since -1 is not in the spectrum). Now since projection operators are diagonalisable, we know that V admits a basis of their eigenvectors, in other words: $\forall v \in V : v = c_\lambda v^\lambda$ (using einstein summation convention) s.t $\hat{P}\hat{Q} = 0$(in V) □

(note that an alternative proof using contradiction exists, i refer you to discord for this one)

Theorem 1.14. *Given a companion matrix*

$$C = \begin{bmatrix} 0 & 1 & 0\dots & 0 & 0 & \\ 0 & 0 & 1\dots & 0 & 0 & \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0\dots & 0 & 1 & \\ -p_0 & -p_1 & -p_2 & \dots - p_{n-2} & -p_{n-1} & \end{bmatrix}$$

write down the eigenvalue equation $Cv = \lambda v$. If we normalize v such that its first component $v^1 = 1$, what are the other components v^k for $k = 2, \dots, n$ of v and what equation does λ need to satisfy?

Proof. C is a companion matrix of some polynomial $\sum_k p_k z^k$, the eigenvalue equation is simply the characteristic polynomial.
if we normalize v^1 the other components are then given by $\tilde{v^k} = \frac{v^k}{v^1}$ s.t we have a monic polynomial. The characteristic polynomial of the companion matrix is then given by:

$$\sum_k p_k \lambda^k \tag{1}$$

If we find the roots of this equation, we find the eigenvalues.

□

Theorem 1.15. *Given a Jordan block corresponding to an eigenvalue λ of size 4,*

$$J_\lambda^{(4)} = \begin{bmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{bmatrix}$$

compute $exp(J_\lambda^{(4)})$

Proof. For a general analytic function f(x) (to say that it has a Taylor series) and a Jordan block of size k of an eigenvalue λ we know that

$$f(J^{(k)}(\lambda)) = \sum_{n=0}^{k-1} \tilde{f}(J^k(0))^n$$

where $\tilde{f}$ is defined as $\frac{f^{(n)}(\lambda)}{n!}$
luckily the n'th derivative of exp is always exp s.t:

$$exp(J^{(4)}(\lambda) = \frac{exp(\lambda)}{0!}I + \frac{exp(\lambda)}{1!}(J^4(0)) + \frac{exp(\lambda)}{2!}(J^4(0))^2 + \frac{exp(\lambda)}{3!}(J^4(0))^3$$

this results in the final matrix: $\begin{bmatrix} e^\lambda & e^\lambda & e^\lambda/2 & e^\lambda/6 \\ 0 & e^\lambda & e^\lambda & e^\lambda/2 \\ 0 & 0 & e^\lambda & e^\lambda \\ 0 & 0 & 0 & e^\lambda \end{bmatrix}$

□

Theorem 1.16. *Reconsider $J_\lambda^{(4)}$ from the previous question, where now $\lambda \in \mathbb{R}_{\geq 0}$. Compute $log(J_\lambda^{(4)})$.*

Proof. Since here we are dealing with only positive eigenvalues, we dont need to choose a branch cut.
saying now that $\lambda = 0$ is not an eigenvalue, we can simply make use of the taylor series expansion: $log(z) = log(\lambda + (z - \lambda)) = log(\lambda) + log(1 + \frac{z-\lambda}{\lambda})$ to define:

$$log(J^k(\lambda)) = log(\lambda)I + \sum_{n=1}^{k-1}(-1)^{n+1}\frac{(J^{(k)}(0))^n}{n\lambda^n}$$

(note that this also follows from the general expression in the previous theorem).
This results in the final matrix: $log(J^4(\lambda)) = \begin{bmatrix} log(\lambda) & \frac{1}{\lambda} & -\frac{1}{2\lambda^2} & \frac{1}{3\lambda^3} \\ 0 & log(\lambda) & \frac{1}{\lambda} & -\frac{1}{2\lambda^2} \\ 0 & 0 & log(\lambda) & \frac{1}{\lambda} \\ 0 & 0 & 0 & log(\lambda) \end{bmatrix}$
□

2 Norms and inner products

Theorem 2.1. *Let $\|v_p\|$ be the standard Hölder p-norm of the vector $v \in \mathbb{F}^n$. Find the lowest value of b such that $\|v\|_1 \leq b\|v\|_2$ is satisfied for all $v \in \mathbb{F}^n$. Can you find a specific vector v for which this inequality is then satisfied?*

Proof. if we write the inequality explicitly:

$$\sum_{i=1}^{n}|v^i| \leq b(\sum_{i=1}^{n}|v^i|^2)^{1/2}$$

but now consider the Cauchy-Schwarz inequality:

$$(\sum_{i=1}^{n}|v^i|)^2 \leq (\sum_{i=1}^{n}|v^i|^2)(\sum_{i=1}^{n}|1|^2)$$

s.t we find that:

$$(\sum_{i=1}^{n}|v^i|)^2 \leq n(\sum_{i=1}^{n}|v^i|^2)$$

from which we trivially see that b $= \sqrt{n}$. The inequality becomes an equality if $\forall i : v^i = c$. () □

Theorem 2.2. *Let $A \in$ Hom(V, W) be a linear map between vector spaces V and W that is bounded, i.e. the induced norm*

$$\|A\|_{V,W} = C < \infty.$$

Show that A is a continuous map between V and W.

Proof. A linear map $\hat{A} \in Hom(V, W)$ is continuous if and only if it is bounded, i.e. if there exists a constant C such that

$$\|\hat{A}v\|_W \leq C\|v\|_V$$

for all $v \in V$. Indeed if $\hat{A}$ is bounded with constant C, then it follows that for any $\|v - v'\|_V < \delta_\epsilon = \epsilon/C$,

$$\|\hat{A}v - \hat{A}v'\|_W = \|\hat{A}(v - v')\|_W \leq C\|v - v'\|_V < \epsilon$$

so that A represents a continuous map. □

Theorem 2.3. *Let* $\mathrm{A} \in \mathrm{Hom}(\mathbb{F}^n, \mathbb{F}^m)$, *i.e.* $\mathrm{A} \in \mathbb{F}^{m\times n}$, *where furthermore on $W = \mathbb{F}^m$ and $V = \mathbb{F}^n$ we consider the 1-norm. Consider the induced norm for A, given by*

$$\|\mathrm{A}\|_{1,1} = \sup\left\{\frac{\|\mathrm{A}v\|_1}{\|v\|_1} = \frac{\sum_{i=1}^{m}\left|\sum_{j=1}^{n} A^i_j v^j\right|}{\sum_{j=1}^{n}|v^j|}, v \in \mathbb{F}^n, \neq \mathbf{0}\right\}.$$

First prove the upper bound

$$\frac{\|\mathrm{A}v\|_1}{\|v\|_1} \leq \max_{j=1,\dots,n} \sum_{i=1}^{m} \left|A^i{}_j\right|$$

for all $v \in \mathbb{F}^n$. Also show that you can find a vector v for which this inequality is satisfied, so that this upper bound is tight. Together, these results imply that

$$\|\mathrm{A}\|_{1,1} = \max_{j=1,\dots,n} \sum_{i=1}^{m} \left|A^i_j\right|.$$

Proof.

$$\begin{aligned} \sum_j |\sum_i A^i_j x^j| \leq \sum_j \sum_i |A^i_j x^j| = \sum_j \sum_i |A^i_j||x^j| \\ \sum_j \sum_i |A^i_j||x^j| \leq \max_j (\sum_i |A^i_j|)\|x\|_1 \\ \sum_j |\sum_i A^i_j x^j| \leq \max_j (\sum_i |A^i_j|)\|x\|_1 \end{aligned} \tag{2}$$

where now dividing both sides by $\|x\|_1$ gives the desired result. (the first term of these inequalities is the one norm of the result vector of Ax written explicitly. if you are not convinced, write out for some matrix and vector and you should see that it is indeed the same.) ☐

Theorem 2.4. *Given a submultiplicative norm on the space of linear operators $\hat{A}, \hat{B} \in \mathrm{End}(V)$, i.e. a norm that satisfies $\|\hat{A}\hat{B}\| \leq \|\hat{A}\|\|\hat{B}\|$. Show that* $\lim_{n\to\infty} \left\|\hat{A}^n\right\|^{1/n} = \rho_{\hat{A}}$ *with $\rho_{\hat{A}}$ the spectral radius (the magnitude of the largest magnitude eigenvalue) of $\hat{A}$. This result is known as Gelfand's formula.*

Proof. submultiplicative norms satisfy $\|\hat{A}^n\| \leq \|\hat{A}\|^n$ and thus $\|\hat{A}\| < 1$ implies that $lim_{n\to\infty}\hat{A}^n \to \hat{0}$. On the other hand, from the Jordan decomposition, we know that $lim_{n\to\infty}\hat{A}^n \to \hat{0}$ is in one-to-one correspondence with all eigenvalues having a magnitude smaller than 1, i.e. $\rho_{\hat{A}} < 1$. Now consider an operator $\hat{B}_\epsilon = (\rho_{\hat{A}} + \epsilon)^{-1}\hat{A}$. It has a spectral radius $\rho_{\hat{B}_\epsilon} = \rho_{\hat{A}}(\rho_{\hat{A}} + \epsilon)^{-1} < 1$ and thus $lim_{n\to\infty}(\hat{B}_\epsilon)^n \to \hat{0}$ for any $\epsilon > 0$. Hence, $lim_{n\to\infty}\|(\hat{B}_\epsilon)^n\| \to 0$, and there exist some N_ϵ such that $\|(\hat{B}_\epsilon)^n\| < 1$ for all $n > N_\epsilon$. Equivalently, $\left\|\hat{A}^n\right\| < \left(\rho_{\hat{A}} + \varepsilon\right)^n$ or thus $\left\|\hat{A}^n\right\|^{1/n} < \rho_{\hat{A}} + \varepsilon$ for all $n > N_\varepsilon$. Combined with $\|\hat{A}^n\|^{1/n} \geq \rho_{\hat{A}}$, we obtain the Gelfand formula:

$$lim_{n\to\infty}\|\hat{A}^n\|^{1/n} = \rho_{\hat{A}}$$

☐

Theorem 2.5. *Consider the invertible linear operator* $\hat{A} \in \mathrm{End}(V)$ *on the normed vector space* $(V, \|\cdot\|)$*, and consider the linear system* $\hat{A}x = y$*. Consider a perturbed linear system* $\hat{A}(x + \Delta x) = (y + \Delta y)$*. Show that you can bound the relative norm of the error* $\|\Delta x\|/\ \|x\|$ *in terms of the relative norm of the perturbation* $\|\Delta y\|/\ \|y\|$ *and the condition number* $\kappa(\hat{A}) = \|\hat{A}\| \left\|\hat{A}^{-1}\right\|$*, where for* $\hat{A} \in \mathrm{End}(V)$ *we use the induced norm. Also show that* $\kappa(\hat{A}) \geq 1$*.*

