COMPLETELY REGULAR MULTIVARIATE STATIONARY PROCESSES AND THE MUCKENHOUPT CONDITION.

S. TREIL AND A. VOLBERG

1. Introduction

In this paper we shall give a necessary and sufficient condition for a multivariate stationary stochastic process to be *completely regular*.

Let us recall that a multivariate stationary stochastic process with discrete time [1] is a sequence of d-tuples $x(n) = (x_1(n), x_2(n), \dots, x_d(n))$, $n \in \mathbb{Z}$ of scalar random variables such that $\mathbb{E}|x_j(n)|^2 < \infty$ and the *correlation matrix* $Q(n,k)$

$$Q(n,k) = \{Q(n,k)_{i,j}\}_{1\le i,j\le d} := \left\{\mathbb{E}x_i(n)\overline{x_j(k)}\right\}_{1\le i,j\le d}$$

depends only on the difference $n-k$; here $\mathbb{E}$ denotes mathematical expectation.

It is well known (see [1]) that there exists a matrix-valued non-negative measure M on the unit circle $\mathbb{T}$ whose Fourier coefficients coincide with entries of the correlation matrix

$$Q(n,k) = \widehat{M}(n-k)\,, \qquad n,k \in \mathbb{Z}\,.$$

The measure M is called the *spectral measure* of the process $\{x(n)\}_{n\in\mathbb{Z}}$.

The random variables $x_j(n)$ can be treated as elements of Hilbert space $L^2(\Omega, dP)$, where Ω is the probability space and P is the probability, so $x(n)$ can be treated as elements of the $\mathbb{R}^d$-valued L^2 space $L^2_{\mathbb{R}^d}(\Omega, dP)$ For a moment n of time we can consider the past $\mathcal{X}_n$ and the future $\mathcal{X}^n$ of the process, which are defined as the subspaces

$$\mathcal{X}_n = \operatorname{span}\{x_j(k) \,:\, 1\le j\le d,\ k<n\}$$
$$\mathcal{X}^n = \operatorname{span}\{x_j(k) \,:\, 1\le j\le d,\ k\ge n\}$$

of $L^2(\Omega, dP)$.

A process is called *regular* if $\cap_{n\ge 0}\mathcal{X}^n = \{0\}$. In this case (see [1]) the spectral measure M of the process is absolutely continuous with respect to Lebesgue measure. Let W be the density of M with respect to Lebesgue measure. The matrix-valued function W is called the spectral density of the process.

A process $\{x(n)\}_{n\in\mathbb{Z}}$ is called *completely regular* if its past is asymptotically orthogonal to the future, namely if

$$\sup\left\{|\mathbb{E}(\xi\eta)| \,:\, \xi\in\mathcal{X}_0,\ \eta\in\mathcal{X}^n,\ \mathbb{E}|\xi|^2\le 1,\ \mathbb{E}|\eta|^2\le 1\right\} \longrightarrow 0 \qquad \text{as } n\to\infty\,.$$

Of course, complete regularity implies regularity. If the process is Gaussian (i.e. all random variables $x_j(k)$ have normal distribution) then the complete regularity means simply that past and future are almost independent. The problem we are dealing with is to characterize completely regular processes in terms of spectral measure.

[1] To avoid non-essential to the problem but unpleasant technicalities, we shall consider in this paper only stationary processes with discrete time.

It is well known (see again [1]) that if the process is completely regular, then its spectral measure is absolutely continuous, $dM = W dm$ where dm is the normalized ($m(\mathbb{T}) = 1$) Lebesgue measure on the unit circle $\mathbb{T}$.

The reader is referred to [1] once more to see that there exists $d_0 \le d$ (the rank of the process) such that the spectral density $W(t)$ has rank d_0 for almost all $t \in \mathbb{T}$. If $d_0 = d$ then the process $\{x(n)\}$ is said to be a *full rank.*

The study of processes of arbitrary rank can be easily reduced to the study of the processes of full rank, see [2]. So in this paper we shall consider only processes of full rank.

For the scalar case the description of completely regular processes was obtained by Helson an Sarason, see [3, 7]. To state their result we need a couple of definitions.

Let us recall that a function f on the unit circle $\mathbb{T}$ belongs to the space BMO (bounded mean oscillation) if

$$\sup_I \frac{1}{|I|}\int_I |f - f_I| dm = \|f\|_{\mathrm{BMO}} < \infty\,;$$

here f_I denotes the mean value of f on the interval I: $f_I := |I|^{-1}\int_I f dm$ and the *supremum* is taken over all subarcs I of $\mathbb{T}$.

The space VMO (vanishing mean oscillation) consists of all function $f \in$ BMO such that

$$\sup_I \frac{1}{|I|}\int_I |f - f_I| dm \longrightarrow 0 \qquad \text{as } |I| \to 0\,.$$

Theorem 1.1 (Helson, Sarason)**.** *Let w be the spectral density of a scalar stationary process. Then the process is completely regular if and only if w admits a representation*

$$w = |p|^2 e^{\varphi}\,,$$

where p is a polynomial with roots on the unit circle $\mathbb{T}$ and φ is a real-valued function in VMO.

It was conjectured by V. Peller in [4] that the same result holds for multivariate stationary processes. Namely he conjectured that a multivariate stationary process is completely regular if and only if its spectral density W admits the following representation

$$W = P^* e^{\Phi} P,$$

where P is a polynomial matrix whose determinant has roots on $\mathbb{T}$ and the matrix function $\Phi = \Phi^*$ belongs VMO.

In this direction he was able to prove the following theorem

Theorem 1.2. *A multivariate stationary process is completely regular if and only if its spectral density W admits the factorization*

$$W = P^* W_1 P,$$

where P is a polynomial matrix whose determinant has roots on $\mathbb{T}$ and W_1 is the density of a completely regular stationary process such that $W_1^{-1} \in L^1$.

1.1. **The main result.** Let us recall that a measure μ on the unit disk $\mathbb{D}$ is called Carleson if

$$\sup_I \mu(Q(I)) \le C \cdot |I|$$

and is called the vanishing Carleson measure if

$$\limsup_{|I|\to 0} \mu(Q(I))/|I| = 0$$

where *limsup* is taken over all subarcs I of T. Here $Q(I)$ denotes the "Carleson square" for the arc I,

$$Q(I) = \{z \in \mathbb{D} : z/|z| \in I, 1 - |I| \leq |z| < 1\}$$

For a function F on the unit circle let $F(\lambda)$, $\lambda \in \mathbb{D}$, denote its harmonic extension at the point λ.

