8. Polynomials in free semicirculars and linearization

8.1. Remarks: 1) In 7.2. we saw that we can deal with linear matrices

$$S = b_1 \otimes S_1 + \dots + b_d \otimes S_d \quad (b_1, \dots, b_d \in M_n(\mathbb{C}))$$

in free semicirculars $S_1, \dots, S_d$.

Note that we can also consider "affine" matrices by adding a constant $b_0 \in M_n(\mathbb{C})$, since this gives only a shift in the argument of the Cauchy transform:

$$G_{b_0 \otimes 1 + S}(b) = id \otimes \varphi [\underbrace{b - (b_0 \otimes 1 + S)}_{(b - b_0) - S}]^{-1}$$

$$= G_S(b - b_0)$$

Since we consider selfadjoint random variables we need $b_0 = b_0^*$ and thus

$\operatorname{Im}(b - b_0) = \operatorname{Im} b \in H^+(M_n(\mathbb{C}))$.

So we can calculate $G_S(b - b_0)$ (at least numerically) and from this also the scalar-valued Cauchy transform

$$g_{b_0 \otimes 1 + S}(z) = tr\, G_{b_0 \otimes 1 + S}(z \cdot 1) = tr\, G_S(z \cdot 1 - b_0)$$

2) In 5.15 we saw that also for arbitrary selfadjoint polynomials $p \in \mathbb{C}\langle x_1, \dots, x_d \rangle$, the distribution of this polynomial applied to our free semicirculars, $p(S_1, \dots, S_d)$, is uniquely determined; however, it is not clear how to calculate this. We will now see that we can do this by relating this problem with a corresponding problem in affine matrices.

<u>8.2. Example:</u> Let us consider the example

$$p(x_1, x_2) = x_1 x_2 + x_2 x_1 + x_1^2$$

i.e.

$$P := p(S_1, S_2) = S_1 S_2 + S_2 S_1 + S_1^2$$

The distribution of P is given by its Cauchy transform (note $P = P^*$)

$$G_P(z) = \varphi[(z-P)^{-1}] \qquad \text{for } z \in \mathbb{H}^+ \subset \mathbb{C})$$

We left the problem of calculating the inverse now from the ground-level $\mathbb{C}$ to matrices by finding there a factorization of P into affine terms.

$$\underbrace{\begin{pmatrix} s_1 & \frac{s_1}{2}+s_2 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}}_{\begin{pmatrix} -\frac{s_1}{2}+s_2 & -s_1 \end{pmatrix}} \begin{pmatrix} s_1 \\ \frac{s_1}{2}+s_2 \end{pmatrix}$$

$$= -\left[\frac{s_1^2}{2} + s_2 s_1 + \frac{s_1^2}{2} + s_1 s_2 \right]$$

$$= -P$$

Let us denote

$$U = \begin{pmatrix} s_1 & \frac{s_1}{2}+s_2 \end{pmatrix} \qquad V = \begin{pmatrix} s_1 \\ \frac{s_1}{2}+s_2 \end{pmatrix}$$

$$Q^{-1} = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$$

then we have

$$P = -U Q^{-1} V$$

where U, Q, V are affine in s_1, s_2

This does not give directly a factorization for P^{-1}, since U, V are not invertible, but we get a factorization of a lifted version of $z - P$ into invertible factors

$$\begin{pmatrix} z-P & 0 \\ 0 & -Q \end{pmatrix} =$$

$$\underbrace{\begin{pmatrix} 1 & -UQ^{-1} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} z & -U \\ -V & -Q \end{pmatrix}}_{\begin{pmatrix} z-P & 0 \\ -V & -Q \end{pmatrix}} \begin{pmatrix} 1 & 0 \\ -Q^{-1}V & 1 \end{pmatrix}$$

We write $\hat{P} = \begin{pmatrix} 0 & U \\ V & Q \end{pmatrix}$

Since the first and third term are always invertible

$$\begin{pmatrix} 1 & A \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & -A \\ 0 & 1 \end{pmatrix}$$

we have

$$\begin{pmatrix} (z-P)^{-1} & 0 \\ 0 & -Q^{-1} \end{pmatrix} = \begin{pmatrix} z-P & 0 \\ 0 & -Q \end{pmatrix}^{-1} =$$

$$= \begin{pmatrix} 1 & 0 \\ Q^{-1}V & 1 \end{pmatrix} \begin{pmatrix} z & -U \\ -V & -Q \end{pmatrix}^{-1} \begin{pmatrix} 1 & UQ^{-1} \\ 0 & 1 \end{pmatrix}$$

If we put

$$\hat{P} = \begin{pmatrix} 0 & u \\ v & Q \end{pmatrix}, \quad \Lambda(z) = \begin{pmatrix} z & 0 \\ 0 & 0 \end{pmatrix}$$

then we have

$$\begin{pmatrix} (z-P)^{-1} & 0 \\ 0 & -Q^{-1} \end{pmatrix} = \begin{pmatrix} [(\Lambda(z)-\hat{P})^{-1}]_{1,1} & * \\ * & * \end{pmatrix}$$

i.e.

