Chapter IV: Linear quadratic optimal control

The cost (or energy supply) functional

Set $\Delta_k^q(\lambda) := \begin{bmatrix} I_q \\ \lambda I_q \\ \vdots \\ \lambda^{k-1} I_q \end{bmatrix} \in \mathbb{C}[\lambda]^{qK,q}$ and with this introduce

the notation $\Delta_K y := \Delta_K^q\left(\frac{d}{dt}\right) y = \begin{bmatrix} y \\ \dot{y} \\ \vdots \\ y^{(k-1)} \end{bmatrix} \in \mathcal{C}_\infty^{q\cdot K}$

for $y \in \mathcal{C}_\infty^q$. To save writing we may also use $\Delta y := \Delta_K y$ if K is clear from the context.

Let $P \in \mathbb{C}[\lambda]^{p,q}$ and let $z \in \mathfrak{L}e(P)$.

Here we measure the cost that z causes (or the power which is supplied to the system along z) at time point $t \in \mathbb{R}$ via an expression of the form

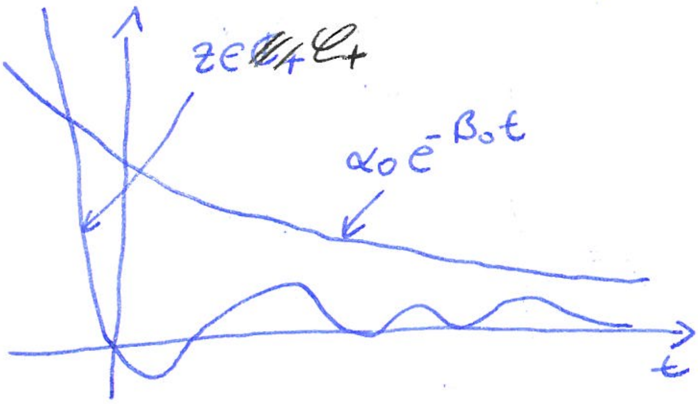
$$\begin{bmatrix} z(t) \\ \dot{z}(t) \\ \vdots \\ z^{(k-1)}(t) \end{bmatrix}^* \underbrace{\begin{bmatrix} H_{1,1} & H_{1,2} & \cdots & H_{1,K} \\ H_{1,2}^* & H_{2,2} & \cdots & H_{2,K} \\ \vdots & & & \\ H_{1K}^* & H_{2K}^* & \cdots H_{k-1,k}^* & H_{k,k} \end{bmatrix}}_{=:H} \begin{bmatrix} z(t) \\ \dot{z}(t) \\ \vdots \\ z^{(k-1)}(t) \end{bmatrix} = (\Delta_k z(t))^* H (\Delta_k z(t)),$$

where $K \in \mathbb{N}$ is some fixed constant and $H_{i,i} = H_{i,i}^*$ for $i = 1, \ldots, K$. This implies $H = H^*$.

The cost caused (or the energy supplied to the system) by z in some time interval $[t_0, t_1]$ is then given by

$$\int_{t_0}^{t_1} \underbrace{(\Delta_k z(t))^* H (\Delta_k z(t))}_{\text{power}} \underbrace{dt}_{\cdot\,\text{time}\,=\,\text{energy}} \overset{\text{Notation}}{=} \int_{t_0}^{t_1} (\Delta z)^* H (\Delta z)\, dt$$

Denote by $\mathcal{C}_+^q$ all elements of $\mathcal{C}_\infty^q$ which *decay exponentially*, i.e., those $z \in \mathcal{C}_\infty^q$ such that for every $i \in \mathbb{N}_0$ there exist $\alpha_i, \beta_i > 0$ with $\| z^{(i)}(t) \| \le \alpha_i e^{-\beta_i t}$ for $t \ge 0$.