The main result of the paper is the following theorem.

Theorem 1.3. *Let the density W of a stationary process satisfy $W^{-1} \in L^1$. Then the the following are equivalent*

1. *The process is completely regular;*
2. *W^{-1} is the spectral density of a completely regular process;*
3. $\limsup\limits_{|I|\to 0} \left\| \Big(\frac{1}{|I|}\int_I W dm\Big)^{1/2} \Big(\frac{1}{|I|}\int_I W^{-1} dm\Big)^{1/2} \right\| = 1$*; here* supremum *is taken over all subarcs I of $\mathbb{T}$;*
4. $\limsup\limits_{|\lambda|\to 1} \left\| \Big(W(\lambda)\Big)^{1/2} \Big(W^{-1}(\lambda)\Big)^{1/2} \right\| = 1$*, where $W(\lambda)$ and $W^{-1}(\lambda)$ are harmonic extensions of functions $W \Big| \mathbb{T}$ and $W^{-1} \Big| \mathbb{T}$ respectively at point $\lambda \in \mathbb{D}$.*
5. $\limsup\limits_{|\lambda|\to 1} \Big\{\det\Big(W(\lambda)\Big) \exp\Big(-\big[\log\det W\big](\lambda)\Big)\Big\} = 1$*, where $W(\lambda)$ and $\big[\log\det W\big](\lambda)$ are harmonic extensions of functions $W \Big| \mathbb{T}$ and $\log\det W \Big| \mathbb{T}$ respectively at point $\lambda \in \mathbb{D}$.*
6. *The measures*

$$\left\| W(z)^{-1/2} \left(\frac{\partial}{\partial x} W(z)\right) W(z)^{-1/2} \right\|^2 (1 - |z|^2) dx dy$$

and

$$\left\| W(z)^{-1/2} \left(\frac{\partial}{\partial y} W(z)\right) W(z)^{-1/2} \right\|^2 (1 - |z|^2) dx dy$$

are vanishing Carleson measures

Together with Theorem 1.2 the above theorem yields the complete description of completely regular stationary processes

Theorem 1.4. *A stationary process with spectral density $\mathcal{W}$ is completely regular if and only if $\mathcal{W}$ admits the representation*

$$\mathcal{W} = P^* W P,$$

where P is a polynomial matrix whose determinant has roots on $\mathbb{T}$ and the matrix-function W satisfies $W^{-1} \in L^1$ and one of equivalent conditions 3–6 of Theorem 1.3

Let us discuss the main result (Theorem 1.3) a little bit. First of all it is not difficult to show directly that in the scalar case the conditions 3–6 of Theorem 1.3 are equivalent to $W = e^{\varphi}$, $\varphi \in \mathrm{VMO}$. We are leaving this as an exercise for the reader.

Usually in probability only real valued stationary processes are considered. In that case the spectral density of a process should satisfy $W(\overline{z}) = W(z)$, and only such functions can be realized as densities of stationary processes.

If one allow complex-valued processes, any non-negative matrix function is the spectral density of some stationary process.

Our theorem deals with arbitrary non-negative matrix-functions and can be applied to complex-valued processes (as well as to real-valued).

2. Scheme of the proof of the main result

The diagram of the proof will be the following: $1 \Longrightarrow 4 \Longrightarrow 5 \Longrightarrow 6 \Longrightarrow 1$. Then we will show that $1 \Longrightarrow 2$ and so automatically $2 \Longrightarrow 1$.

And in this section we will show that $3 \Longleftrightarrow 4$.

Lemma 2.1. *For a scalar weight w the following conditions are equivalent:*

1. $\limsup_{|I|\to 0}\Big(\frac{1}{|I|}\int_I w\Big)\Big(\frac{1}{|I|}\int_I w^{-1}\Big) = 1$;
2. $\limsup_{|\lambda|\to 1} w(\lambda)w^{-1}(\lambda) = 1$, *where $w(\lambda)$ and $w^{-1}(\lambda)$ denote the harmonic extensions of w and w^{-1} respectively at the point λ;*
3. $w = e^{\varphi}$, *where $\varphi \in$ VMO.*

Proof. First of all let us rewrite condition 1. Let $\varphi := \log w$. For a function f let f_I denote its average over the arc I, $f_I := |I|^{-1}\int_I f$. Then clearly

$$w_I \cdot (w^{-1})_I = \big[w_I \exp(-\varphi_I)\big] \cdot \big[(w^{-1})_I \exp(\varphi_I)\big].$$

By Jensen inequality (geometric mean $\le$ arithmetic mean) the expressions in brackets are at least 1, so the condition ! splits into the following 2 conditions

$$\limsup_{|I|\to 0}\big[w_I \exp(-\varphi_I)\big] = 1, \qquad \text{and} \qquad \limsup_{|I|\to 0}\big[w_I^{-1} \exp(\varphi_I)\big] = 1.$$

Let f_+ denote the positive part of the function f, $f_+(x) := \max(f(x), 0)$. Then the inequality

$$x \le e^x - 1 \qquad \text{for } x \ge 0$$

implies

$$\frac{1}{|I|}\int_I (\varphi - \varphi_I)_+ \le \frac{1}{|I|}\int_I \Big(\exp(\varphi - \varphi_I) - 1\Big) = w_I \exp(-\varphi_I) - 1 \to 0 \quad \text{as } |I| \to 0.$$

Since $\int_I |\varphi - \varphi_I| = 2\int_I (f - f_I)_+$, one can conclude that $\varphi \in$ VMO.

Similarly, using Poisson averages instead of averages over intervals one can get from condition 2 of the lemma that harmonic extension of $|\varphi - \varphi(\lambda)|$ at the point λ tends to 0 as $\lambda \to 1$. But that is an equivalent definition of VMO, so the condition 2 also implies that $\varphi \in$ VMO.