Proof. For a small variation or error $y \to y + \Delta y$, the solution will change tot $x \to x + \Delta x$ with $\Delta x = \hat{A}^{-1}\Delta y$, which can be bound as $\|\Delta x\| \leq \|\hat{A}^{-1}\|\|\Delta y\|$. OFten, we are not interested in the absolute error, but rather relative error $\|\Delta x\|/\|x\|$. From $\|A\|\|x\| \geq \|y\|$, we obtain

$$\frac{\|\Delta x\|}{\|x\|} \leq \frac{\|\hat{A}^{-1}\|\|\Delta y\|}{\|x\|} \leq \frac{\|\hat{A}^{-1}\|\|\hat{A}\|\|\Delta y\|}{\|y\|} = \kappa(\hat{A})\frac{\|\Delta y\|}{\|y\|}$$

Where we have introduced the condition number $\kappa(\hat{A}) = \|\hat{A}^{-1}\|\|\hat{A}\| \geq \|\hat{1}\| \geq 1$ □

Theorem 2.6. *Given an approximate solution* $\tilde{x}$ *for the linear system* $\hat{A}x = y$*, so that* $r = y - \hat{A}\tilde{x}$ *is not exactly the zero vector. Bound the relative error* $\|x - \tilde{x}\|/\|x\|$ *between the exact and approximate solution in terms of the condition number of* $\hat{A}$ *and the norms of* r *and* y

Proof.

$$\frac{\|x - \tilde{x}\|}{\|x\|} \leq \frac{\|\hat{A}^{-1}\|\|r\|}{\|x\|}\frac{\|\hat{A}\|\|x\|}{\|y\|} = \kappa(\hat{A})\frac{\|r\|}{\|y\|}$$

□

Theorem 2.7. *Given a vector space* V *with a positive definite inner product* $\langle , \rangle$*. Prove Cauchy-Schwartz-Bunjakowski inequality*

$$|\langle v, w\rangle|^2 \leq \langle v, v\rangle\langle w, w\rangle$$

for all $v, w \in V$

Proof.

$$\begin{aligned}
0 &\leq \langle av - w, av - w\rangle \\
&\leq |a|^2\langle v, v\rangle + \langle w, w\rangle - \overline{a}\langle v, w\rangle - a\langle w, v\rangle \\
&\leq |\frac{\langle w, v\rangle}{\langle v, v\rangle}|^2\langle v, v\rangle + \langle w, w\rangle - \frac{\overline{\langle w, v\rangle}}{\langle v, v\rangle}\langle v, w\rangle - \frac{\langle w, v\rangle}{\langle v, v\rangle}\langle w, v\rangle \qquad (3)\\
&\leq \langle w, w\rangle\langle v, v\rangle - |\langle v, w\rangle|^2 \\
|\langle v, w\rangle|^2 &\leq \langle w, w\rangle\langle v, v\rangle
\end{aligned}$$

where we let a $= \frac{\langle w,v\rangle}{\langle v,v\rangle}$ and we multiplied both sides by $\langle v, v\rangle$ □

Theorem 2.8. *Given a vector space V with a positive definite inner product $\langle,\rangle$, and consider a set of orthonormal vectors $\{e_i, i = 1, \dots, n\}$ in V (which does not necessarily constitute a complete set). Prove Bessel's inequality*

$$\|v\|^2 \geq \sum_{i=1}^{n} |\langle e_i, v\rangle|^2$$

where we use the inner product norm $\|v\|^2 = \langle v, v\rangle$

Proof. Firstly we have a general vector $v \in V$, we find

$$\begin{aligned} 0 \leq \|v - \sum_{i=1}^{n} a^i e_i\|^2 = \|v\|^2 + \sum_{i=1}^{n} |a^i|^2 - \sum_{i=1}^{n} (\bar{a^i}\langle e_i, v\rangle + a^i\langle v, e_i\rangle) \\ = \|v\|^2 + \sum_{i=1}^{n} |a^i - \langle e_i, v\rangle|^2 - \sum_{i=1}^{n} |\langle e_i, v\rangle|^2 \end{aligned}$$

now due to the basis set e_i being orthonormal, we know $a^i = \langle e_i, v\rangle$ which gives Bessel's inequality. □

Theorem 2.9. *Let $(V, \langle,\rangle_V)$ and $(W, \langle,\rangle_W)$ be Euclidean/unitary spaces, and consider a bounded linear map $\hat{A} \in \mathrm{Hom}(V, W)$. Show that, with respect to the induced norms on $\mathrm{Hom}(V, W)$ and $\mathrm{Hom}(W, V)$, we have $\|\hat{A}\| = \left\|\hat{A}^\dagger\right\|$. Make sure to properly introduce or motivate your definition for the induced norm!*

Proof.

$$\|\hat{A}\| = sup_{\|v\|_V=1} \|\hat{A}v\|_W = sup_{\|v\|_V=1 \textbf{ en } \|w\|_W=1} |\langle w, \hat{A}v\rangle|$$

but since $|\langle w, \hat{A}^\dagger v\rangle| = |\langle v, \hat{A}^\dagger w\rangle|$ this definition also gives rise to $\|\hat{A}^\dagger\|$ the equality of the second and final expression comes from cauchy-schwarz as:

$$\langle w, \hat{A}v\rangle \leq \sqrt{\langle w, w\rangle\langle \hat{A}v, \hat{A}v\rangle}$$

now by definition the norm of w is one, and this inequality is saturated in the case of the supremum □

Theorem 2.10. *Consider a bounded linear map $\hat{A} \in \mathrm{Hom}(V, W)$ between the Euclidean/unitary spaces $(V, \langle,\rangle_V)$ and $(W, \langle,\rangle_W)$. Show that $W = \mathrm{im}(\hat{A}) \oplus \ker\left(\hat{A}^\dagger\right)$*

Proof. Consider the null space of $\hat{A}^\dagger$: for any $w \in ker(\hat{A}^\dagger)$, we find

$$\hat{A}^\dagger w = 0 \in V \Rightarrow \langle \hat{A}^\dagger w, v\rangle_V = \langle w, \hat{A}v\rangle_W = 0, \forall v \in V$$

And thus, since the second vector in the last inner product can be any vector in $im(\hat{A})$, we have $ker(\hat{A}^\dagger) = im(\hat{A})^\perp$. As $ker(\hat{A})$, like any null space, is a closed subspace, we find

$$W = ker(\hat{A}^\dagger) \oplus im(\hat{A})$$

□

Theorem 2.11. *Consider a bounded linear map* $\hat{A} \in \mathrm{Hom}(V, W)$ *between the Euclidean/unitary spaces* $(V, \langle,\rangle_V)$ *and* $(W, \langle,\rangle_W)$*. Using the norm associated with these positive definite inner products, we define the distance functions* $d_W(w, w') = \|w - w'\|_W$ *for all* $w, w' \in W$ *and similarly* $d_V(v, v') = \|v - v'\|_V$ *for all* $v, v' \in V$*. Show that the isometry condition* $d_W\left(\hat{A}v, \hat{A}v'\right) = \left\|\hat{A}v - \hat{A}v'\right\|_W = d_V(v, v') = \|v - v'\|_V$ *for all* $v, v' \in V$ *requires that* $\hat{A}^\dagger\hat{A} = \hat{1}_V$ *(necessary and sufficient condition).*

Proof. if $\hat{A}$ represents an isometric mapping with respect to the metrics $d_W(w', w) = \|w' - w\|_W$ and $d_V(v', v) = \|v - v'\|_V$, then $d_W(\hat{A}v', \hat{A}v) = d_V(v', v)$ implies

$$\langle \hat{A}v, \hat{A}v \rangle_W = \langle v, \hat{A}^\dagger \hat{A}v \rangle_V = \langle v, v \rangle_V$$

This is only possible if $\hat{A}^\dagger\hat{A} = \hat{1}v$. □

Theorem 2.12. *Given a (bounded) linear operator* $\hat{A} \in \mathrm{End}(V)$ *on a Euclidean/unitary space* $(V, \langle,\rangle)$*, that is self-adjoint:* $\hat{A} = \hat{A}^\dagger$*. Let* $\hat{A}$ *have an eigenvector* v *with eigenvalue* λ*. Show that* λ *is real.*

Proof. If (λ, v) is an eigenvalue-eigenvector pair of a self-adjoint operator $\hat{A}$, we find

$$\langle v, \hat{A}v \rangle = \langle v, \lambda v \rangle = \lambda \langle v, v \rangle = \langle \hat{A}v, v \rangle = \langle \lambda v, v \rangle = \bar{\lambda} \langle v, v \rangle$$

Since an eigenvector is not the zero vector, $\langle v, v \rangle > 0$, and we thus have $\lambda = \bar{\lambda}$, this proofs λ is real. □

Theorem 2.13. *Given a (bounded) linear operator* $\hat{A} \in \mathrm{End}(V)$ *on a Euclidean/unitary space* $(V, \langle,\rangle,$ *, for which* $\|\hat{A}v\| = \left\|\hat{A}^\dagger v\right\|$ *for all* $v \in V$*. Show that* $\hat{A}$ *is normal, i.e.* $\hat{A}\hat{A}^\dagger = \hat{A}^\dagger\hat{A}$

Proof.

$$\begin{aligned} \|\hat{A}\| &= \|A^\dagger\| \\ \langle \hat{A}v, \hat{A}v \rangle &= \langle A^\dagger v, A^\dagger v \rangle \\ \langle v, A^\dagger \hat{A}v \rangle &= \langle v, \hat{A}A^\dagger v \rangle \\ A^\dagger \hat{A} &= \hat{A}A^\dagger \end{aligned} \tag{4}$$

□

Theorem 2.14. *Given a (bounded) linear operator $\hat{A} \in \mathrm{End}(V)$ on a Euclidean/unitary space $(V, \langle , \rangle)$, that is normal, i.e. $\hat{A}\hat{A}^\dagger = \hat{A}^\dagger\hat{A}$. Let $\hat{A}$ have an eigenvector v with eigenvalue λ. Show that v is also an eigenvector of $\hat{A}^\dagger$. What is the corresponding eigenvalue?*

Proof. Consider an eigenvector v with eigenvalue λ of a normal operator $\hat{A}$. Using this relation; we find that $\|(\hat{A} - \lambda)v\| = 0$ implies $\|(\hat{A}^\dagger - \bar{\lambda})v\| = 0$, i.e. v is also an eigenvector of $\hat{A}^\dagger$ with eigenvalue $\bar{\lambda}$ □

Theorem 2.15. *Given a (bounded) linear operator $\hat{A} \in \mathrm{End}(V)$ on a Euclidean/unitary space $(V, \langle , \rangle)$, that is normal, i.e. $\hat{A}\hat{A}^\dagger = \hat{A}^\dagger\hat{A}$. Decompose $\hat{A} = \hat{A}_1 + \mathrm{i}\hat{A}_2$ where $\hat{A}_1$ and $\hat{A}_2$ are self-adjoint: express $\hat{A}_1, \hat{A}_2$ in terms of $\hat{A}$. What can you say about $\left[\hat{A}_1, \hat{A}_2\right]$. Given an eigenvector v of $\hat{A}$ with eigenvalue λ. Is v also an eigenvector of $\hat{A}_1$ and $\hat{A}_2$? If so, what are the corresponding eigenvalues? Does this result also hold if $\hat{A}$ is not normal?*