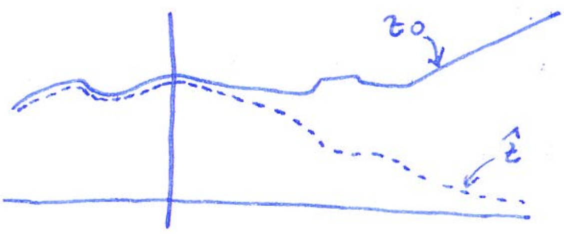


We denote by
$$\mathfrak{L}_+(P) := \left\{ z \in \mathcal{C}_+^q \;\middle|\; P\left(\tfrac{d}{dt}\right) z = 0 \right\} = \mathfrak{L}(P) \cap \mathcal{C}_+^q$$
the *(exponentially) decaying behavior*.

Then for every *(exponentially) decaying trajectory* $z \in \mathfrak{L}_+(P)$ the expression
$$\int_0^\infty (\Delta_k z(t))^* H (\Delta_k z(t))\, dt \overset{\text{Notation}}{=} \int_0^\infty (\Delta z)^* H (\Delta z)\, dt$$
is well defined.

The optimal control problem which we are going to study here is then the following:

$$(LQ) \begin{cases} \text{Given } z_0 \in \mathfrak{L}_+(P) \text{ find } \hat{z} \in \mathfrak{L}_+(P) \text{ such that} \\ \quad \hat{z}(t) = z_0(t) \text{ for } t \le 0 \text{ and} \\ \quad \displaystyle\int_0^\infty (\Delta_k \hat{z}(t))^* H (\Delta_k \hat{z}(t))\, dt = \inf_{\substack{z \in \mathfrak{L}_+(P) \\ z(t) = z_0(t),\, t \le 0}} \int_0^\infty (\Delta_k z(t))^* H (\Delta_k z(t))\, dt \end{cases}$$



In this case $\hat{z}$ is called the optimal trajectory (with respect to H).

Before we can state the central results related to (LQ) we need to introduce adjoint equations and the notion of non-negativity.

Adjoint equations

Definition: For $R \in \mathbb{C}(\lambda)^{p,q}$ we define the para-conjugate transposed $R^\sim \in \mathbb{C}(\lambda)^{q,p}$ as

$$R^\sim(\lambda) := (R(-\bar{\lambda}))^* \underset{\text{Notation.}}{=:} R^*(-\bar{\lambda}) = R(-\bar{\lambda})^*.$$

Furthermore, if $p = q$, we call R para-Hermitian if $R = R^\sim$.

For example with $R(\lambda) := \left[\lambda - a \;\vdots\; \frac{1}{\lambda - b}\right]$ where $a, b \in \mathbb{C}$ we have $R^\sim(\lambda) = \left[-\bar{\lambda} - a \;\vdots\; \frac{1}{-\bar{\lambda} - b}\right]^* = \begin{bmatrix} -\lambda - \bar{a} \\ \frac{1}{-\lambda - \bar{b}} \end{bmatrix} \in \mathbb{C}(\lambda)^{2,1}$

The $\sim$-operator behaves very much like the $*$-operator for matrices in $\mathbb{C}^{p,q}$. For example if $A \in \mathbb{C}(\lambda)^{p,p}$, $B \in \mathbb{C}(\lambda)^{p,q}$, $C \in \mathbb{C}(\lambda)^{q,r}$ we have:

1) $(A^{-1})^\sim = (A^\sim)^{-1} =: A^{-\sim}$, if A is invertible (over $\mathbb{C}(\lambda)$)

2) $(BC)^\sim = C^\sim B^\sim$

3) $(B^\sim)^\sim = B$

4) If A is para-Hermitian then also $B^\sim A B$ is.

5) If $A \in \mathbb{C}[\lambda]^{p,p}$ is polynomial and unimodular then also $A^\sim$ is.