$$(z-P)^{-1} = [(\Lambda(z)-\hat{P})^{-1}]_{1,1}$$

$\uparrow$ (1,1) entry of the 3×3-matrix

and thus

$$\begin{aligned} G_P(z) &= \varphi[(z-P)^{-1}] \\ &= \varphi\{[(\Lambda(z)-\hat{P})^{-1}]_{1,1}\} \\ &= \{\underbrace{id\otimes\varphi\,[(\Lambda(z)-\hat{P})^{-1}]}_{G_{\hat{P}}(\Lambda(z))}\}_{1,1} \end{aligned}$$

Note that

$$\hat{P} = \begin{pmatrix} 0 & S_1 & \frac{S_1}{2}+S_2 \\ S_1 & 0 & -1 \\ \frac{S_1}{2}+S_2 & -1 & 0 \end{pmatrix}$$

is an affine matrix in free semicirculars, thus an $M_3(\mathbb{C})$ operator-valued semicircular, shifted by a constant, for which we can calculate its $M_3(\mathbb{C})$-valued Cauchy transform $G_{\hat{p}}(b)$.

Note also that $\Lambda(z) = \begin{pmatrix} z & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ is not in $H^+(M_3(\mathbb{C}))$, so our theory from Sect 6 for solving for $G_{\hat{p}}[\Lambda(z)]$ does not apply directly. But since $\Lambda(z) - \hat{P}$ is invertible, by the above calculation, the function $G_{\hat{p}}(b)$ is holomorphic, hence continuous in a neighborhood of $b = \Lambda(z)$ and we have

$$G_{\hat{p}}(\Lambda(z)) = \lim_{\substack{\varepsilon \to 0 \\ \varepsilon > 0}} G_{\hat{p}}(\Lambda_\varepsilon(z))$$

where

$$\Lambda_\varepsilon(z) = \underbrace{\begin{pmatrix} z & 0 & 0 \\ 0 & i\varepsilon & 0 \\ 0 & 0 & i\varepsilon \end{pmatrix}}_{\substack{\in H^+(M_3(\mathbb{C})) \\ \forall \varepsilon > 0}} \xrightarrow{\varepsilon \searrow 0} \Lambda(z)$$

<u>8.3. Remark:</u> The main ingredient in the above calculation was that we can factorize our polynomial as $P = -U Q^{-1} V$, with affine U, Q, V. This works for all polynomials and is actually independent from having semicircular elements as variables. Let us now do the general case on the level of formal variables $\mathbb{C}\langle x_1, \dots, x_d \rangle$.

<u>8.4. Definition:</u> Let $p \in \mathbb{C}\langle x_1, \dots x_d \rangle$ be given. A matrix

$$\hat{p} = \begin{pmatrix} 0 & u \\ v & q \end{pmatrix} \in M_n(\mathbb{C}\langle x_1, \dots, x_d \rangle)$$

where

- $n \in \mathbb{N}$
- $q \in M_{n-1}(\mathbb{C}\langle x_1 \dots x_d \rangle)$ is invertible as a matrix over polynomials
- $u \in M_{1,n}(\mathbb{C}\langle x_1, \dots, x_d \rangle)$ is a row and $v \in M_{n,1}(\mathbb{C}\langle x_1 \dots x_d \rangle)$ is a column

is called a <u>linearization</u> of p, if the following conditions are satisfied:

i) $\hat{p}$ is an affine matrix in $x_1, \dots, x_d$, i.e. there are $b_0, b_1, \dots, b_d \in M_n(\mathbb{C})$ s. th.

$$\hat{p} = b_0 \otimes 1 + b_1 \otimes x_1 + \dots + b_d \otimes x_d$$

ii) $p = -u q^{-1} v$

<u>8.5. Theorem:</u> For any polynomial $p \in \mathbb{C}\langle x_1, \dots, x_d \rangle$ there exists a linearization. If p is selfadjoint, then there is also a selfadjoint linearization.

<u>8.6 Remark:</u> Such linearizations are not unique; it is interesting to find minimal ones, where the matrix size n is as small as possible. Our algorithm in the following proof will not produce minimal realizations in general.

<u>Proof of 8.5:</u> 1) For monomials we have linearizations:

i) for degree 0, $p = \alpha \quad (\alpha \in \mathbb{C})$

$$\hat{p} = \begin{pmatrix} 0 & \alpha \\ 1 & -1 \end{pmatrix} \in M_2$$

$$-(\alpha (-1) 1) = \alpha$$

ii) for degree 1, $p = \alpha x_i$

$$\hat{p} = \begin{pmatrix} 0 & \alpha x_i \\ 1 & -1 \end{pmatrix} \in M_2$$

$$-\alpha x_i \cdot (-1) \cdot 1 = \alpha x_i$$

iii) for degree $k \geq 2$, $p = \alpha x_{i_1} \cdots x_{i_k}$

$$\begin{pmatrix} & & & \alpha x_{i_1} \\ & 0 & x_{i_2} & -1 \\ & x_{i_3} & -1 & \\ & \iddots & \iddots & 0 \\ x_{i_k} & -1 & & \end{pmatrix} \in M_k$$