Proof. Any operator can be split into two self-adjoint parts acoording to

$$\hat{A} = \frac{\hat{A} + \hat{A}^\dagger}{2} + i\frac{\hat{A} - \hat{A}^\dagger}{2i} = \hat{A}_1 + i\hat{A}_2$$

which could be considered as the analogue of separating a complex number into its real and complex part. Note that this way of writing $\hat{A}$ then implies that $\hat{A}^\dagger = \hat{A}_1 - i\hat{A}_2$. Imposing that $\hat{A}$ is normal amounts to $[\hat{A}_1, \hat{A}_2] = \hat{0}$. Combined with the result on eigenvectors, this implies that any eigenvector v of $\hat{A}$ with eigenvalue λ is also an eigenvector of $\hat{A}_1$ and $\hat{A}_2$ with $\hat{A}_1 v = Re(\lambda)v$ and $\hat{A}_2 v = Im(\lambda)v$ note that this result hinges on the commutator being zero, s.t this does not hold for non normal operators □

Theorem 2.16. *Given a (bounded) linear operator $\hat{A} \in \mathrm{End}(V)$ on a Euclidean/unitary space $(V, \langle , \rangle)$, that is normal, i.e. $\hat{A}\hat{A}^\dagger = \hat{A}^\dagger\hat{A}$. Show that $\|\hat{A}\| = \rho_{\hat{A}}$, i.e. the operator norm associated with the inner product norm equals the spectral radius.*

Proof. We generally have that $\|\hat{A}^n\| \leq \|\hat{A}\|^n$. Equality is obtained if we can also show $\|\hat{A}^n\| \leq \|\hat{A}\|^n$, for which we prove, by induction, that $\|\hat{A}^n v\| \geq \|\hat{A}v\|^n$ for any vector $v \in V$ with normalisation $\|v\| = 1$. For $n = 2$, we have $\|\hat{A}^2 v\| = \|\hat{A}(\hat{A}v)\| = \|\hat{A}^\dagger\hat{A}v\| \geq \langle v, \hat{A}^\dagger\hat{A}v\rangle = \|\hat{A}v\|^2$ by the Cauchy-Schwarz inequality, the inductive step then follows from

$$\|\hat{A}^{n+1}v\| = \|\hat{A}^n \frac{\hat{A}v}{\|\hat{A}v\|}\|\|\hat{A}v\| \geq \|\hat{A}\frac{\hat{A}v}{\|\hat{A}v\|}\|^n\|\hat{A}v\| = \|\hat{A}^2 v\|^n\|\hat{A}v\|^{1-n} \geq \|\hat{A}v\|^{n+1}$$

A consequence of this result is that $\rho_{\hat{A}} = \|\hat{A}\|$. □

Theorem 2.17. *Given a (bounded) linear operator $\hat{A} \in \mathrm{End}(V)$ on a Euclidean/unitary space $(V, \langle , \rangle)$, that satisfies $\hat{A}^2 = \hat{0}$. Show that $\hat{A}$ cannot be a normal operator.*

Proof. Suppose $\hat{A}$ is normal and nilpotent with indesx s = 2. For any vector $v \in V$, we would find $0 = \|\hat{A}^2 v\| = \|\hat{A}^\dagger \hat{A} v\|$, and thus $\hat{A}^\dagger \hat{A} v = o$. Indeed, any vector $\hat{A}v$ is either the zero vector or a nonzero eigenvector of $\hat{A}$ with eigenvalue zero. But any eigenvector of $\hat{A}$ with eigenvalue zero is also an eigenvector of $\hat{A}^\dagger$ with eigenvalue zero. But this means that $\langle v, \hat{A}^\dagger \hat{A} v \rangle = \|\hat{A}v\|^2 = 0$, so that $\hat{A}v = 0$ for all $v \in V$, and thus that $\hat{A} = \hat{0}$, which contradicts the assumption that $\hat{A}$ is nilpotent with index 2. □

3 Unitary similarity and unitary equivalence

Theorem 3.1. *Given a matrix* $\mathrm{A} \in \mathbb{F}^{n\times n}$. *Under which condition on* A *is* $\mathrm{U} = \exp(\mathrm{A})$ *a unitary matrix?*

Proof. Since we have $U(t) = exp(tA)$ we have $U(t)^{-1} = exp(-tA)$ and $U(t)^H = exp(tA^H)$, we obtain the conditions $A^H = -A$ i.e. A is an anti-Hermitian matrix. □

Theorem 3.2. *Given a matrix* $A \in \mathbb{R}^{n\times n}$ *that satisfies* $A^\top = -A$. *What can you say about* $\det(\exp(\mathrm{A}))$

Proof. We use the relation $det(exp(A)) = exp(tr(A))$ to find that $det(exp(A)) = 1$. □

Theorem 3.3. *Given a Householder matrix* $\mathrm{H} = \mathrm{I} - 2\frac{vv^H}{v^H}$, *with* $v \in \mathbb{C}^n$ *and thus* $\mathrm{H} \in \mathrm{C}^{n\times n}$. *How should you choose* v *such that* Hw *becomes parallel to* e_1, *for a given vector* w. *Here,* e_1 *is the first coordinate vector:* $e_1 = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \end{bmatrix}^\top$.

Proof. Given a vector $w = (w^1, w^2, \dots, w^n)$, we will try to make all entries w^k for $k = 2, \dots, n$ zero by acting with H, thus moving all the 'weight' of the vector to the first component. We thus want

$$Hw = w - \frac{2v^H w}{v^H v} v = e^{i\theta}\|w\|e_1$$

for some value of the phase θ, from which we observe that v will need to be a linear combination of w and $e_1 : v = w + \alpha e_1$. □

Theorem 3.4. *Consider a matrix* $\mathrm{A} \in \mathbb{C}^{n\times n}$ *that is circulant, i.e.* $A^i_j = f_{(j-i) \bmod n}$ *with* $f_0, f_1, \dots, f_{n-1} \in C$. *Construct the eigenvectors and eigenvalues of* A.

Proof. It can easily be seen that the columns of the Fourier matrix U diagonalise such a matrix. When denoting columns of U as $u_k = (n^{-1/2}\omega^{kj})_{j=0,\dots,n-1}$ for $k = 0, \dots, n-1$, we obtain

$$\begin{aligned}(Au_k)^i &= \frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} A^i_j \omega^{kj} = \frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} f_{(j-1) \text{ mod n}} \omega^{k(j-i+i)} \\ &= \frac{\omega^{ki}}{\sqrt{n}} \sum_{j=0}^{n-1} f_{(j-1) \text{ mod n}} \omega^{k(j-i)} = (u_k)^i \lambda_k\end{aligned}$$

so that u_k is an eigenvector with corresponding eigenvalue λ_k given by

$$\lambda_k = \sum_{j=1}^{n} f_{(j-1) \text{ mod n}} \omega^{k(j-i)} = \sum_{j=0}^{n-1} f_j \omega^{kj} = \sum_{j=0}^{n-1} f_j e^{-i\frac{2\pi}{n}kj}$$

□

Theorem 3.5. *Prove that any matrix* $\mathrm{A} \in \mathrm{C}^{n\times n}$ *admits a Schur decomposition* $\mathrm{A} = \mathrm{ZTZ^H}$, *where* $\mathrm{Z} \in \mathbb{C}^{n\times n}$ *is unitary and* $\mathrm{T} \in \mathbb{C}^{n\times n}$ *is upper triangular* $\left(T^i_j = 0 \text{ if } i > j\right)$.

Proof. It is definitely valid for the case n=1, with A = T and Z = 1. Assume it holds for all matrix in $\mathbb{F}^{(n-1)\times(n-1)}$ and let $A \in \mathbb{F}^{n\times n}$. Being a finite-dimensional matrix, A has at least one eigenvalue λ with at least one eigenvector v, which we can normalize to be a unit vector $u = v/\|v\|$. Complete $\{u\}$ to be a complete set, or equivalently, construct a unitary matrix U whose first column is u, i.e. $U^i_1 = u^i$. We now have

$$AU = U \begin{bmatrix} \lambda & \dots \\ O_{(n-1)\times 1} & \tilde{A} \end{bmatrix} \Leftrightarrow A = U \begin{bmatrix} \lambda & \dots \\ O_{(n-1)\times 1} & \tilde{A} \end{bmatrix} U^H$$

with $\tilde{A} \in \mathbb{F}^{(n-1)\times(n-1)}$, so that it has a Schur decomposition $\tilde{A} = \tilde{Z}\tilde{T}\tilde{Z}^H$. We thus find

$$A = U \begin{bmatrix} 1 & O_{1\times(n-1)} \\ O_{(n-1)\times 1} & \tilde{Z} \end{bmatrix} \begin{bmatrix} \lambda & \dots \\ O_{(n-1)\times 1} & \tilde{T} \end{bmatrix} \begin{bmatrix} 1 & O_{1\times(n-1)} \\ O_{(n-1)\times 1} & \tilde{Z}^H \end{bmatrix} U^H$$

First matrix is Z, second T and last Z^H, where clearly T is defined as upper triangular and Z is unitary. □

Theorem 3.6. *Given a matrix* $\mathrm{A} \in \mathrm{C}^{n\times n}$ *that is normal. Show that it can be unitarily diagonalised, i.e.* $\mathrm{A} = \mathrm{UDU^H}$ *with* $\mathrm{U} \in \mathbb{C}^{n\times n}$ *unitary and* $\mathrm{D} \in \mathrm{C}^{n\times n}$ *diagonal* $\left(D^i_j = 0 \text{ if } i \neq j\right)$. *You can use the existence of the Schur decomposition from the previous question.*

Proof. Schur decomposition of a normal matrix must satisfy $T^H T = T T^H$. Taking row i and column j of this matrix equation, we find

$$\sum_{k=1}^{min(i,j)} \bar{T}^k_i T^k_j = \sum_{k=min(i,j)}^{n} T^i_k \bar{T}^j_k$$

Starting with for example $(i,j) = (1,1)$, we find $|T^1_1|^2 = \sum_{k=1}^n |T^1_k|^2$, which is only possible if $T^1_k = 0$ for $k = 2, \dots, n$. Continuing along these lines, we see that a normal matrix A must have a Schur decomposition where T is diagonal. Hence, the Schur decomposition coincides with the eigenvalue decompostion, and a normal matrix is unitarily diagonalisable. □

Theorem 3.7. *Given a matrix* $\mathrm{A} \in \mathrm{C}^{m\times n}$ *(assuming* $m \geq n$ *) and its (thin) QR decomposition* $\mathrm{A} = \mathrm{QR}$ *with* $\mathrm{Q} \in \mathbb{F}^{m\times n}$ *isometric and* $\mathrm{R} \in \mathbb{F}^{n\times n}$ *upper triangular. Relate the factors in the (thin) singular value decomposition of A with those in the the singular value decomposition of* R.

Proof. We see the following:

$$\begin{aligned} A &= QR \\ A &= U_A S_A V_A^H \\ R &= U_R S_R V_R^H \\ U_A S_A V_A^H &= Q U_R S_R V_R^H \\ \Rightarrow U_A = QU_R, \; S_A &= S_R, \; V_A = V_R \end{aligned}$$

□

Theorem 3.8. *Given a matrix $A \in C^{n\times n}$ and its singular value decomposition $A = USV^H$ with $\mathrm{U}, \mathrm{V} \in \mathbb{F}^{n\times n}$ unitary, and $\mathrm{S} \in \mathbb{F}^{n\times n}$ diagonal. Construct the unitary (!) matrix that diagonalises the hermitian matrix*

$$\left[\begin{array}{cc} \mathrm{O} & \mathrm{A^H} \\ \mathrm{A} & \mathrm{O} \end{array} \right].$$

Show that your answer is unitary, and that it diagonalises this matrix.