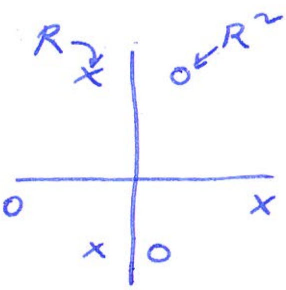
6) $\operatorname{rank}_{\mathbb{C}(\lambda)}(B) = \operatorname{rank}_{\mathbb{C}(\lambda)}(B^\sim)$

7) If B has full row rank then $B^\sim$ has full column rank.

(Proof is Homework) (S. 9, T. 1)

<u>Lemma 4.1:</u> For every $R \in \mathbb{C}(\lambda)^{p,q}$ we have

$$\mathcal{Z}(R^\sim) = -\overline{\mathcal{Z}(R)} \quad \text{and} \quad \mathcal{P}(R^\sim) = -\overline{\mathcal{P}(R)}.$$



<u>Proof:</u> Let the MacMillan form of R be

$$R = S\begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}T \quad ; D = \begin{bmatrix} \alpha_1/\beta_1 & & \\ & \ddots & \\ & & \alpha_r/\beta_r \end{bmatrix}$$

and assume that $\alpha_1,\dots,\alpha_r \in \mathbb{C}[\lambda]$ are given as

$$\alpha_i(\lambda) = c_i \prod_{k=1}^{m_i} (\lambda - a_k^{(i)}) \quad \text{for } i = 1,\dots,r$$

where $c_i, a_k^{(i)} \in \mathbb{C}$, $m_i \in \mathbb{N}_0$. Then we have

$$\mathcal{Z}(R^\sim) = \mathcal{Z}(\underbrace{T^\sim \begin{bmatrix} D^\sim & 0 \\ 0 & 0 \end{bmatrix} S^\sim}_{\text{This is a MacMillan form of } R^\sim}) = \mathcal{Z}\left(\begin{bmatrix} \frac{\alpha_1^\sim}{\beta_1^\sim} & & \\ & \ddots & \\ & & \frac{\alpha_r^\sim}{\beta_r^\sim} \end{bmatrix}\right)$$

$$= \bigcup_{i=1}^{r} \mathcal{Z}(\alpha_i^\sim) = \bigcup_{i=1}^{r} \mathcal{Z}\Big(\bar{c}_i \prod_{k=1}^{m_i} (-\lambda - \overline{a_k^{(i)}})\Big) = \bigcup_i \bigcup_{k=1}^{m_i} \{\overline{-a_k^{(i)}}\}$$

$$= -\overline{\bigcup_i \bigcup_{k=1}^{m_i} \{a_k^{(i)}\}} = -\overline{\bigcup_i \mathcal{Z}\Big(c_i \prod_{k=1}^{m_i} (\lambda - a_k^{(i)})\Big)} = -\overline{\bigcup_i \mathcal{Z}(\alpha_i)}$$

$$= -\overline{\mathcal{Z}(D)} = -\overline{\mathcal{Z}(R)}$$

The other identity can be shown analogously ∎

Lemma 4.1 especially shows that for para-Hermitian $R = R^\sim$ the zeros $\mathcal{Z}(R) \underset{R=R^\sim}{=} \mathcal{Z}(R^\sim) = \overline{-\mathcal{Z}(R)}$

(and also the poles) are symmetric with respect to the imaginary axis.

Definition: Let $P(\lambda) = \left(\sum_{i=0}^{k} \lambda^i \underbrace{P_i}_{\in \mathbb{C}^{p,q}} \right) \in \mathbb{C}[\lambda]_k^{p,q}$ and let $\ell \in \mathbb{N}_0$.

Then we define the ℓ-times shifted polynomial

$P^{<\ell>} \in \mathbb{C}[\lambda]_{k-\ell}^{p,q}$ through

$$P^{<\ell>}(\lambda) := \sum_{i=\ell}^{k} \lambda^{i-\ell} P_i = \sum_{i=0}^{k-\ell} \lambda^i P_{i+\ell} .$$

We have $P = P^{<0>}$ for all $P \in \mathbb{C}[\lambda]^{p,q}$.