On the other hand, if $\varphi \in$ VMO, John–Nirenberg Theorem (see [5][Chapter VI], the measure of the set $\{t \in I : |\varphi(t) - \varphi_I| > a\}$ is estimated from above by Ce^{-Ka}, where $K = K_I \to \infty$ as $|I| \to 0$. Therefore for $x > 1$ the measure of the set $\{t \in I : \exp(\varphi(t) - \varphi_I) > x\}$ is estimated from above by Cx^{-K}. Integrating this distribution function one can get that $\limsup_{|I|\to 0} w_I \exp(-\varphi_I) \le 1$ (in fact, it is 1, because by Jensen inequality $w_I \exp(-\varphi_I) \ge 1$.

Similarly, $\limsup_{|I|\to 0}(w^{-1})_I \exp(\varphi_I) = 1$. Multiplying the above two inequalities one gets condition 1.

The proof that $3 \Longrightarrow 2$ is similar. For a point $\lambda \in \mathbb{D}$ let I_λ be an interval with center at $\lambda/|\lambda|$ of length $\sqrt{1-|\lambda|}$. Since the Poisson Kernel $P_\lambda(z) = (1-|\lambda|^2)\cdot|1-\overline{\lambda}z|^{-2}$ satisfies $\sup_{z\in\mathbb{T}\setminus I_\lambda} P_\lambda(z) \to 0$ as $|\lambda| \to 1$, the distribution inequality for φ on I_λ implies that $w(\lambda)\cdot\exp(-\varphi(\lambda)) \to 0$ as $|\lambda| \to 1$, and therefore the condition 2 of the lemma. □

The following Lemma is probably well known and can be easily from the distribution function inequality for VMO (John–Nirenberg Theorem).

Lemma 2.2. *For $\lambda \in \mathbb{D}$ let I_λ be an interval centered at $\lambda/|\lambda|$ of length $1-|\lambda|$. If $\varphi \in VMO$, then $\varphi_{I_\lambda} - \varphi(\lambda) \to 0$ as $|\lambda| \to 1$.*

Corollary 2.3. *Let $\varphi \in VMO$ and let $w = e^\varphi$. Then for I_λ as in the above lemma we have*

$$\lim_{|\lambda|\to 1} \frac{w(\lambda)}{w_{I_\lambda}} = 1.$$

Proof. By the above lemma $\lim_{|\lambda|\to 1} \exp(w(\lambda))/\exp(w_{I_\lambda}) = 1$. On the other hand it follows from the proof of Lemma 2.1 that

$$\lim_{|\lambda|\to 1} w(\lambda)/\exp(\varphi(\lambda)) = 1 \qquad \text{and} \qquad \lim_{|I|\to 0} w_I/\exp(\varphi_I) = 1.$$

Taking the ration of the last 2 identities (with $I = I_\lambda$)we get the statement we need. □

Now to show equivalence of condition 3 and 4 of the Theorem 1.3 is is enough to show that these conditions imply that for a fixed vector $e \in \mathbb{C}^d$ scalar weight $w(z) = (W(z)e, e)$ satisfies conditions 1 and 2 of Lemma 2.1. Then Corollary 2.3 implies that the averages W_{I_λ} and $W(\lambda)$ are equivalent, the same holds for W^{-1}, and we are done.

It remains now to show that the scalar weight $w(z) = (W(z)e, e)$ satisfies condition 1 (equivalently 2) of Lemma 2.1. The easiest way to do that is to recall where the Muckenhoupt condition (A_2) came from, see [8].

Recall that the quantity $\|[W_I]^{1/2}[(W^{-1})_I]^{1/2}\|$ is just the norm of the averaging operator $f \mapsto f_I \cdot \chi_I$ in the weighted space $L^2(W)$, see [8][Lemma 2.1]. Then $[w_I]^{1/2}[(w^{-1})_I]^{1/2}$ is is the norm of the restriction of the above averaging operator onto the subspace of $L^2(W)$ consisting of functions of form fe where f is a scalar function. Therefore

$$1 \le [w_I]^{1/2}[(w^{-1})_I]^{1/2} \le \|[W_I]^{1/2}[(W^{-1})_I]^{1/2}\|$$

so the weight w satisfies condition 1 of the lemma.

Similarly, the quantity $\|W(\lambda)^{1/2}W^{-1}(\lambda)^{1/2}\|$ is just the norm of another averaging operator $\left(f \mapsto \int_{\mathbb{T}} f k_\lambda\right) k_\lambda$, where k_λ is the normalized reproducing kernel of H^2, $k_\lambda(z) = (1-|\lambda|^2)^{1/2}(1-\overline{\lambda}z)^{-1}$, see [8][Lemma 2.1], so condition 4 of the theorem imply condition 2 of the lemma for the weight w.

3. Eliminating probability

The problem of description of completely regular processes can be now stated without mentioning any probability theory at all.

First of all notice that without loss of generality we can assume that the process is complex-valued. Namely, if we have a real stationary process $\{x(n)\}_{n\in\mathbb{Z}}$ we can consider its comlexification, namely the same process but in the complex Hilbert space $L^2_{\mathbb{C}^d}(\Omega, dP)$. Considers the comlexificated past $(\mathcal{X}_n)_{\mathbb{C}}$ and future $(\mathcal{X}^n)_{\mathbb{C}}$

$$(\mathcal{X}_n)_{\mathbb{C}} = \operatorname{span}\{x_j(k) \,:\, 1 \le j \le d,\, k < n\}$$
$$(\mathcal{X}^n)_{\mathbb{C}} = \operatorname{span}\{x_j(k) \,:\, 1 \le j \le d,\, k \ge n\}$$

where span now means the closed linear span in the complex Hilbert space $L^2_{\mathbb{C}^d}(\Omega, dP)$. It is easy to see that

$$\sup\left\{|\mathbb{E}(\xi\eta)| \,:\, \xi \in \mathcal{X}_0,\, \eta \in \mathcal{X}^n,\, \mathbb{E}|\xi|^2 \le 1,\, \mathbb{E}|\eta|^2 \le 1\right\} = \\ = \sup\left\{|\mathbb{E}(\xi\bar\eta)| \,:\, \xi \in (\mathcal{X}_0)_{\mathbb{C}},\, \eta \in (\mathcal{X}^n)_{\mathbb{C}},\, \mathbb{E}|\xi|^2 \le 1,\, \mathbb{E}|\eta|^2 \le 1\right\},$$

so a process and its comlexification are completely regular simultaneously. So we indeed can assume from the beginning that our process is complex valued.