Proof. The unitary matrix that diagonalises this is:

$$A = \frac{1}{\sqrt{2}} \begin{bmatrix} U & V \\ U & -V \end{bmatrix}$$

this matrix is indeed unitary, as:

$$\begin{aligned} A^H A &= \frac{1}{2} \begin{bmatrix} U & V \\ U & -V \end{bmatrix}^H \begin{bmatrix} U & V \\ U & -V \end{bmatrix} \\ &= \frac{1}{2} \begin{bmatrix} U^H & U^H \\ V^H & -V^H \end{bmatrix} \begin{bmatrix} U & V \\ U & -V \end{bmatrix} \\ &= \frac{1}{2} \begin{bmatrix} U^H U + U^H U & 0 \\ 0 & V^H V + V^H V \end{bmatrix} \\ &= I \end{aligned} \tag{5}$$

and that it diagonalises the matrix is a trivial calculation □

Theorem 3.9. *Given a matrix $A \in \mathbb{C}^{m\times n}$; consider the linear system $Ax = y$ that is potentially overdetermined. Suppose in particular that A is not full rank, i.e. $\rho(\mathrm{A}) = p < \min(m, n)$. Construct a solution x that minimises $\|\mathrm{A}x - y\|$ using the standard Euclidean norm. Explain wether or not this solution unique.*

Proof. For an overdetermined system $Ax = y$ where A can be a general rectangular matrix and $y \notin im(A)$, we can write the least squares solution x that minimises $\|Ax - y\|$ using the compact singular value decomposition as

$$x = V_p S_p^{-1} U_p^H y = A^+ y$$

where we again introduced the notation A^+ for the Moore-Penrose pseudoinverse. ($A^+ = (A^H A)^{-1} A^H$) note that this is not unique as the SVD decomposition is not unique □

Theorem 3.10. *Given a matrix* A $\in \mathbb{C}^{m \times n}$*; we want to approximate* A *by a matrix* B *that has at most rank* $r < \min(m, n)$*. Prove that the operator norm* $\|$A $-$ B$\|$ *is minimised by choosing* B = U_r S_r V_r^H*, the reduced singular value decomposition obtained by retaining only the r largest singular values (and corresponding left and right singular vectors) of A. This result is known as the Eckhart-Young-Mirsky theorem. (There is no need to prove the equivalent statement using the Frobenius norm.)*

Proof. Any rank-r matrix B has $\nu(B) = n - r$. Hence, the subspace spanned by the frist r + 1 colmuns of V cannot be disjoint from ker(B) because of dimensionality; let $w = \sum_{i=1}^{r+1} a^i v_i$ denoted a unit vector in this intersection, with v_i the ith column of V. Using $\sum_{i=1}^{r+1} |a^i|^2 = 1$, we find

$$\|A - B\| \geq \|(A - B)w\| = \|Aw\| = \|SV^H w\| = [\sum_{i=1}^{r+1} |\sigma_i ai|^2]^{1/2} \geq \sigma_{r+1}$$

The lower bound is exactly saturated by choosing B equal to the truncated singular value decomposition of A. □

4 Multilinear algebra

Theorem 4.1. *Consider two square matrices* $\mathrm{A}_1 \in \mathbb{F}^{m\times m}$ *and* $\mathrm{A}_2 \in \mathbb{F}^{n\times n}$ *that are diagonalisable. What would be the eigenvalues and eigenvectors of the Kronecker product* $\mathrm{A}_1 \otimes \mathrm{A}_2$ *?*

Proof.

$$A = A_1 \otimes A_2 = \begin{bmatrix} (A_1)^1_1 A_2 & (A_1)^1_2 A_2 & \dots & (A_1)^1_{n_1} A_2 \\ (A_1)^2_1 A_2 & (A_1)^2_2 A_2 & \dots & (A_1)^2_{n_1} A_2 \\ \vdots & \vdots & \ddots & \vdots \\ (A_1)^{m_1}_1 A_2 & (A_1)^{m_1}_2 A_2 & \dots & (A_1)^{m_1}_{n_1} A_2 \end{bmatrix}$$

Let us observe the linear operators A_1 and A_2 with eigenvectors x_i and y_j respectively. For simplicity's sake we will look at just two eigenvectors x_i and y_j with corresponding eigenvalues λ_i and μ_j:

$$A_1 x_i = \lambda_i x_i$$
$$A_2 y_j = \mu_j y_j$$

Taking $(A_1 \otimes A_2)(x_i \otimes y_j)$ yields:

$$\begin{aligned}(A_1 \otimes A_2)(x_i \otimes y_j) &= (A_1 x_i) \otimes (A_2 y_j) \\ &= (\lambda_i x_i) \otimes (\mu_j y_j) \\ &= \lambda_i \mu_j (x_i \otimes y_j)\end{aligned}$$

Thus the eigenvalues of $A_1 \otimes A_2$ are given by: $\lambda_i \mu_j \ \forall 1 \leq i \leq m, 1 \leq j \leq n$ and its eigenvectors by: $(x_i \otimes y_j) \ \forall 1 \leq i \leq m, 1 \leq j \leq n$ □

Theorem 4.2. *Let* V *be an* n*-dimensional vector space with basis* $\{e_i, i = 1, \dots, n\}$. *Consider the linear operator* $\hat{S} \in \mathrm{End}(V \otimes V)$ *that acts on the basis vectors as* $\hat{S} e_i \otimes e_j = e_j \otimes e_i$. *Show that* $\hat{S}^2 = \hat{1}_{V\otimes V}$. *This equation implies that* $\hat{S}$ *can only have two different eigenvalues; which are those? Can you determine the dimensionality of the associated eigenspaces (geometric multiplicity) and discuss the structure of the associated eigenvectors.*

Proof. showing $\hat{S}^2 = 1_{(V\otimes V)}$ is rather trivial, we can simply look at the action on the basis vectors :

$$\hat{S}^2(e_i \otimes e_j) = \hat{S}(\hat{S}(e_i \otimes e_j)) = \hat{S}(e_j \otimes e_i) = \hat{1}_{V\otimes V}$$

now, if we consider this property and simply apply it to the eigenvalue equation it shows us that the only possible eigenvalues are ± 1 (this also follows directly from the fact that S is a permutation operator). We know that eigenvectors of the operator with eigenvalue one are vectors v that are structured as: $(e_i \otimes e_j) + (e_j \otimes e_i)$ where the dimensionality is then equal to C^n_2 and the vectors with eigenvalue -1 are those that are as: $(e_i \otimes e_j) - (e_j \otimes e_i)$ where the dimensionality follows directly from the fact that the direct sum of eigenspaces must have dimension n^2

$\square$

5 Function spaces

Theorem 5.1. *Let $f(x)$ be a continuous function that is periodic with period 2π and is square integrable over a period, for example $I = [0, 2\pi]$, so that it admits a Fourier series representation $f(x) = \frac{1}{\sqrt{2\pi}} \sum_{k=-\infty}^{+\infty} F^k \exp(\mathrm{i}kx)$. Can you find a condition that the Fourier coefficients F^k need to satisfy to ensure that $f'(x) = \frac{\mathrm{d}f}{\mathrm{d}x}(x)$ is also square integrable on $[0, 2\pi]$.*

Proof. Knowing the conditions they must satisfy is a simple case of explicitly taking the derivative:

$$\begin{aligned}
f(x) &= \frac{1}{\sqrt{2\pi}} \sum_{-\infty}^{\infty} F^k exp(ikx) \\
f(x)' &= (\frac{1}{\sqrt{2\pi}} \sum_{-\infty}^{\infty} F^k exp(ikx))' \\
&= \frac{1}{\sqrt{2\pi}} \sum_{-\infty}^{\infty} (F^k exp(ikx))' \\
&= \frac{1}{\sqrt{2\pi}} \left(\sum_{-\infty}^{\infty} (F^k)' exp(ikx) + \sum_{-\infty}^{\infty} F^k (exp(ikx))' \right) \\
&= \frac{1}{\sqrt{2\pi}} \sum_{-\infty}^{\infty} ikF^k (exp(ikx))
\end{aligned} \tag{6}$$

s.t the following must hold if we want f' $\in L^2$: $(F^k)' \wedge kF^k \in L^2$

□

Theorem 5.2. *Let $f(x)$ be a continuous function that is periodic with period 2π and is square integrable over a period, for example $I = [0, 2\pi]$, so that it admits a Fourier series representation $f(x) = \frac{1}{\sqrt{2\pi}} \sum_{k=-\infty}^{+\infty} F^k \exp(\mathrm{i}kx)$. Can you express $f(x)$ as a linear combination of* sin *and* cos *functions, namely as*

$$f(x) = A^0 + \sum_{k=1}^{+\infty} A^k \cos(kx) + \sum_{k=1}^{+\infty} B^k \sin(kx).$$

What is the value of the coefficients A^0, A^k and B^k (for $k = 1, 2, \dots$) expressed as an integral over $f(x)$?

Proof. using the euler form of complex numbers :

$$e^{ikx} = cos(kx) + isin(kx)$$

it is rather trivial to expand and find these coefficients where A^0 is simply the case of k = 0 (duh), and A^k and B^k are defined as:

$$\begin{aligned} A^k &= \int_0^{2\pi} f(x)cos(kx)dx \\ B^k &= \int_0^{2\pi} f(x)sin(-kx)dx \end{aligned} \tag{7}$$

where through simple symmetry arguments of odd and even functions we know that $A^k = 0$ if f(x) is odd, and $B^k = 0$ if f(x) is even. □

Theorem 5.3. *Let* $\{p_n(x), n = 0, 1, 2, \ldots\}$ *be a family of real-valued univariate polynomials, where* $p_n(x)$ *is of degree* n*, that is orthogonal* ($\langle p_n, p_m\rangle_w = N_n\delta_{n,m}$) *with respect to an inner product*

$$\langle p, q\rangle_w = \int_I w(x)p(x)q(x)\mathrm{d}x$$

on the interval I (which could be finite or infinite). Prove that this family of polynomials must satisfy a recursion relation of the form

$$b_{n+1}p_{n+1}(x) + a_n p_n(x) + c_{n-1}p_{n-1}(x) = xp_n(x).$$

Proof. Since $xp_n(x)$ is a polynomial of degree n + 1, it can be written as a linear combination of the polynomials $\{1, p_1(x), \ldots, p_n(x), p_{n+1}(x)\}$. However, given that $p_n(x)$ is orthogonal to any polynomial q of degree $k \leq n-2$, due to the symmetry of the inner product formula

$$\int_I w(x)(q(x))(xp_n(x))dx = \int_I w(x)(xq(x))p_n(x)dx$$

Combining these two observations leads to the result that orthogonal polynomials are governed by a recursion relation of the form

$$b_{n+1}p_{n+1}(x) + a_n p_n(x) + c_{n-1}p_{n_1}(x) = xp_n(x)$$

as

$$xp_n(x) = \sum_{k=0}^{n+1} a^k p_k(x) = \sum_{k=n-1}^{n+1} a^k p_k(x) \tag{8}$$

where the last equality follows from that orthogonality relation saying any expansion coefficient for $k \leq n-2$ is zero

Slightly clearer derivation:

Let us observe polynomials p(x) and q(x) of degree $\leq n$. These are simply linear combinations of the following form: $f(x) = \sum_{i=0}^{n} \alpha_i x_i$. Thus taking the inner product between these two will result in 0 as they are orthogonal polynomials. Let's see what happens when we take $\langle p_k(x), xp_n(x)\rangle$. This is equal to $\langle xp_k(x), p_n(x)\rangle$. The bra side is of degree k+1 and the ket side is of degree n, thus this inner product will equal zero when $k < n-1$. We find that the only terms surviving are those of order n-1, n and n+1 (since all other coefficients must equal zero as $\langle p_k(x), p_k(x)\rangle \neq 0$). Which allows us to write $xp_n(x)$ as:

$$xp_n(x) = b_{n+1}p_{n+1}(x) + a_n p_n(x) + c_{n-1}p_{n-1}(x)$$

□

Theorem 5.4. *Let $\{p_n(x), n = 0, 1, 2, \ldots\}$ be a family of real-valued univariate polynomials, where $p_n(x)$ is of degree n, that is orthogonal ($\langle p_n, p_m\rangle_w = N_n\delta_{n,m}$) with respect to an inner product*

$$\langle p, q\rangle_w = \int_I w(x)p(x)q(x)\mathrm{d}x$$

on the interval I (which could be finite or infinite). Prove that $p_n(x)$ must have exactly n distinct (and thus simple) roots in the interval I.