Example: Let $M, D, K, F, G \in \mathbb{C}^{p,q}$

For $P(\lambda) = \lambda^2 M + \lambda D + K$ we have

$$P^{<1>}(\lambda) = \lambda M + D, \quad P^{<2>}(\lambda) = M, \quad P^{<\ell>}(\lambda) = 0 \ \forall \ell \geq 3$$

For $P(\lambda) = \lambda F + G$ we have $P^{<1>}(\lambda) = F$, $P^{<\ell>}(\lambda) = 0 \ \forall \ell \geq 2$.

Lemma 4.2:

Let $P \in \mathbb{C}[\lambda]^{p,q}$, $y \in \mathcal{C}_\infty^p$, $z \in \mathcal{C}_\infty^q$, and let $t_0, t_1 \in \mathbb{R}$ be with $t_0 \leq t_1$. Then we have

$$\int_{t_0}^{t_1} z^*(t) \left(P^\sim\left(\tfrac{d}{dt}\right) y(t) \right) dt = \int_{t_0}^{t_1} \left(P\left(\tfrac{d}{dt}\right) z(t) \right)^* y(t)\, dt$$

$$+ \sum_{\ell=1}^{\infty} (-1)^\ell \left(P^{<\ell>}\left(\tfrac{d}{dt}\right) z(t) \right)^* y^{(\ell-1)}(t) \Bigg|_{t_0}^{t_1} .$$

where the infinite series is indeed only a finite sum since for ℓ big enough $P^{<\ell>} = 0$

<u>Proof:</u> Write P in the form $P(\lambda)=\sum_{i=0}^{K} \lambda^i P_i$. Using repeated partial integration we see that for $i=0,\dots,K$ we have

$$\int_{t_0}^{t_1} \underbrace{z^*(t)}_{v} \underbrace{P_i^* y^{(i)}(t)}_{u} \, dt = z^*(t) P_i^* y^{(i-1)}(t)\Big|_{t_0}^{t_1} - \int_{t_0}^{t_1} \dot{z}^*(t) P_i^* y^{(i-1)}(t)\,dt$$

$$= (-1)^0 z^*(t) P_i^* y^{(i-1)}(t)\Big|_{t_0}^{t_1} + (-1)^1 \dot{z}^*(t) P_i^* y^{(i-2)}(t)\Big|_{t_0}^{t_1} + \dots$$

$$\dots + (-1)^2 \int_{t_0}^{t_1} \ddot{z}^*(t) P_i^* y^{(i-2)}(t)\,dt = \dots =$$

$$= \sum_{j=0}^{i-1} (-1)^j (z^{(j)}(t))^* P_i^* y^{(i-j-1)}(t)\Big|_{t_0}^{t_1} + (-1)^i \int_{t_0}^{t_1} (z^{(i)}(t))^* P_i^* y(t)\,dt$$

Using the formula $P^{\sim}(\lambda) = \sum_{i=0}^{K} (-1)^i \lambda^i P_i^*$ (Homework S.9, T.1)

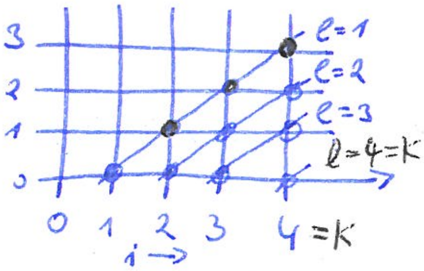
This implies

$$\int_{t_0}^{t_1} z^*(t) \left(P^{\sim}\left(\frac{d}{dt}\right) y(t)\right) dt = \sum_{i=0}^{K} (-1)^i \int_{t_0}^{t_1} z^*(t) P_i^* y^{(i)}(t)\,dt$$