Consider now the vector space $L^2(W)$ of $\mathbb{C}^d$-valued functions on the unit circle with the norm

$$\|f\|^2_{L^2(W)} = \int_{\mathbb{T}} (W(\xi)f(\xi), f(\xi))_{\mathbb{C}^d}\, dm(\xi)$$

(of course we have to take the quotient space over the functions of norm 0). The mapping $x_j(k) \mapsto z^k e_j$, where e_j, $j = 1, ..., d$ is the standard orthonormal basis in $\mathbb{C}^d$, is an isometric isomorphism between $\operatorname{span}\{x_j(k) \,:\, 1 \le j \le d, k \in \mathbb{Z}\}$ and $L^2(W)$.

The past $\mathcal{X}_n$ and future $\mathcal{X}^n$ are mapped to the spaces X_n and X^n of $L^2(W)$

$$X_n = \operatorname{span}\{z^k\mathbb{C}^d \,:\, k < n\} \tag{3.1}$$
$$X^n = \operatorname{span}\{z^k\mathbb{C}^d \,:\, k \ge n\}. \tag{3.2}$$

So the problem of describing completely regular stationary processes can be reformulated as follows: *describe all matrix weights W such that the spaces X_0 and X^n are asymptotically (as $n \to \infty$) orthogonal to each other,*

$$\rho_n = \sup\left\{|(\xi,\eta)_{L^2(W)}| \,:\, \xi \in X_0,\, \eta \in X^n,\, \|\xi\|_{L^2(W)} \le 1,\, \|\eta\|_{L^2(W)} \le 1\right\} \longrightarrow 0, \tag{3.3}$$

as $n \to \infty$.

4. Necessity ($1 \Longrightarrow 4$)

In this section we are going to prove the implication $1 \Longrightarrow 4$ (see Theorem 4.1 below) and the equivalence $1 \Longleftrightarrow 2$ (see Lemma 4.4).

For a function F defined on the unit circle $\mathbb{T}$ let $F(\lambda)$ denote its harmonic extension a the point $\lambda \in \mathbb{D}$.

Theorem 4.1. *Let W be a matrix valued weight such that $W^{-1} \in L^1$. Suppose the "past" X_0 and "future" X^n defined by* (3.1), (3.2) *are asymptotically orthogonal, that means*

$$\rho_n = \sup\left\{|(\xi,\eta)_{L^2(W)}| \,:\, \xi \in X_0,\, \eta \in X^n,\, \|\xi\|_{L^2(W)} \le 1,\, \|\eta\|_{L^2(W)} \le 1\right\} \longrightarrow 0$$

as $n \to \infty$. Then

$$\limsup_{|\lambda|\to 1} \left\| \Big(W(\lambda)\Big)^{1/2} \Big(W^{-1}(\lambda)\Big)^{1/2} \right\| = 1.$$

Proof. First of all let us show that if W^{-1} is completely regular and $W^{-1} \in L^1$ then W satisfies the Muckenhoupt (A_2) condition

$$\sup_{\lambda\in\mathbb{D}} \left\| \Big(W(\lambda)\Big)^{1/2} \Big(W^{-1}(\lambda)\Big)^{1/2} \right\| < \infty\,. \tag{A_p}$$

Recall that $\left\| \Big(W(\lambda)\Big)^{1/2} \Big(W^{-1}(\lambda)\Big)^{1/2} \right\|$ is exactly the norm of the operator $f \mapsto (f, k_\lambda)k_\lambda$ in the weighted space $L^2(W)$; here k_λ denotes the normalized reproducing kernel for H^2,

$$k_\lambda(z) := \frac{(1-|\lambda|^2)^{1/2}}{1-\overline{\lambda}z}\,, \qquad \lambda \in \mathbb{D},$$

$\|k_\lambda\|_2 = 1$. Note that $k_0 \equiv 1$. So if $W^{-1} \in L^1$ the operator $f \mapsto (f,1)1$ is bounded in $L^2(W)$, and therefore by translation invariance the operators $f \mapsto (f, z^n)z^n = \hat{f}(n)z^n$ are bounded as well (they all have the same norm).

We know that the spaces X_0 and X^n are asymptotically orthogonal, so we can say that for large enough N the operator P_+ restricted onto $\operatorname{span}\{X_0, X^N\} = \operatorname{span}\{z^n\mathbb{C}^d : n \notin [0,N]\}$ is bounded, say by 2,

$$\|P_+ f\|_{L^2(W)} \le 2\|f\|_{L^2(W)}\,, \qquad \forall f \in \operatorname{span}\{X_0, X^N\} = \operatorname{span}\{z^n\mathbb{C}^d : n \notin [0,N]\}\,.$$

Since $f - \sum_{n=0}^N \hat{f}(n)z^n \in \operatorname{span}\{X_0, X^N\} = \operatorname{span}\{z^n\mathbb{C}^d : n \notin [0,N]\}$, one can conclude that the operator P_+ is bounded in $L^2(W)$ and so the weight satisfy the Muckenhoupt condition (A_2).

We will need the following simple lemma about Muckenhoupt weights.

Lemma 4.2. *If w is a scalar Muckenhoupt weight, then its harmonic extension $w(\lambda)$ can't decay to fast near the boundary of the disk. Namely, if the Muckenhoupt norm of w is at most M there is a function $\alpha = \alpha_M$, $\alpha : [0,1) \to (0,\infty)$, $\alpha(t) \searrow 0$ as $t \to 1+$ such that*

$$\frac{(1-|\lambda|^2)w(0)}{w(\lambda)} \le \alpha(|\lambda|).$$

Proof of the lemma. For an arc $I \subset \mathbb{T}$ and $k > 0$ let kI denote the arc of length $k|I|$ with the same center as I.

We are going to show that for a Muckenhoupt weight w with the Muckenhoupt norm at most M

$$w_{2^nI} \le M^2(2-\varepsilon)^n w_I\,, \qquad \varepsilon = \varepsilon(M) > 0\,. \tag{4.1}$$

Applying this formula in the case $2^nI = \mathbb{T}$ and using the trivial estimate

$$w(\lambda) \ge C w_{I_\lambda}$$

where I_λ is the arc with center at the point $\lambda/|\lambda|$, $|I_\lambda| = 1 - |\lambda|^2$ and C is an absolute constant, we can get from there (recall that $|I_\lambda| = 1 - |\lambda|^2 = 2^{-n}$)

$$w(\lambda) \ge c(2-\varepsilon)^{-n} \cdot w(0) = c(2-\varepsilon)^{\log_2(1-|\lambda|^2)} \cdot w(0) = c \cdot (e-\delta)^{\log(1-|\lambda|^2)} \cdot w(0),$$

where $\delta = \delta(\varepsilon) > 0$; here e is the base of the natural logarithm, not a vector in $\mathbb{C}^d$. This estimate implies the conclusion of the lemma with $\alpha(t) = c^{-1}(1-t^2)\cdot(e-\delta)^{-\log(1-t^2)}$.