Proof. Let $x_1, x_2, \ldots, x_m$ be the distinct roots of odd order of $p_n(x)$ which lie within the interval I, with thus $m \leq n$. Define $q(x) = (x-x_1)(x-x_2)\ldots(x-x_m)$, so that $q(x)p_n(x)$ must be a function with constant sign for all $x \in I$. This implies that $\langle q, p_n\rangle = \int_I w(x)q(x)p_n(x)dx$ cannot be zero. However, if $m < n$, then q is a linear combination of $\{p_0, p_1, \ldots, p_m\}$ and must be orthogonal to p_n. Hence, the only possibility is that $m = n$, which implies that $p_n(x)$ has n simple roots lying strictly within the interval I, and in particular cannot have complex roots. □

Theorem 5.5. *The Legendre polynomials $\{P_n(x), n = 0, 1, 2, \ldots\}$ are defined by the generating function*

$$\frac{1}{\sqrt{1-2tx+t^2}} = \sum_{n=0}^{+\infty} P_n(x)t^n$$

Prove that the Legendre polynomials satisfy the recursion relation

$$(n+1)P_{n+1}(x) + nP_{n-1}(x) = (2n+1)xP_n(x)$$

Proof. page 154 exercise (use derivative of generating function) (will type this out soon..) □

Theorem 5.6. *Using the recursion relation for the Legendre polynomials $P_n(x)$ from the previous question, together with $P_0(x) = 1$ and $P_1(x) = 1$, prove that they satisfy the Rodrigues formula*

$$P_n(x) = \frac{1}{2^n n!}\frac{\mathrm{d}^n}{\mathrm{d}x^n}\left(x^2-1\right)^n.$$

Proof. page 154 plop functions in recursion relation □

Theorem 5.7. *The Hermite polynomials* $\{H_n(x), n = 0, 1, 2, \dots\}$ *are defined by the generating function*

$$\exp\left(2xt - t^2\right) = \sum_{n=0}^{+\infty} \frac{H_n(x)}{n!} t^n$$

Show that the Hermite polynomials satisfy

$$\langle H_n(x), H_m(x) \rangle_w = \int_{-\infty}^{+\infty} \mathrm{e}^{-x^2} H_m(x) H_n(x) \mathrm{d}x = 2^n n! \sqrt{\pi} \delta_{m,n}$$

Proof. The Rodriguez representation tells us that:

$$H_n(x) = (-1)^n \mathrm{e}^{x^2} \frac{\mathrm{d}^n}{\mathrm{d}x^n} \mathrm{e}^{-x^2}$$

Using this we can rewrite $\langle H_n(x), H_m(x) \rangle$ as:

$$\langle H_n(x), H_m(x) \rangle_w = (-1)^n \int_{-\infty}^{+\infty} H_m(x) \frac{\mathrm{d}^n}{\mathrm{d}x^n} \mathrm{e}^{-x^2} \mathrm{d}x$$

Let us look at the case where $m \neq n$, and let us assume $n > m$ (note that it doesn't matter whether you choose $m > n$ or $m < n$, just substitute the polynomial of the highest order):
Integrating by parts n times yields:

$$(-1)^n \int_{-\infty}^{+\infty} H_m(x) \frac{\mathrm{d}^n}{\mathrm{d}x^n} \mathrm{e}^{-x^2} \mathrm{d}x = (-1)^{2n} \int_{-\infty}^{+\infty} \frac{\mathrm{d}^n}{\mathrm{d}x^n} H_m(x) \mathrm{e}^{-x^2} \mathrm{d}x = 0$$

Since $H_m(x)$ is a polynomial of order m, taking its nth derivative will result in 0 since $n > m$.

Now we need to look at the case where $n = m$. Hermite polynomials also possess the neat property that: $\frac{\mathrm{d}^n}{\mathrm{d}x^n} H_n(x) = 2n \frac{\mathrm{d}^{n-1}}{\mathrm{d}x^{n-1}} H_{n-1}(x) = 2^n n!$
After n integrations by parts we obtain:

$$(-1)^{2n} \int_{-\infty}^{+\infty} \frac{\mathrm{d}^n}{\mathrm{d}x^n} H_n(x) \mathrm{e}^{-x^2} \mathrm{d}x = 2^n n! \int_{-\infty}^{+\infty} \mathrm{e}^{-x^2} \mathrm{d}x = 2^n n! \sqrt{\pi}$$

Which completes the proof. □

Theorem 5.8. *The Hermite polynomials* $\{H_n(x), n = 0, 1, 2, \dots\}$ *are defined by the generating function*

$$\exp\left(2xt - t^2\right) = \sum_{n=0}^{+\infty} \frac{H_n(x)}{n!} t^n$$

Show that the Hermite polynomials satisfy the recursion relation

$$H_{n+1}(x) + 2nH_{n-1}(x) = 2xH_n(x).$$

Proof. page 155 same as theorem 6.6 □

Theorem 5.9. *Using the recursion relation from the previous question and* $H_0(x) = 1, H_1(x) = 2x$*, show that the Hermite polynomial* $H_n(x)$ *satisfies the Rodrigues representation*

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2}.$$

Proof. page 156 □

Theorem 5.10. *The Laguerre polynomials* $\{L_{nt}(x), n = 0, 1, 2, \ldots\}$ *are defined by the generating function*

$$\frac{1}{1-t} \exp\left(-x\frac{t}{1-t}\right) = \sum_{n=0}^{+\infty} L_n(x)t^n.$$

Show that the Laguerre polynomials satisfy

$$\langle L_n(x), L_m(x)\rangle_w = \int_0^{+\infty} e^{-x} L_m(x)L_n(x)dx = \delta_{m,n}$$

Proof. page 155 □

Theorem 5.11. *The Laguerre polynomials* $\{L_n(x), n = 0, 1, 2, \ldots\}$ *are defined by the generating function*

$$\frac{1}{1-t} \exp\left(-x\frac{t}{1-t}\right) = \sum_{n=0}^{+\infty} L_n(x)t^n$$

Show that the Laguerre polynomials satisfy the recursion relation

$$(n+1)L_{n+1}(x) + nL_{n-1}(x) = (2n+1-x)L_n(x).$$

Proof. page 156 same as theorem 6.5 □

Theorem 5.12. *Using the recursion relation from the previous question and* $L_0(x) = 1, L_1(x) = 1- x$*, show that the Laguerre polynomial* $L_n(x)$ *satisfies the Rodrigues representation*

$$L_n(x) = \frac{e^x}{n!} \frac{d^n}{dx^n} \left(x^n e^{-x}\right)$$

Proof. page 156 same as theorem 6.6 □

Theorem 5.13. *The Chebyshev polynomials are given by* $T_n(x) = \cos(n \arccos(x))$*. Show that they satisfy the recursion relation* $T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$*. Work out* $T_0(x)$ *and* $T_1(x)$ *and use this to show that* $T_n(x)$ *is indeed a polynomial of degree* n*.*

Proof. page 156 - 157 □

Theorem 5.14. *The Chebyshev polynomials are given by* $T_{nt}(x) = \cos(n \arccos(x))$. *Show that*

$$\langle T_n, T_m \rangle_w = \int_{-1}^{+1} \frac{1}{\sqrt{1-x^2}} T_n(x) T_m(x) \mathrm{d}x = \delta_{n,m} \begin{cases} \pi, & n = 0 \\ \frac{\pi}{2}, & n > 0 \end{cases}.$$

Proof. page 156 - 157 □

Theorem 5.15. *Using the recursion relation of the Chebyshev polymials from question 13 , show that they are generated via*

$$\frac{1-xt}{1-2xt+t^2} = \sum_{n=0}^{+\infty} T_n(x) t^n.$$

Proof.

$$T_n(\cos\theta) = \cos n\theta = Re\{e^{in\theta}\}$$

$$\begin{aligned} \sum_{n=0}^{+\infty} (e^{i\theta} t)^n &= \frac{1}{1-e^{i\theta}t} \text{convergence? We're physicists, everything converges if you wait long enough} \\ &= \frac{1}{1-e^{i\theta}t} \cdot \frac{1-e^{-i\theta}t}{1-e^{-i\theta}t} \\ &= \frac{1-e^{-i\theta}t}{1-2t\cos\theta+t^2} \end{aligned}$$

Taking only the real part and substituting $\cos\theta$ for x yields the wanted generating function. □

Theorem 5.16. *Given an integral operator* $\hat{A} \in \mathrm{End}\left(L^2(I)\right)$ *that acts as* $(\hat{A}f)(x) = \int_1 A(x,y) f(y)\ dy$. *Show that* $\hat{A}$ *is a bounded operator if (sufficient condition)*

$$\int_1 \int_1 |A(x,y)|^2\ \mathrm{d}x\ \mathrm{d}y < \infty.$$

Proof. For integral operators, we can easily show that

$$\begin{aligned} \|\hat{A}f\|^2 &= \int_I |\int_I A(x,y) f(y) dy|^2 dx \\ &\leq \int_I \left(\int_I |A(x,y) f(y) dy| \right)^2 dx \\ &= \int_I \left(\int_I |A(x,y)||f(y)| dy \right)^2 dx \\ &\leq \int_I \left(\int_I |A(x,y)|^2 dy \right) \left(\int_I |f(y)|^2 dy \right) dx \\ &= \left(\int_I \int_I |A(x,y)|^2 dy dx \right) \|f\|^2 \end{aligned}$$

where in the first transition, we applied the Cauchy-Schwarz inequality for every fixed value of x. Hence, a sufficient condition for $\hat{A}$ to be bounded is that the integral between the parentheses on the last line is finite. This integral can be interpret as the (squared) 2-norm of the kernel $A(x,y)$, when interpreted as a function in $L^2(I \times I)$ □