$$= \sum_{i=0}^{K} (-1)^i \left[\sum_{j=0}^{i-1} (-1)^j (z^{(j)}(t))^* P_i^* y^{(i-1-j)}(t)\Big|_{t_0}^{t_1} + (-1)^i \int_{t_0}^{t_1} \underbrace{(z^{(i)}(t))^* P_i^*}_{=(P_i z^{(i)}(t))^*} y(t)\,dt \right]$$

$$= \sum_{i=0}^{K} \sum_{j=0}^{i-1} (-1)^{i+j} (z^{(j)}(t))^* P_i^* y^{(i-1-j)}(t)\Big|_{t_0}^{t_1} + \int_{t_0}^{t_1} \left(P\left(\frac{d}{dt}\right) z(t)\right)^* y(t)\,dt$$

Finally, using reindexing as in 3 we see that for all $f: (\mathbb{N}_0)^2 \to \mathbb{C}$ we have



$$\sum_{i=0}^{K} \sum_{j=0}^{i-1} f(i,j) = \sum_{\ell=0}^{K} \sum_{m=0}^{K-\ell} f(\ell+m, m)$$

and thus

$$\sum_{i=0}^{K}\sum_{j=0}^{i-1}(-1)^{i+j}\left(z^{(j)}(t)\right)^* P_i^*\, y(t)^{(i-1-j)}\Big|_{t_0}^{t_1}$$

$$=\sum_{\ell=1}^{K}\sum_{m=0}^{K-\ell}(-1)^{\ell+m+m}\left(z^{(m)}(t)\right)^* P_{\ell+m}^*\, y^{(\ell+m-1-m)}(t)\Big|_{t_0}^{t_1}$$

$$=\sum_{\ell=1}^{K}(-1)^{\ell}\underbrace{\left(\sum_{m=0}^{K-1} P_{\ell+m}\, z^{(m)}(t)\right)}_{=P^{<\ell>}(\frac{d}{dt})z(t)} y^{(\ell-1)}(t)\Big|_{t_0}^{t_1}$$

which proves the claim. ☒

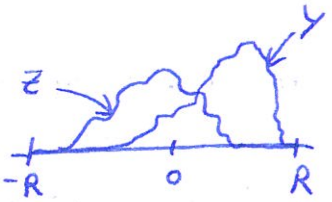
On the vector space $\mathcal{C}_c^q$ the expression

$$\langle f, g\rangle_{\mathcal{C}_c^q} := \int_{-\infty}^{\infty} f^*(t)\, g(t)\, dt$$

defines a scalar product (pre-Hilbert space), for all $q \in \mathbb{N}$.

With the notation of Lemma 4.2 we have for the linear operator $P(\frac{d}{dt}) : \mathcal{C}_c^q \to \mathcal{C}_c^p$ that

$$\langle P(\tfrac{d}{dt})z, y\rangle_{\mathcal{C}_c^p} = \int_{\cancel{-\infty}\ -R}^{\cancel{\infty}\ R \text{ (compact support)}} \left(P(\tfrac{d}{dt})z(t)\right)^* y(t)\, dt$$



$$\overset{\text{Lemma 4.2.}}{=} \int_{\cancel{-R}\ -\infty}^{\cancel{R}\ \infty} z^*(t)\left(P^{\sim}(\tfrac{d}{dt})y(t)\right)dt + \sum_{k=1}^{\infty}(-1)^k \Big(\underbrace{P^{<k>}(\tfrac{d}{dt})z(t)}_{=0\ (\text{boundary})}\Big)^* \underbrace{y^{(k-1)}(t)}_{=0}\Big|_{-R}^{R}$$

$$= \langle z, P^{\sim}(\tfrac{d}{dt})y\rangle_{\mathcal{C}_c^q}.$$

This means (by definition; as given in functional analysis) that

$$P^{\sim}(\tfrac{d}{dt}) : \mathcal{C}_c^p \to \mathcal{C}_c^q$$

is the adjoint operator of $P(\frac{d}{dt})$ (w.r.t. the given pre-Hilbert space).