To prove (4.1) we notice the since the weight w^{-1} is the Muckenhoupt (A_2) weight with the same Muckenhoupt norm as w, it is doubling and therefore

$$(w^{-1})_{2I} \geq (2-\varepsilon)^{-1}(w^{-1})_I\,,$$

where ε depends only on the Muckenhoupt norm of w. Iterating this inequality n times we get

$$(w^{-1})_{2^nI} \geq (2-\varepsilon)^{-n}(w^{-1})_I\,.$$

The last estimate and the Muckenhoupt condition imply

$$w_{2^nI} \leq M/(w^{-1})_{2^nI} \leq M\cdot(2-\varepsilon)^n/(w^{-1})_I \leq M^2 w_I$$

and that is exactly what we need. □

Corollary 4.3. *If a matrix weight W satisfies the Muckenhoupt condition (A_2) with the Muckenhoupt norm at most M then for any $e\in\mathbb{C}^d$*

$$(1-|\lambda|^2)\cdot\frac{(W(0)e,e)_{\mathbb{C}^d}}{(W(\lambda)e,e)_{\mathbb{C}^d}} \leq \alpha(|\lambda|)\ \to 0 \qquad \text{as } |\lambda|\to 1,$$

where $\alpha=\alpha_M$ is the function from Lemma 4.2.

Proof of the corollary. The proof follows immediately from the fact that the scalar weight w, $w(\xi) = \big(W(\xi)e,e\big)_{\mathbb{C}^d}$ is the Muckenhoupt (A_2) weight with the Muckenhoupt norm at most M (see [9], proof of Corollary 2.4). □

We now return to the proof of the theorem.

The condition $W^{-1}\in L^1$ implies that $\int_{\mathbb{T}}\log\det W(\xi)dm(\xi) > -\infty$, hence (see [6]) there exists a factorization of W of the form $W = F^*F$, where F is an outer matrix function in H^2.

Take $e\in\mathbb{C}^d$ and let us compute the distance

$$\operatorname{dist}_{L^2(W)}\{z^{-1}e, \operatorname{span}\{z^n\mathbb{C}^d\ :\ n\geq 0\} = \operatorname{dist}_{L^2(W)}\{e, \operatorname{span}\{z^n\mathbb{C}^d\ :\ n>0\}\,.$$

By the vectorial version of the Szegö theorem (see [6]) this distance is exactly $\|F(0)e\|$. Using the Möbius transformation of the disk one can get from there

$$\operatorname{dist}_{L^2(W)}\{\frac{(1-|\lambda|^2)^{1/2}}{z-\lambda}e, \operatorname{span}\{z^n\mathbb{C}^d\ :\ n\geq 0\}\} = \|F(\lambda)e\|_{\mathbb{C}^d}\,.$$

Writing the Fourier series expansion of $\frac{(1-|\lambda|^2)^{1/2}}{z-\lambda}$

$$\frac{(1-|\lambda|^2)^{1/2}}{z-\lambda} = (1-|\lambda|^2)^{1/2}\sum_{n=0}^{\infty}\lambda^n z^{-(n+1)}$$

one can see that for any fixed $N>0$ the function $\frac{(1-|\lambda|^2)^{1/2}}{z-\lambda}e$ is almost in the "past" X_{-N} as $|\lambda|\to 1$. Namely,

$$f_\lambda = \frac{(1-|\lambda|^2)^{1/2}}{z-\lambda}e = (1-|\lambda|^2)^{1/2}\sum_{n=0}^{N-1}\lambda^n z^{-(n+1)}e + (1-|\lambda|^2)^{1/2}\sum_{n=N}^{\infty}\lambda^n z^{-(n+1)}e = f^1_\lambda + f^2_\lambda\,,$$

where $f_\lambda^2 \in X_{-N}$, and f_λ^1 is small,

$$\frac{\|f_\lambda^1\|_{L^2(W)}}{\|f_\lambda\|_{L^2(W)}} \le \frac{(1-|\lambda|^2)^{1/2}N\cdot\|e\|_{L^2(W)}}{\big(W(\lambda)e,e\big)_{\mathbb{C}^d}^{1/2}} = \frac{(1-|\lambda|^2)^{1/2}N\cdot\big(W(0)e,e\big)_{\mathbb{C}^d}^{1/2}}{\big(W(\lambda)e,e\big)_{\mathbb{C}^d}^{1/2}} \le N\alpha(|\lambda|)^{1/2}\to 0,$$

as $|\lambda|\to 1$, where $\alpha(.)$ is as in Lemma 4.2 and Corollary 4.3.

Since X_0 and X^N are asymptotically orthogonal, the shift invariance implies that the subspaces X_{-N} and X^0 are asymptotically orthogonal as well. Taking $|\lambda|\to 1$ and then $N\to\infty$ we can conclude that

$$\|F(\lambda)e\|_{\mathbb{C}^d}/\|W(\lambda)^{1/2}e\|_{\mathbb{C}^d} =$$
$$= \operatorname{dist}_{L^2(W)}\{\frac{(1-|\lambda|^2)^{1/2}}{z-\lambda}e, \operatorname{span}\{z^n\mathbb{C}^d\ :\ n\ge 0\}\}/\,\|f_\lambda\|_{L^2(W)} \ge 1-\beta(|\lambda|)^{1/2}\to 1\,,$$

where $\beta(.)$ depends only on the Muckenhoupt norm of W and $\beta(|\lambda|)\to 0$ as $|\lambda|\to 1$.