Theorem 5.17. *Given an integral operator $\hat{A} \in \mathrm{End}\left(L^2(I)\right)$ that acts as $(\hat{A}f)(x) = \int_1 A(x,y)f(y)\ dy$ and that is bounded (thus satisfying the condition from the previous question). Can you impose a (sufficient) condition on the kernel $A(x,y)$ so that $\hat{A}$ is a symmetric operator. Is it then also self-adjoint?*

Proof. Writing the inner product explicitly:

$$\begin{aligned}\langle \hat{A}f, f\rangle &= \int\int \overline{A(x,y)f(y)}f(x)dydx \\ &= \int\int \overline{f(y)}[\overline{A(x,y)}f(x)]dydx\end{aligned} \tag{9}$$

imposing A to be symmetric is then equivalent to the following condition:

$$\int \overline{A(x,y)}dy = \int A(x,y)dx$$

which is equivalent to saying:

$$\overline{A(x,y)} = A(y,x)$$

This operator is then also self adjoint as they have the same domain □

Theorem 5.18. *Given two operators $\hat{A}, \hat{B}$ End (H) on some infinite-dimensional (separable) Hilbert space H. Show that, if $[\hat{A},\hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}$ acts as the identity (on vectors that lie within its domain), then it is impossible for both operators $\hat{A}$ and $\hat{B}$ to have a finite operator norm.*

Proof. if $\hat{A}\hat{B} - \hat{B}\hat{A} = \hat{1}$, implies that

$$\hat{A}\hat{B}^n - \hat{B}^n\hat{A} = n\hat{B}^{n-1}$$

However, because of subadditivity and submultiplicativity of the operator norm, we find

$$n\|\hat{B}^{n-1}\| \le \|\hat{A}\hat{B}^n\| + \|\hat{B}^n\hat{A}\| \le 2\|\hat{A}\|\|\hat{B}^n\| \le 2\|\hat{A}\|\|\hat{B}\|\|\hat{B}^{n-1}\|$$

from which we obtain $n \le 2\|\hat{A}\|\|\hat{B}\|$. As this must hold for any n, it is clearly inconsistent with the assumption that both operators are bounded. □

Theorem 5.19. *On the Hilbert space of (equivalence classes of) square integrable functions $L^2(I)$ (with standard inner product) on a compact interval $I = [a,b]$, consider the operator $\hat{P}$ with action $(Pf)(x) = -\mathrm{i}f'(x)$ and domain $\mathcal{D}_{\hat{P}} = \left\{f \in L^2(I) \mid f' \in L^2(I), f(a) = f(b) = 0\right\}$. Is this operator symmetric? Is it self-adjoint?*

Proof. page 165 □

Theorem 5.20. *On the Hilbert space of (equivalence classes of) square integrable functions* $L^2(I)$ *(with standard inner product) on a compact interval* $I = [a, b]$*, consider the operator* $\hat{P}$ *action* $(\hat{P}f)(x) = -\mathrm{i}f'(x)$ *and domain* $\mathcal{D}_{\hat{P}} = \left\{ f \in L^2(I) \mid f' \in L^2(I), f(a) = -f(b) \right\}$*. Is this operator symmetric? Is it self-adjoint?*

Proof. page 165 □

Theorem 5.21. *On the Hilbert space of (equivalence classes of) square integrable functions* $L^2(I)$ *(with standard inner product) on a compact interval* $I = [a, b]$*, consider the operator* $\hat{P}$ *with action* $(\hat{P}f)(x) = -\mathrm{i}f'(x)$ *and domain* $\mathcal{D}_{\hat{P}} = \left\{ f \in L^2(I) \mid f' \in L^2(I), f(a) = 2f(b) \right\}$*. Is this operator symmetric? Is it self-adjoint?*

Proof. page 165 □

Theorem 5.22. *On the Hilbert space of (equivalence classes of) square integrable functions* $L^2(I)$ *(with standard inner product) on a compact interval* $I = [-a, +a]$*, consider the operator* $\hat{A}$ *with action* $(\hat{A}f)(x) = \left(x^2 - 1\right) f(x)$ *and domain* $\mathcal{D}_{\hat{A}} = L^2(I)$*. Is this operator symmetric? Is it self-adjoint? Can you guess what the (discrete, continuous and residual) spectrum of* $\hat{A}$ *is; try to explain or motivate your answer (in words)?*

Proof. exercise □

6 Fourier calculus and distributions

Theorem 6.1. *Let the Fourier transform of a function $f(x) \in L^1(\mathbb{R})$ be given by*

$$\widehat{f}(\xi) = \int_{-\infty}^{+\infty} f(x)\mathrm{e}^{-\mathrm{i}2\pi\xi x}\ \mathrm{d}x.$$

Prove that, for $g(x) = f(x)\mathrm{e}^{-\mathrm{i}2\pi a x}$, it holds that $\widehat{g}(\xi) = \hat{f}(\xi + a)$.

Proof.

$$\begin{aligned}
\widehat{g}(\xi) &= \int_{-\infty}^{+\infty} g(x)\mathrm{e}^{-\mathrm{i}2\pi\xi x}\ \mathrm{d}x \\
&= \int_{-\infty}^{+\infty} f(x)e^{-i2\pi x a}\mathrm{e}^{-\mathrm{i}2\pi\xi x}\ \mathrm{d}x \\
&= \int_{-\infty}^{+\infty} f(x)\mathrm{e}^{-\mathrm{i}2\pi(\xi+a)x}\ \mathrm{d}x \\
&= \hat{f}(\xi + a)
\end{aligned} \tag{10}$$

□

Theorem 6.2. *Let the Fourier transform of a function $f(x) \in L^1(\mathbb{R})$ be given by*

$$\widehat{f}(\xi) = \int_{-\infty}^{+\infty} f(x)\mathrm{e}^{-\mathrm{i}2\pi\xi x}\ \mathrm{d}x.$$

Prove that, for $g(x) = f(x - a)$, it holds that $\widehat{g}(\xi) = \hat{f}(\xi)\mathrm{e}^{-\mathrm{i}2\pi a\xi}$.

Proof.

$$\begin{aligned}
\widehat{g}(\xi) &= \int_{-\infty}^{+\infty} g(x)\mathrm{e}^{-\mathrm{i}2\pi\xi x}\ \mathrm{d}x \\
&= \int_{-\infty}^{+\infty} f(x-a)\mathrm{e}^{-\mathrm{i}2\pi\xi x}\ \mathrm{d}x \\
&= \int_{-\infty}^{+\infty} f(u)\mathrm{e}^{-\mathrm{i}2\pi\xi(u+a)}\ \mathrm{d}u \\
&= \int_{-\infty}^{+\infty} f(u)\mathrm{e}^{-\mathrm{i}2\pi\xi u}e^{-\mathrm{i}2\pi\xi a}\ \mathrm{d}u \\
&= \hat{f}(\bar{\xi})\mathrm{e}^{-\mathrm{i}2\pi a\xi}
\end{aligned} \tag{11}$$

□

Theorem 6.3. *Let the Fourier transform of a function $f(x) \in L^1(\mathbb{R})$ be given by*

$$\widehat{f}(\xi) = \int_{-\infty}^{+\infty} f(x)\mathrm{e}^{-\mathrm{i}2\pi\bar{\xi}x}\ \mathrm{d}x.$$

Prove that, for $g(x) = \overline{f(-x)}$*, it holds that* $\widehat{g}(\xi) = \overline{\widehat{f}(\xi)}$*.*

Proof.

$$\begin{aligned}
\widehat{g}(\xi) &= \int_{-\infty}^{+\infty} g(x)\mathrm{e}^{-\mathrm{i}2\pi\xi x}\ \mathrm{d}x \\
&= \int_{-\infty}^{+\infty} \overline{f(-x)}\mathrm{e}^{-\mathrm{i}2\pi\xi x}\ \mathrm{d}x \\
&= \int_{-\infty}^{+\infty} \overline{f(-x)\mathrm{e}^{-\mathrm{i}2\pi\xi(-x)}\ \mathrm{d}x} \\
&= \overline{\widehat{f}(\bar{\xi})}
\end{aligned} \tag{12}$$

□

Theorem 6.4. *Let the Fourier transform of a function* $f(x) \in L^1(\mathbb{R})$ *be given by*

$$\widehat{f}(\xi) = \int_{-\infty}^{+\infty} f(x)\mathrm{e}^{-\mathrm{i}2\pi\bar{\xi}x}\ \mathrm{d}x.$$

Prove that, for $g(x) = f(x/a)$*, it holds that* $\widehat{g}(\xi) = |a|\widehat{f}(a\bar{\xi})$*.*

Proof. we shall consider two cases, $a > 0$ and $a < 0$ (note a = 0 is trivial)

$$\begin{aligned}
\widehat{g}(\xi) &= \int_{-\infty}^{+\infty} g(x)\mathrm{e}^{-\mathrm{i}2\pi\xi x}\ \mathrm{d}x \\
&= \int_{-\infty}^{+\infty} f(x/a)\mathrm{e}^{-\mathrm{i}2\pi\xi x}\ \mathrm{d}x \\
a > 0 & \\
&= a\int_{-\infty}^{+\infty} f(u)\mathrm{e}^{-\mathrm{i}2\pi\xi au}\ \mathrm{d}u \\
&= |a|\widehat{f}(a\xi) \\
a < 0 & \\
&= a\int_{+\infty}^{-\infty} f(u)\mathrm{e}^{-\mathrm{i}2\pi\xi au}\ \mathrm{d}u \\
&= |a|\int_{-\infty}^{+\infty} f(u)\mathrm{e}^{-\mathrm{i}2\pi\xi au}\ \mathrm{d}u \\
&= |a|\widehat{f}(a\xi)
\end{aligned} \tag{13}$$

where we note the swapping of the boundaries if $a < 0$ when we apply the absolute value, these boundaries swap again thus delivering the desired result □

Theorem 6.5. *Let the Fourier transform of a function $f(t) \in L^1(\mathbb{R})$ be given by*

$$\widetilde{f}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(t)\mathrm{e}^{+\mathrm{i}\omega t}\ \mathrm{d}t.$$

Prove that, for $h(t) = (f * g)(t) = \int_{-\infty}^{+\infty} f(s)g(s-t)\ ds$ with $f, g \in L^1(\mathbb{R})$, it holds that $\tilde{h}(\omega) = \sqrt{2\pi}\tilde{f}(\omega)\tilde{g}(\omega)$.