Non-negativity

If we want to solve a quadratic minimization problem of the form

$$(*) \quad \min_{x \in \mathbb{C}^n} f(x) \text{ with } f(x) := x^* H x \quad , H = H^* \in \mathbb{C}^{n,n}$$

then this is only possible if and only if $0 \leq x^* H x \ \forall x \in \mathbb{C}^n$ since otherwise there exists a $\hat{x}$ with $\hat{x} H x < 0$ and then for $\alpha \in \mathbb{C}$ we have $(\alpha \hat{x})^* H (\alpha \hat{x}) = |\alpha|^2 \hat{x}^* H \hat{x} \xrightarrow{|\alpha| \to \infty} -\infty$ which implies that $(*)$ can not have a minimum (also compare Homework S.9, T.6)

A similar condition is also necessary for the more complicated problem (LQ).

Definition: Let $P \in \mathbb{C}[\lambda]^{p,q}$, $H = H^* \in \mathbb{C}^{kq,kq}$. Then we call H non-negative w.r.t. P if

$$0 \leq \int_{\cancel{-\infty}\,0}^{\infty} (\Delta_k z(t))^* H (\Delta_k z(t)) dt$$

for all $z \in \mathcal{L}e_+(P)$ with $z(t) = 0$ for $t \leq 0$

If $H \geq 0$ this condition is always fulfilled

The optimality system

Let $P \in \mathbb{C}[\lambda]^{p,q}$ and $H = H^* \in \mathbb{C}^{kq,kq}$. In this section we will show (in two theorems) that (LQ) is solvable if and only if H is non-negative w.r.t. P; and in this case the solutions $\hat{z}$ of (LQ) are given as the solutions of the *optimality system*

$$\mathcal{L}_+\left(\begin{bmatrix} 0 & P \\ P^\sim & \tilde{H} \end{bmatrix}\right) \ni \begin{bmatrix} \hat{u} \\ \hat{z} \end{bmatrix} \quad \text{with} \quad \hat{H}(\lambda) := (\Delta_k^q(\lambda))^\sim H (\Delta_k^q(\lambda))$$

where $\hat{u}$ is the so-called *Lagrange-multiplier* or *co-state*. Note that $N = N^\sim$ is para-Hermitian.

<u>Theorem 4.3:</u> Let $z_0 \in \mathcal{L}_+(P)$. (as in (LQ))
If H is non-negative w.r.t. P and $\begin{bmatrix} \hat{u} \\ \hat{z} \end{bmatrix} \in \mathcal{L}_+\left(\begin{bmatrix} 0 & P \\ P^\sim & \tilde{H} \end{bmatrix}\right)$ fullfills $z_0(t) = \hat{z}(t)$ for $t \le 0$ then $\hat{z}$ solves (LQ)

<u>Proof:</u> Let $v \in \mathcal{L}_+(P)$ be arbitrary with $v(t) = z_0(t)$ for $t \le 0$. Then for $s \in \mathbb{R}$ we have

$$z_s(t) := \hat{z}(t) + \underbrace{s(v(t) - \hat{z}(t))}_{=:y \in \mathcal{L}_+(P)} \in \mathcal{L}_+(P)$$

and for $t \le 0$

$$z_s(t) = \underbrace{\hat{z}(t)}_{=z_0(t)} + s(\underbrace{v(t)}_{=z_0(t)} + \underbrace{\hat{z}(t)}_{=z_0(t)}) = z_0(t)$$

Thus z_s is in the set over which the infimum is taken in (LQ).