The last inequality implies

$$\|W(\lambda)^{1/2}F(\lambda)^{-1}\| \le (1-\beta(|\lambda|))^{-1}\,. \tag{4.2}$$

Note that since $\|F(\lambda)e\|_{\mathbb{C}^d}/\|W(\lambda)^{1/2}e\|_{\mathbb{C}^d}\le 1$ for all $e\in\mathbb{C}^d$, we have

$$\|W(\lambda)^{1/2}F(\lambda)^{-1}\|\ge 1\,.$$

We will show a little later that under assumptions of the theorem the subspaces X_0 and X^N in the weighted space $L^2(W^{-1})$ are asymptotically orthogonal as well. The factorization $W=F^*F$ yields the factorization $W^{-1}=F^{-1}(F^{-1})^*$ of W^{-1}. Similarly to the previous case

$$\operatorname{dist}_{L^2(W^{-1})}\{\frac{(1-|\lambda|^2)^{1/2}}{1-\overline{\lambda}z}e, \operatorname{span}\{z^n\mathbb{C}^d\ :\ n\ge 0\}\} = \|F^{-1}(\lambda)^*e\|_{\mathbb{C}^d} = \|F(\lambda)^{-1*}e\|_{\mathbb{C}^d}\,.$$

Acting as before we get

$$\|W^{-1}(\lambda)^{1/2}F(\lambda)^*\|\le (1-\beta_1(|\lambda|))^{-1} \tag{4.3}$$

where $\beta_1(|\lambda|)\to 0$ as $|\lambda|\to 1$.

Combining (4.2) and (4.3) we get

$$\|W(\lambda)^{1/2}W^{-1}(\lambda)^{1/2}\|\le (1-\beta(|\lambda|))^{-1}(1-\beta_1(|\lambda|))^{-1}\to 1 \qquad \text{as } |\lambda|\to 1\,.$$

So, we completed the proof modulo the following lemma. $\square$

This lemma also gives us the equivalence $1 \iff 2$.

Lemma 4.4. *Under assumptions of Theorem 4.1 the weight W^{-1} is a spectral density of a completely regular process, i.e that the spaces X_0 and X^N are asymptotically orthogonal (as $N\to\infty$) in the weighted space $L^2(W^{-1})$.*

Proof. It is enough to show that

$$\|P_+ \mid \operatorname{span}\{X_0, X^N\}\|_{L^2(W^{-1})\to L^2(W^{-1})}\to 1 \qquad \text{as } N\to\infty.$$

The later is true because

$$\begin{aligned}\|P_+ \mid \mathrm{span}\{X_0, X^N\}\|_{L^2(W^{-1})\to L^2(W^{-1})} &= \|W^{-1/2}\big(P_+ \mid \mathrm{span}\{X_0, X^N\}\big)W^{1/2}\|_{L^2\to L^2} = \\ = \|W^{1/2}\big(P_+ \mid \mathrm{span}\{X_0, X^N\}\big)W^{-1/2}\|_{L^2\to L^2} &= \|P_+ \mid \mathrm{span}\{X_0, X^N\}\|_{L^2(W)\to L^2(W)}\end{aligned}$$

and

$$\|P_+ \mid \mathrm{span}\{X_0, X^N\}\|_{L^2(W^{-1})\to L^2(W)} \to 1 \qquad \text{as } N\to\infty$$

(since X_0 and X^N are asymptotically orthogonal in $L^2(W)$). □

5. Vanishing Carleson measures

Recall that $W(\lambda)$ and $W^{-1}(\lambda)$ denote harmonic extensions at the point $\lambda \in \mathbb{D}$ of the weights W and W^{-1} respectively.

Lemma 5.1. *Let a matrix weight W satisfy*

$$\lim_{|\lambda|\to 1} \|W(\lambda)^{1/2}\big(W^{-1}\big)(\lambda)^{1/2}\| = 1.$$

Then

$$\limsup_{|\lambda|\to 1} \Big\{\det\Big(W(\lambda)\Big)\exp\Big(-\big[\log\det W\big](\lambda)\Big)\Big\} = 1\,.$$

Proof. First of all let us notice that the assumption of the lemma implies that $W, W^{-1} \in L^1(\mathbb{T})$, therefore $\log(\det W) \in L^1(\mathbb{T})$. Therefore there exists a factorization $W = F^*F$ a.e. on $\mathbb{T}$, where F is an outer function in $H^2(M_{d\times d})$

Since F is an outer function in H^2, $\det F$ is an outer function in $H^{2/d}$. Therefore

$$|\det F(z)| = \exp\{(\log|\det F|)(z)\} = \exp\left\{\frac{1}{2}(\log\det W)(z)\right\} \tag{5.1}$$

It is well known fact that $F^*(z)F(z) \le W(z)$ for any $z\in D$, where $\le$ means the inequality for quadratic forms. There are many proofs of this fact, for example it admits a very simple operator-theoretic interpretation which is in fact hidden in the proof of Theorem 4.1. Explanation that we present here is more function-theoretic: Direct computation shows that

$$\Delta\left(F(z)^*F(z)\right) = 4\left(\bar\partial F(z)^*\right)\left(\partial F(z)\right) = 4\left(\partial F(z)\right)^*\left(\partial F(z)\right) \ge 0\,,$$

so for any $e\in\mathbb{C}^d$ the function $\|F(z)e\|^2$ is subharmonic and coincide with $(W(\xi)e, e)$ on $\mathbb{T}$.

We can do the same factorization for W^{-1}. Namely, let G be an outer matrix-valued function in $H^2(M_{d\times d})$ such that $W^{-1} = G^*G$ on $\mathbb{T}$. We should point out to the reader that in general G does not necessarily coincide with F^{-1}. However, applying (5.1) to G one can conclude that

$$|\det G(z)| = \exp\left\{\frac{1}{2}\left(\log\det W^{-1}\right)(z)\right\} = |\det F(z)|^{-1} \tag{5.2}$$

Now we are in position to prove the lemma. By the assumption

$$\lim_{|z|\to 1}\left\|W(z)^{1/2}(W^{-1})(z)^{1/2}\right\| = 1, \tag{5.3}$$

and therefore,

$$\lim_{|z|\to 1}\left|\det(W(z))\det\left((W^{-1})(z)\right)\right| = 1$$

Using (5.2) one can rewrite the last identity as

$$\lim_{|z|\to 1} \left\{ \left[\det W(z)/|\det F(z)|^2\right] \left[\det W^{-1}(z)/|\det G(z)|^2\right] \right\} = 1$$