Proof. $h = f * g = g * f$:

$$h(x) = (f * g)(x) = \int_{-\infty}^{+\infty} f(x-y)g(y)dy = \int_{-\infty}^{+\infty} f(y)g(x-y)dy$$

It is easy to show that $h \in L^1(\mathbb{R})$. We now obtain the convolution theorem:

$$\begin{aligned}
\widehat{h}(\zeta) &= \int_{-\infty}^{+\infty} dx e^{-i2\pi\zeta x} \int_{-\infty}^{+\infty} dy f(x-y)g(y) \\
&= \int_{-\infty}^{+\infty} dy (\int_{-\infty}^{+\infty} dx e^{-i2\pi\zeta(x-y)} f(x-y))g(y)e^{-i2\pi\zeta y} \\
&= \widehat{f}(\zeta)\widehat{g}(\zeta)
\end{aligned}$$

Due to fourier conventions, can the factor change (here it is 1). □

Theorem 6.6. *Let the Fourier transform of a function $f(t) \in L^1(\mathbb{R})$ be given by*

$$\widetilde{f}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(t)\mathrm{e}^{+\mathrm{i}\omega t}\ \mathrm{d}t.$$

What is the Fourier transform $\tilde{g}(\omega)$ of the function $g(t) = f'(t)$ with $f \in L^1(\mathbb{R})$ such that also $f' \in L^1(\mathbb{R})$?

Proof. $g = f' \in L^1(\mathbb{R})$, we find:

$$\begin{aligned}
\widehat{g}(\omega) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f'(t)e^{i\omega t}dt \\
&= \frac{1}{\sqrt{2\pi}} [f(t)e^{i\omega t}]_{x=-\infty}^{x=+\infty} - i\omega \int_{-\infty}^{+\infty} f(t)e^{i\omega t}dt \\
&= -\frac{1}{\sqrt{2\pi}} i\omega \widehat{f}(\omega)
\end{aligned}$$

□

Theorem 6.7. *Let the (Plancherel)-Fourier transform of a function $f(x) \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ be given by*

$$\widehat{f}(\xi) = (\hat{F}f)(\xi) = \int_{-\infty}^{+\infty} f(x)\mathrm{e}^{-\mathrm{i}2\pi\xi x}\mathrm{d}x.$$

What does the Parseval relation $\int_{-\infty}^{+\infty} |\hat{f}(\xi)|^2 \mathrm{d}\xi = \int_{-\infty}^{+\infty} |f(x)|^2 \ \mathrm{d}x$ tell you about the nature of the Plancheral-Fourier operator $\hat{F}$?

Proof. The Parseval relation says that $|f\|_2 = \|\hat{F}(f)\|_2$. This says that the Fourier transform operator an isometric linear operator is. Which immediately implies that we habe the general Parseval relation:

$$\langle f, g\rangle = \langle \hat{F}(f), \hat{F}(g)\rangle = \langle \widehat{f}, \widehat{g}\rangle$$

□

Theorem 6.8. *Let $H_a[\varphi] = \int_a^{+\infty} \varphi(x)\mathrm{d}x$ be the (shifted) Heaviside distribution. Show that the distributional derivative $H_a' = \delta_a$, with $\delta_a[\varphi] = \varphi(a)$ the shifted Dirac distribution.*

Proof.

$$H'[\varphi] = -\int_a^{+\infty} \varphi'(x)dx = \varphi(a) = \delta[\varphi]$$

where, using partial integration of the functions 1, and φ we achieve the desired result as $\varphi(\infty) = 0$ and the derivative of a constant function is also zero st the second integral vanishes □

Theorem 6.9. *Using the result from the previous question, and the fact that the distributional derivative of the regular distribution T_f associated to $f(x) = \log(|x|)$ is the Cauchy principle value $T_f' = \mathrm{Pv}\,\frac{1}{x}$, prove the Sokhotsky-Plemelj formula*

$$\lim_{s\to 0^\dagger} \frac{1}{x \pm is} = \mathrm{Pv}\,\frac{1}{x} \mp \mathrm{i}\pi\delta(x)$$

Proof. Consider the principal branch of the logarithm function. This function is defined for all $z = \in \mathbb{C}\backslash\mathbb{R}_{\neq 0}$, or thus, for $z = re^{i\phi}$ with $r = |z| \in [0, +\infty)$ and $\phi = \arg(z) = arctan(\frac{s}{x}) \in (-\pi, +\pi)$ and act as

$$\log(z) = \log(r) + i\phi = \log\left(\sqrt{x^2 + s^2}\right) + i \arctan\left(\frac{s}{x}\right)$$

Note that:

$$\lim_{s\to 0^+} \arctan\left(\frac{s}{x}\right) = \begin{cases} 0 & x > 0 \\ \pi & x < 0 \end{cases} = \pi H(-x) \tag{14}$$

$$\lim_{s\to 0^+} \log\left(\sqrt{x^2 + s^2}\right) = \log|x| \tag{15}$$

we then find

$$\lim_{s\to 0^+} \log(x \pm is) = \log|x| \pm i\pi H(-x)$$

Thus taking the distributional derivative yields:

$$\lim_{s\to 0^+} \frac{1}{x \pm is} = Pv\frac{1}{x} \mp i\pi\delta(x)$$

This result is known as the Sokhotski-Plemelj theorem □

Theorem 6.10. *Use this result to show that the distributional Fourier transform $\hat{H}$ of the Heaviside distribution $H[\varphi] = \int_0^{+\infty} \varphi(x)\mathrm{d}x$ is given by*

$$\hat{H}(\xi) = -\frac{\mathrm{i}}{2\pi} \operatorname{Pv} \frac{1}{\xi} + \frac{1}{2}\delta(\xi).$$

Proof. $H(x) = \lim_{s\to 0^+} e^{-sx}H(x)$, we find:

$$\begin{aligned}
\widehat{H}(\xi) &= \lim_{s\to 0^+} \int_{-\infty}^{+\infty} e^{-sx}H(x)e^{-i2\pi\xi x}dx \\
&= \lim_{s\to 0^+} \int_{0}^{+\infty} e^{-(s+i2\pi\xi)x}dx \\
&= \lim_{s\to 0^+} \frac{1}{s+i2\pi\xi} \\
&= -i \lim_{s\to 0^+} \frac{1}{2\pi\xi - is} \\
&= -\frac{i}{2\pi}Pv\frac{1}{\xi} + \frac{1}{2}\delta(\xi)
\end{aligned} \tag{16}$$

□

7 Applications of Linear Differential Operators

Theorem 7.1. *Consider the second order linear differential operator $\hat{L}$ that acts as*

$$(\hat{L}f)(x) = a_2(x)\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}(x) + a_1(x)\frac{\mathrm{d}f}{\mathrm{d}x}(x) + a_0(x)f(x).$$

Define the action of the formal adjoint $\hat{L}^\dagger$with respect to the standard $L^2(\mathbb{R})$ inner product $\langle f, g\rangle = \int_{-\infty}^{+\infty} \overline{g(x)} f(x)\mathrm{d}x$ and show that

$$j(x) = \bar{g}(x)(\hat{L}f)(x) - \overline{\left(\hat{L}^\dagger g\right)(x)} f(x)$$

equals a total derivative, i.e. $j(x) = \frac{d}{dx}J(g(x), f(x))$, where you define what $J(g(x), f(x))$ is.

Proof. We define the formal adjoint as

$$\hat{L}^\dagger = \sum_{j=0}^{p} (-1)^j \hat{D}^j a_j(\hat{x})$$

or thus

$$(\hat{L}^\dagger v)(x) = \sum_{j=0}^{p} (-1)^j \frac{d^j}{dx^j}(a_j(x)v(x))$$

Consider the following equality

$$\frac{d}{dx}[\sum_{k=0}^{j-1} (-1)^k (\frac{d^k}{dx^k}v(x))(\frac{d^{j-1-k}}{dx^{j-1-k}}u(x))] = v(x)\frac{d^j u}{dx^j}(x) - (-1)^j u(x)\frac{d^j v}{dx^j}(x)$$

which follows by applying the Leibniz (read: product) rule and canceling the contribution where the additional derivative acts on $v(x)$ in the term j with the contribution where the additional derivative acts on $u(x)$ in the term $(j+1)$. If we substitute $v(x) \to a_k(x)\bar{v}(x)$ and sum for k from 0 to p, we find the Lagrange identity

$$\overline{v(x)}(\hat{L}u)(x) - \overline{(\hat{L}^\dagger v)(x)}u(x) = \frac{d}{dx}J(u(x), v(x))$$

where we have defined the bilinear concomitant (it actually is sesquilinear (fucking chad jutho))

$$J(u(x), v(x)) = \sum_{j=0}^{p}\sum_{k=0}^{j-1} (-1)^k (\frac{d^k}{dx^k}[a_j(x)\overline{v(x)}])(\frac{d^{j-1-k}}{dx^{j-1-k}}u(x))$$

□

Theorem 7.2. *Repeat the previous question with respect to a weighted inner product $\langle f, g\rangle_w = \int_{-\infty}^{+\infty} w(x)\overline{g(x)}f(x)\mathrm{d}x$. Under which conditions on $a_2(x), a_1(x)$ and $a_0(x)$ (which you can assume to be real-valued) is $\hat{L}$ formally self-adjoint?*

Proof. The formal adjoint of the derivative is given by

$$(\hat{D}^\dagger v)(x) = -\frac{1}{w(x)}\frac{d}{dx}[w(x)v(x)]$$

We also need to change the Lagrange identity to

$$w(x)\overline{v(x)}(\hat{L}^\dagger u)(x) - w(x)\overline{(\hat{L}^\dagger v)(x)}u(x) = \frac{d}{dx}J(u(x), v(x))$$

with bilinear concomitant now given by

$$J(u(x), v(x)) = \sum_{j=0}^{p}\sum_{k=0}^{j-1}(\frac{d^k}{dx^k}[w(x)a_k(x)\overline{v(x)}](\frac{d^{j-1-k}}{dx^{j-1-k}}u(x))$$

□

Theorem 7.3. *Consider the second order differential operator $\hat{L} = -\hat{D}^2$, i.e. $(\hat{L}f)(x) = f''(x)$ with a domain $\mathcal{D}_L = \{f \in L^2([a,b]) \mid f'' \in L^2([a,b]), f(a) = f(b), f'(a) = 0\}$. What is the action and domain of the adjoint $\hat{L}^\dagger$. Is $\hat{L}$ self-adjoint with these boundary conditions?*

Proof. Starting with the formal adjoint, we find its action to be

$$\hat{L}^\dagger = -\hat{D}^2$$

now using the (not so bilinear) bilinear concomitant to find the boundary conditions for $\hat{L}^\dagger$, we find:

$$[\overline{g(b)}\frac{\partial}{\partial x}(f(b)) - \frac{\partial}{\partial x}(\overline{g(b)})f(b)] - [\overline{g(a)}\frac{\partial}{\partial x}(f(a)) - \frac{\partial}{\partial x}(\overline{g(a)})f(a)] = 0$$

now using the boundary conditions in $\mathcal{D}_L$ all terms are zero if we impose following(boundary) conditions in $\mathcal{D}_{L^\dagger}$:

$$\mathcal{D}_{L^\dagger} = \{g \in L^2(I) \mid g''(x) \in L^2(I), g' \in L^2(I) \wedge \overline{g(b)} = 0\}$$

from which we see that L is not self-adjoint but it is symmetric. □

Theorem 7.4. *Consider the Sturm-Liouville operator $(\hat{L}u)(x) = \frac{\mathrm{d}}{\mathrm{d}x}\left(p(x)\frac{\mathrm{d}}{\mathrm{d}x}u(x)\right) + q(x)u(x)$. Show that it is self-adjoint with respect to the standard inner product on $L^2([a,b])$ if using the separated boundary conditions $f(a) + \alpha f'(a) = 0$ and $f(b) + \beta f'(b) = 0$ where $\alpha, \beta \in \mathbb{R}$.*