Define $\phi_v : \mathbb{R} \to \mathbb{R}$ through

$$\phi_v(s) := \int_0^\infty (\Delta z_s)^* H (\Delta z_s)\, dt$$

$$= \int_0^\infty (\Delta \hat{z})^* H (\Delta \hat{z}) + 2s\, \mathrm{Re}\left\{(\Delta y)^* H (\Delta \hat{z}\right\} + s^2 (\Delta y)^* H (\Delta y)\, dt$$

$$= \int_0^\infty (\Delta \hat{z})^* H (\Delta \hat{z})\, dt + 2s \int_0^\infty \mathrm{Re}\left\{(\Delta y)^* H (\Delta \hat{z})\right\} dt + s^2 \int_0^\infty (\Delta y)^* H (\Delta y)\, dt$$

In the following we are going to show that ϕ_v has a minimum at $s = 0$. Since $z_s|_{s=1} = v$ and $z_s|_{s=0} = \hat{z}$ this means that in the sense of (LQ) the trajectory $\hat{z}$ is better than (or at least as good as) v. Since v is assumed to be arbitrary this proves that $\hat{z}$ is indeed an optimal solution of (LQ).

Since ϕ_v is a quadratic function in s it suffices to show that $\frac{d}{ds}\phi_v(0) = 0$ and $\left(\frac{d}{ds}\right)^2 \phi_v(0) \geq 0$.

We have

$$\frac{d}{ds}\phi_v(0) = 2 \int_0^\infty \mathrm{Re}\left\{(\Delta y)^* H (\Delta \hat{z})\right\} dt$$

$$= 2\, \mathrm{Re}\left\{\int_0^\infty (\Delta y)^* H (\Delta \hat{z})\, dt\right\}.$$

Since for $t \leq 0$

$$(\mathrm{I}) \quad y(t) = \underbrace{v(t)}_{= z_1(t)} - \underbrace{\hat{z}(t)}_{= z_0(t)} = 0 \quad \text{and} \quad y \in \mathfrak{L}_+(P)$$

we obtain

$$\int_0^\infty (\Delta y)^* H (\Delta \hat{z})\, dt$$

$$= \int_0^\infty \left(\Delta_k^q\left(\tfrac{d}{dt}\right) y\right)^* \left(H \Delta_k^q\left(\tfrac{d}{dt}\right) \hat{z}\right) dt$$

$$\overset{\text{Lemma 4.2.}}{=} \int_0^\infty y^* \left(\Delta_k^{q\sim}\!\left(\tfrac{d}{dt}\right) H \, \Delta_k^{q}\!\left(\tfrac{d}{dt}\right) \hat{z} \right) dt$$

$$- \sum_{\ell=1}^{\infty} (-1)^\ell \Big(\underbrace{\Delta_k^{q\,\langle \ell \rangle}\!\left(\tfrac{d}{dt}\right) y(t)}_{=0,\ \text{with (II)}} \Big)^* \Big(H \Delta_k^{q}\!\left(\tfrac{d}{dt}\right) \hat{z}^{(\ell-1)}(t) \Big) \Big|_0^\infty$$

$$= \int_0^\infty y^* \left(\tilde{H}\!\left(\tfrac{d}{dt}\right) \hat{z} \right) dt = \int_0^\infty y^* \left(P^\sim\!\left(\tfrac{d}{dt}\right) \hat{\mu} \right) dt$$

$$\overset{\text{Lemma 4.2}}{=} - \int_0^\infty \Big(\underbrace{P\!\left(\tfrac{d}{dt}\right) y}_{=0,\ \text{since } y \in \mathfrak{B}_+(P)} \Big)^* \hat{\mu} \, dt - \sum_{\ell=1}^{\infty} (-1)^\ell \Big(\underbrace{P^{\langle \ell \rangle}\!\left(\tfrac{d}{dt}\right) y(t)}_{=0,\ \text{with (II)}} \Big)^* \hat{\mu}^{(\ell-1)}(t) \Big|_0^\infty$$

$$= 0 \qquad \Rightarrow \quad \frac{d}{ds} \phi_v(0) = 2 \operatorname{Re}\{0\} = 0.$$

Finally, the second derivative is

$$\left(\frac{d}{ds}\right)^2 \phi_v(s) = \int_0^\infty (\Delta y)^* H (\Delta y) \, dt \geq 0,$$

Since we assumed non-negativity of H and we have (II).

∎