Since $F(z)^*F(z) \le W(z)$ and $G(z)^*G(z) \le W^{-1}(z)$, expressions in brackets are at least 1, so, taking into account (5.1)

$$\lim_{|z|\to 1} \left[\det W(z)/\exp\left\{(\log\det W)(z)\right\}\right] = 0$$

or equivalently

$$\lim_{|z|\to 1} \log\left\{\det(W(z))\right\} - (\log\det W)(z) = 0\,. \tag{5.4}$$

□

Theorem 5.2. *A matrix weight W satisfy*

$$\limsup_{|\lambda|\to 1} \left\{ \det\Big(W(\lambda)\Big) \exp\Big(-\big[\log\det W\big](\lambda)\Big) \right\} = 1$$

if and only if the measures

$$\left\| W(z)^{-1/2} \left(\frac{\partial}{\partial x} W(z)\right) W(z)^{-1/2} \right\|^2 (1-|z|^2)dxdy$$

and

$$\left\| W(z)^{-1/2} \left(\frac{\partial}{\partial y} W(z)\right) W(z)^{-1/2} \right\|^2 (1-|z|^2)dxdy$$

are vanishing Carleson measures.

The implication $3 \Longrightarrow 4$ of Theorem 1.3 follows immediately from Theorem 5.2 and Lemma 5.1.

To prove the theorem we need the following well known description of vanishing Carleson measures

Lemma 5.3. *A measure μ in the unit disk $\mathbb{D}$ is a vanishing Carleson measure if and only if*

$$\lim_{|\lambda|\to 1} \int_{\mathbb{D}} \frac{1-|\lambda|^2}{|1-\bar{\lambda} z|^2} d\mu(z) = 0.$$

We also need the following lemma that was proved in [9], see Lemma 3.1 there.

Lemma 5.4. *Let W be a harmonic function of n variables with values in the space of strictly positive $d\times d$ matrices ($W(x) = W(x)^* > 0\ \forall x$). Then*

$$\Delta\left(\log(\det W)\right) = -\sum_{j=1}^{n} \operatorname{trace}\left((W^{-1/2}\frac{\partial W}{\partial x_j} W^{-1/2})^2 \right)$$

Proof of Theorem 5.2. The proof below follows the lines of the proof of Theorem 3.2 of [9].

By Green's formula and Lemma 5.4

$$\log\{\det(W(s))\} - (\log\det W)(s) = -\frac{1}{2\pi}\iint_{\mathbb{D}} \log\left|\frac{1-\overline{s}z}{z-s}\right| \Delta\log\{\det(W(z))\}\, dxdy =$$

$$= \frac{1}{4\pi}\iint_{\mathbb{D}} \left\{ \operatorname{trace}\left(W(z)^{-1/2}\frac{\partial W(z)}{\partial x}W(z)^{-1/2}\right)^2 + \right.$$

$$\left. + \operatorname{trace}\left(W(z)^{-1/2}\frac{\partial W(z)}{\partial y}W(z)^{-1/2}\right)^2 \right\} \log\left|\frac{1-\overline{s}z}{z-s}\right|^2 dxdy$$

Using an elementary inequality $\log(1/a) \geq 1-a$ for $0 < a \leq 1$ and the fact that $\|A\| \leq \operatorname{trace} A$ for a non-negative matrix A, the last integral is at least

$$\frac{1}{4\pi}\iint_{\mathbb{D}} \left\|W(z)^{-1/2}\frac{\partial W(z)}{\partial x}W(z)^{-1/2}\right\|^2 \log\left|\frac{1-\overline{s}z}{z-s}\right|^2 dxdy \geq$$

$$\frac{1}{4\pi}\iint_{\mathbb{D}} \left\|W(z)^{-1/2}\frac{\partial W(z)}{\partial x}W(z)^{-1/2}\right\|^2 \left(1 - \left|\frac{1-\overline{s}z}{z-s}\right|^2\right) dxdy =$$

$$= \iint_{\mathbb{D}} \left\|W(z)^{-1/2}\frac{\partial W(z)}{\partial x}W(z)^{-1/2}\right\|^2 \cdot \frac{(1-|s|^2)(1-|z|^2)}{|1-\overline{s}z|^2}\, dxdy$$

Together with (5.4) this imply

$$\lim_{|s|\to 1}\iint_{\mathbb{D}} \frac{(1-|s|^2)}{|1-\overline{s}z|^2} \cdot \left\|W(z)^{-1/2}\frac{\partial W(z)}{\partial x}W(z)^{-1/2}\right\|^2 (1-|z|^2)\, dxdy = 0$$

that yields that the measure $\left\|W(z)^{-1/2}\left(\frac{\partial}{\partial x}W(z)\right)W(z)^{-1/2}\right\|^2(1-|z|^2)\,dxdy$ is a vanishing Carleson measure.

The measure $\left\|W(z)^{-1/2}\left(\frac{\partial}{\partial y}W(z)\right)W(z)^{-1/2}\right\|^2(1-|z|^2)\,dxdy$ is treated similarly.

To prove the opposite implication, let us estimate the integral

$$\iint_{\mathbb{D}} \operatorname{trace}\left(W(z)^{-1/2}\frac{\partial W(z)}{\partial x}W(z)^{-1/2}\right)^2 \log\left|\frac{1-\overline{s}z}{z-s}\right|^2 dxdy$$

(the integral with $\partial W/\partial y$ can be estimated absolutely the same way). Denote by b_s a Blaschke factor with zero at the point s, $b_s(z) = (z-s)(1-\overline{s}z)^{-1}$.

First of all, we can estimate the trace by $d\cdot\|\cdot\|$, where d is dimension of the space. So we can estimate the integral by

$$C\iint_{\mathbb{D}} \left\|W(z)^{-1/2}\frac{\partial W(z)}{\partial x}W(z)^{-1/2}\right\|^2 \log|b_s(z)|^{-2}dxdy = \iint_{|b_s(z)|<\varepsilon}\dots + \iint_{|b_s(z)|\geq\varepsilon}\dots$$

To estimate the second integral we notice that

$$\log|b_s(z)|^{-2}dxdy \leq C(\varepsilon)\frac{(1-|s|^2)(1-|z|^2)}{|1-\overline{s}z|^2}$$

for $|b_s(z)| \geq \varepsilon$, and since the measure is a vanishing Carleson measure we can make the integral as small as we want when $|s| \to 1$.