Proof. page 210 under 7.32, it can be verified that these conditions on u imply the same conditions on v in order to have a vanishing boundary term, so that $\hat{L}$ then becomes self-adjoint. □

Theorem 7.5. *Consider the first order vector-valued homogeneous differential equation*

$$\frac{\mathrm{d}z}{\mathrm{d}t}(t) = \mathrm{A}(t)z(t)$$

with $z(t) \in \mathbb{F}^n$ *and* $\mathrm{A}(t) \in \mathbb{F}^{n\times n}$*. Let* $\mathrm{Z}(t)$ *be a fundamental solution matrix, i.e. for every* $t, Z(t) \in \mathbb{F}^{n\times n}$ *such that*

$$\frac{dZ}{dt}(t) = A(t)Z(t)$$

with furthermore $\det(Z(t)) \neq 0$*, so that the column of* $Z(t)$ *represent* n *linearly independent solutions. Let* $\bar{Z}(t)$ *be another fundamental solution matrix; show that there exists a constant matrix* $C \in \mathbb{F}^{n\times n}$ *such that* $\bar{Z}(t) = Z(t)C$*.*

Proof. Which follows from

$$\begin{aligned}\frac{dC}{dt} &= \frac{d}{dt}[Z^{-1}(t)\tilde{Z}(t)] = -Z^{-1}(t)\frac{dZ}{dt}(t)Z^{-1}(t)\tilde{Z}(t) + Z^{-1}(t)\frac{d\tilde{Z}}{dt}(t) \\ &= -Z^{-1}(t)A(t)\tilde{Z}(t) + Z^{-1}(t)A(t)\tilde{Z}(t) = 0\end{aligned}$$

where the first term comes from the fact that:

$$\frac{\partial ZZ^{-1}}{\partial t} = \frac{\partial Z}{\partial t}Z^{-1} + Z\frac{\partial Z^{-1}}{\partial t} = 0$$

□

Theorem 7.6. *Consider the first order vector-valued homogeneous differential equation*

$$\frac{\mathrm{d}z}{\mathrm{d}t}(t) = \mathrm{A}(t)z(t)$$

with $z(t) \in \mathbb{F}^n$ *and* $\mathrm{A}(t) \in \mathbb{F}^{n\times n}$*, where furthermore* $\mathrm{A}(t)$ *is a periodic function of time, i.e.* $\mathrm{A}(t) = \mathrm{A}(t+T)$ *for some period* T*. Let* $\mathrm{Z}(t)$ *be a fundamental solution matrix as defined in the previous question. Show Floquet's theorem, namely that* $Z(t)$ *can be expressed as*

$$Z(t) = Q(t)\exp(Bt)$$

where $\mathrm{Q}(t)$ *is also periodic with period* T *and* B *is a constant matrix. You can use the result from the previous question.*

Proof. If Z(t) is a fundamental solution matrix, then so is $\tilde{Z}(t) = Z(t+T)$, as it satisfies

$$\frac{d\tilde{Z}(t)}{dt} = \frac{d}{dt}Z(t+T) = A(t+T)Z(t+T) = A(t)\tilde{Z}(t)$$

Hence, there exists a constant matrix C such that $Z(t+T) = Z(t)C$; it is given by

$$C = Z(t)^{-1}Z(t+T)$$

and known as the monodromy matrix. With C being invertible, it has a (non-unique) logarithm, from which we define $B = \frac{1}{T} logC$ so that $C = e^{TB}$. If we now define $Q(t) = Z(t)e^{-tB}$ then we find

$$Q(t+T) = Z(t+T)e^{-(t+T)B} = Z(t)Ce^{-TB}e^{-tB} = Q(t)$$

Hence, any fundamental solution matrix take the form $Z(t) = Q(t)e^{tB}$ with Q periodic with period T. □

Theorem 7.7. *Consider the second order scalar-valued homogeneous differential equation*

$$a_2(t)\ddot{u}(t) + a_1(t)\dot{u}(t) + a_0(t)u(t) = 0$$

and consider two solutions $u(t)$ and $v(t)$. Define the Wronskian

$$W(t) = \det\left(\begin{bmatrix} u(t) & v(t) \\ \dot{u}(t) & \dot{v}(t) \end{bmatrix}\right)$$

Prove Abel's formula $W(t) = W(t_0)\exp\left(-\int_{t_0}^t a_1(\tau)/a_2(\tau)\mathrm{d}\tau\right)$.

Proof. If we have two solutions u(t) and v(t), we find $W(t) = u(t)\ddot{v}(t) - \dot{u}(t)v(t)$ and thus

$$\dot{W}(t) = u(t)\ddot{v}(t) - \ddot{u}(t)v(t) = -\frac{a_1(t)}{a_2(t)}(u(t)\dot{v}(t) - \dot{u}(t)v(t)) = -\frac{a_1(t)}{a_2(t)}W(t)$$

$-\frac{a_1}{a_2}$ is bounded and integrable , s.t:

$$V(t) = W(t)exp(\int_{t_0}^t \frac{a_1(\tau)}{a_2(\tau)}d\tau)$$

is well defined, if we now differentiate both sides, and use our known expression for the time derivative of W(x):

$$\dot{V}(t) = \dot{W}(t) - W(t)\frac{a_1}{a_2}exp(\int_{t_0}^t \frac{a_1(\tau)}{a_2(\tau)}d\tau) = 0$$

V is thus constant, and $V(t_0) = W(t_0) = V(t)$ s.t solving the definiton of V(t) for W gives abel's formula □

Theorem 7.8. *Consider the first order vector valued initial value problem*

$$\frac{\mathrm{d}z}{\mathrm{d}t}(t) = \mathrm{A}(t)z(t) + \boldsymbol{b}(t), \quad z(0) = \xi.$$

Let $\mathrm{Z}(t, t_0)$ the principal fundamental solution of the homogeneous differential equation, i.e. for every $t_0, \mathrm{Z}(t, t_0)$ is the fundamental solution matrix (thus satisfying $\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{Z}(t, t_0) = \mathrm{A}(t)\mathrm{Z}(t, t_0)$) which satisfies the specific condition $\mathrm{Z}(t_0, t_0) = \mathrm{I}$. *Verify that the solution to the initial value problem is given by*

$$\boldsymbol{z}(t) = \mathrm{Z}(t, 0)\boldsymbol{\zeta} + \int_0^t \mathrm{Z}(t, s)\boldsymbol{b}(s)\mathrm{d}s$$

Proof. page 219, maybe a little before this as well but idk

$$
\begin{aligned}
z(t) &= Z(t,t_0)\zeta + \sum_{n=0}^{+\infty}\int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt2 \cdots \int_{t_0}^{t_{n-1}} dt_n \int_{t_0}^{t_n} d\tau A(t_1)A(t_2)\dots A(t_n)\mathbf{b}\tau \\
&= Z(t,t_0)\zeta + \int_{t_0}^{t} d\tau \sum_{n=0}^{+\infty}\int_{\tau}^{t} dt_1 \int_{\tau}^{t_1} dt2 \cdots \int_{t_0}^{t_{n-1}} dt_n A(t_1)A(t_2)\dots A(t_n)\mathbf{b}\tau \\
&= Z(t,t_0)\zeta + \int_{t_0}^{t} Z(t,\tau)\mathbf{b}(\tau)d\tau
\end{aligned}
$$

□

Theorem 7.9. *Given a second order linear differential operator $\hat{L}$ that acts as*

$$(\hat{L}u)(x) = a_2(x)\frac{\mathrm{d}^2u}{\mathrm{d}x^2}(x) + a_1(x)\frac{\mathrm{d}u}{\mathrm{d}x}(x) + a_0(x)u(x)$$

Consider the homogeneous differential equation $(\hat{L}u)(x) = 0$ with separated homogeneous boundary conditions $u(a)+\alpha u'(a) = 0$ and $u(b)+\beta u'(b) = 0$. Explain why the solution space of this completely homogeneous boundary value problem can be at most one-dimensional.

Proof. For separated boundary conditions, the solution space can at most be one-dimensional. If the two linearly independent fundamental solutions u_1 and u_2 would both satisfy $B_1[u_1] = 0$ and $B_1[u_2] = 0$, then both vectors $[u_1(a), u_1'(a)]^T$ and $[u_2(a)u_2'(a)]^T$ would be in the kernel of a two-dimensional linear form, which is one-dimensional. Hence, both vectors would be linearly dependent, which is in violation with u_1 and u_2 being linearly independent solutions, as this requires that at any point x, the Wronskian $W(x) = u_1(x)u_2'(x) - u_2(x)u_1'(x) \neq 0$. □

Theorem 7.10. *Given a second order linear differential operator $\hat{L}$ that acts as*

$$(Lu)(x) = a_2(x)\frac{\mathrm{d}^2u}{\mathrm{d}x^2}(x) + a_1(x)\frac{\mathrm{d}u}{\mathrm{d}x}(x) + a_0(x)u(x).$$

Consider the inhomogeneous differential equation $(\hat{L}u)(x) = f(x)$ with separated homogeneous boundary conditions $u(a) + \alpha u'(a) = 0$ and $u(b) + \beta u'(b) = 0$. Let $u_a(x)$ and $u_b(x)$ be solutions of the homogeneous differential equation $\left(\hat{L}u_a\right)(x) =$ Assume that $u_a(x)$ and $u_b(x)$ are linearly independent solutions, so that

$$W(x) = \det\left(\begin{bmatrix} u_a(x) & u_b(x) \\ u_a'(x) & u_b'(x) \end{bmatrix}\right) \neq 0.$$

Show that $u_f(x) = \int_a^b g(x,y)f(y)\ dy$ with

$$g(x,y) = \begin{cases} \frac{u_a(x)u_b(y)}{u_2(y)W(y)}, & a < x < y \\ \frac{u_a(y)u_b(x)}{a_2(y)W(y)}, & y < x < b \end{cases} = \frac{u_a(\min(x,y))u_b(\max(x,y))}{a_2(y)W(y)}.$$

is a particular solution to the inhomogeneous differential equation with homogeneous boundary conditions.

Proof. We can construct the fundamental solution $u_1(x)$ such that it satisfies the left boundary condition $B_1[u_1] = \alpha_1 u_1(a) + \alpha_2 u_1'(a) = 0$, namely by choosing $u_1(a) = \alpha_2$ and $u_1'(a) = -\alpha_1$ as initial conditions. We similarly construct the second fundamental solutions u_2 such that $B_2[u_2] = \beta_1 u_2(b) + \beta_2 u_2'(b) = 0$, namely by choosing $u_2(b) = \beta_2$ and $u_2'(b) = -\beta_1$ as 'initial condition' and integerating to the left we can then set

$$g(x,y) = H(y-x)c(y)u_1(x) + H(x-y)d(y)u_2(x)$$

and find

$$\begin{aligned} d(y)u_2(y) - c(y)u_1(y) &= 0 \\ d(y)u_2'(y) - c(y)u_1'(y) &= \frac{1}{a_2(y)} \end{aligned}$$

the solution can be written using the Wronskian $W(y) = u_1(y)u_2'(y) - u_1'(y)u_2(y)$ with g(x,y) the given function. □