e.g. $k = 3$

$$\begin{pmatrix} 0 & \boxed{0 \quad \alpha x_{i_1}} \\ \boxed{\begin{matrix} 0 \\ x_{i_3} \end{matrix}} & \boxed{\begin{matrix} x_{i_2} & -1 \\ -1 & 0 \end{matrix}} \end{pmatrix}$$

such matrices are invertible

$$\begin{pmatrix} q & -1 \\ -1 & 0 \end{pmatrix}^{-1} = \begin{pmatrix} 0 & -1 \\ -1 & -q \end{pmatrix}$$

$$-\begin{pmatrix} 0 & \alpha x_{i_1} \end{pmatrix} \begin{pmatrix} 0 & -1 \\ -1 & -x_{i_2} \end{pmatrix} \begin{pmatrix} 0 \\ x_{i_3} \end{pmatrix} = \alpha x_{i_1} x_{i_2} x_{i_3}$$

2) If we have linearizations $\begin{pmatrix} 0 & u_i \\ v_i & q_i \end{pmatrix}$ for polynomials p_i $(i=1,\dots,r)$, then $p_1+\dots+p_r$ has a linearization

$$\begin{pmatrix} 0 & u_1 & u_2 & \cdots & u_r \\ v_1 & q_1 & 0 & & 0 \\ v_2 & 0 & q_2 & & 0 \\ \vdots & & & \ddots & \\ v_r & 0 & 0 & & q_r \end{pmatrix}$$

3) If p has linearization $\begin{pmatrix} 0 & u \\ v & q \end{pmatrix}$, then p^* has linearization $\begin{pmatrix} 0 & v^* \\ u^* & q^* \end{pmatrix}$

If $p = p^*$ we want to take a linearization of $p = \frac{1}{2}(p+p^*)$. The construction in (2) however does not give a s.a. $\hat{p}$. Instead we take

$$\frac{1}{2}\begin{pmatrix} 0 & u & v^* \\ u^* & 0 & q^* \\ v & q & 0 \end{pmatrix}$$

$\square$

8.7. Remark: The linearization and the calculations for expressing the Cauchy transform of P in terms of the Cauchy transform of $\hat{P}$ are independent of the concrete nature of our random variables, neither freeness nor being semicircular is important. Thus we have the following.

8.8. Theorem: Let $(\mathcal{A}, \varphi)$ be a C^*-probability space and consider selfadjoint $X_1, \dots, X_d \in \mathcal{A}$. For a selfadjoint polynomial $p \in \mathbb{C}\langle x_1, \dots x_d \rangle$ let $\hat{p} = b_0 \otimes 1 + b_1 \otimes x_1 + \dots + b_d \otimes x_d$, with $b_0, \dots, b_d \in M_n(\mathbb{C})$, be a selfadjoint linearization of p. Put

$$P := p(X_1, \dots, X_d) \in \mathcal{A}$$

and

$$\hat{P} := \hat{p}(X_1, \dots, X_d) = b_0 \otimes 1 + b_1 \otimes X_1 + \dots + b_d \otimes X_d$$

$$\in M_n(\mathcal{A})$$

Then we have for $z \in H^+(\mathbb{C})$

$$G_p(z) = \left[G_{\hat{p}}(\Lambda(z))\right]_{1,1}$$

$$= \lim_{\varepsilon \searrow 0} \left[G_{\hat{p}}(\Lambda_\varepsilon(z))\right]_{1,1}$$

with

$$\Lambda(z) = \begin{pmatrix} z & & 0 \\ & 0 & \\ 0 & & 0 \end{pmatrix} \quad \text{and}$$

$$\Lambda_\varepsilon(z) = \begin{pmatrix} z & & & 0 \\ & i\varepsilon & & \\ & & \ddots & \\ 0 & & & i\varepsilon \end{pmatrix} \in H^+(M_n(\mathbb{C}))$$

8.9. Remark: Thus in order to deal with polynomials of variables we need to understand linear matrices in the variables. If the variables are free semicirculars we understand linear matrices in them, as they are operator-valued semicirculars. But how about general free variables: If $X_1, \dots, X_d$ are free, what can we say about

$$X = b_0 \otimes 1 + b_1 \otimes X_1 + \dots + b_d \otimes X_d$$

Note:

i) $b_0 \otimes 1$ is just a shift of the argument in G_X, thus easy to deal with

ii) the operator-valued Cauchy transform of $b_i \otimes X_i$ is determined (theoretically and numerically) in terms of the Cauchy transform of X_i

iii) if $X_1, \dots, X_d$ are free in $(\mathcal{A}, \varphi)$, then $b_1 \otimes X_1, \dots, b_d \otimes X_d$ are free in $(M_n(\mathcal{A}), M_n(\mathbb{C}), id \otimes \varphi)$ (by 5.7); hence we have to understand how to deal with sums of free variables on an operator-valued level

$\rightarrow$ operator-valued free convolution