To estimate the first integral let make a trivial observation: if $w \in L^1(\mathbb{T})$, $w \geq 0$ and $w(z)$ denotes its harmonic extension at the point z, then for all z such that $|z| \leq 1/2$ (and therefore for all z such that $|z| < \varepsilon \leq 1/2$)

$$\frac{\partial}{\partial x} w(z) \leq C w(0)\,,$$

where C is an absolute constant. Combining this observation with the Harnack inequality $w(0) \leq C' w(z)$, $|z| \leq 1/2$, and applying it to functions $w(.) = \big(W(\cdot)e, e\big)_{\mathbb{C}^d}$ we get the inequality for quadratic forms

$$\frac{\partial}{\partial x} W(z) \leq C\varepsilon W(0) \leq C_1 W(z)\,.$$

It in turn implies

$$\left\| W(z)^{-1/2} \Big(\frac{\partial}{\partial x} W(z)\Big) W(z)^{-1/2} \right\| \leq C_1\,, \qquad \forall z:\ |z| < \varepsilon \leq 1/2\,.$$

Using the Möbius transformation $z \mapsto b_s(z)$ we get

$$\left\| W(z)^{-1/2} \Big(\frac{\partial}{\partial x} W(z)\Big) W(z)^{-1/2} \right\| \leq C\varepsilon\,, \qquad \forall z:\ |b_s(z)| < \varepsilon \leq 1/2\,.$$

Since

$$\iint_{|b_s(z)| \leq \varepsilon} \log |b_s(z)|^{-2} dxdy \leq C\varepsilon^2 \log \frac{1}{\varepsilon}\,,$$

we can estimate the first integral by $C\varepsilon^2 \log(1/\varepsilon)$; we can make this number as small as we want by picking sufficiently small ε.

□

6. Embedding theorem and equivalent norms

By analogy with the scalar case (see [10]) we will say that a matrix weight W satisfies the *invariant* A_∞ *condition* if

$$\sup_{s \in \mathbb{D}} \Big\{ \det\Big(W(s)\Big) \exp\Big(-\big[\log\det W\big](s)\Big) \Big\} < \infty\,. \qquad (invA_\infty)$$

The *supremum* is called the invariant A_∞ norm of W.

Theorem 5.2 implies that if the measures

$$\left\| W(z)^{-1/2} \left(\frac{\partial}{\partial x} W(z)\right) W(z)^{-1/2} \right\|^2 (1 - |z|^2) dxdy$$

and

$$\left\| W(z)^{-1/2} \left(\frac{\partial}{\partial y} W(z)\right) W(z)^{-1/2} \right\|^2 (1 - |z|^2) dxdy$$

are vanishing Carleson measures then the weight W satisfies the invariant A_∞ condition.

Literally repeating the proof of Theorem 5.2 one can obtain that the weight W satisfies the invariant A_∞ condition if and only if the above measures are Carleson.

We will need the following "embedding theorem". More general result was proved in [9], Lemma 4.1.

Lemma 6.1. *Let W be a matrix weight satisfying the invariant A_∞ condition, and let μ be a Carleson measure with the Carleson norm $\|\mu\|_C$. Then for any analytic (or antianalytic) vector-function f, the following inequality holds,*

$$\iint_{\mathbb{D}} (W(z)f(z), f(z))\, d\mu(z) \le C\|\mu\|_C \int_{\mathbb{T}} (W(\xi)f(\xi), f(\xi)) dm(\xi)\,,$$

where the constant C depends the dimension d and the invariant A_∞ norm of W.

Proof. The invariant A_∞ condition implies that $\log\det W \in L^1$, so there exists (see [6]) an outer function $F \in H^2(M_{d\times d})$ such that $W = F^*F$. It is well known (see again [6]) that

$$|\det F(z)| = \exp\left\{\frac{1}{2}\big[\log\det W\big](z)\right\}\,.$$

It is well known and it was already shown it in the proof of Lemma 5.1 that $F(z)^*F(z) \le W(z)$. Hence

$$\|W(z)^{1/2}F(z)^{-1}e\| \ge \|e\|, \qquad e \in \mathbb{C}^d\,. \tag{6.5}$$

Since

$$\left|\det\left\{W(z)^{1/2}F(z)^{-1}\right\}\right| = \left\{\det\Big(W(\lambda)\Big)\exp\Big(-\big[\log\det W\big](\lambda)\Big)\right\}^{1/2} \le C$$

we can estimate

$$\|W(z)^{1/2}F(z)^{-1}e\| \le C\,.$$

Together with (6.5) it implies that $(W(z)e, e)$ and $\|F(z)e\|^2$ are equivalent in a sense of two-sided estimate. Therefore

$$\iint_{\mathbb{D}} (W(z)f(z), f(z))\, d\mu(z) \le C \iint_{\mathbb{D}} (F(z)f(z), F(z)f(z))\, d\mu(z) \le$$
$$\le C\|\mu\|_C \int_{\mathbb{T}} (F(\xi)f(\xi), F(\xi)f(\xi)) dm(\xi) = C\|\mu\|_C \int_{\mathbb{T}} (W(\xi)f(\xi), f(\xi)) dm(\xi)\,.$$

□

We also need the following simple lemma.

Lemma 6.2 (equivalence of weighted norms)**.** *Let W be a matrix weight satisfying the invariant A_∞ condition. There exist a constant C such that for any analytic or antianalytic vector-function f in $L^2(W)$ satisfying $f(0) = 0$*

$$\frac{1}{C}\int_{\mathbb{T}} (Wf, f) dm \le \iint_{\mathbb{D}} (W(z)f'(z), f'(z)) \log\frac{1}{|z|}\, dxdy \le C\int_{\mathbb{T}} (Wf, f) dm$$

Proof. Let us recall the the operators ∂ and $\overline{\partial}$ are defined as

$$\partial f = \frac{1}{2}\left(\frac{\partial f}{\partial x} - i\frac{\partial f}{\partial y}\right), \qquad \overline{\partial} f = \frac{1}{2}\left(\frac{\partial f}{\partial x} + i\frac{\partial f}{\partial y}\right)\,.$$

Recall that for analytic functions $\partial f = f'$ and $\overline{\partial} f = 0$.

Let f be an analytic function, $f(0) = 0$. Using the Green's formula and taking into account that $f(0) = 0$ and $\Delta = 4\partial\overline{\partial} = 4\overline{\partial}\partial$ we get

$$\int_{\mathbb{T}} \big(Wf, f\big) dm = \frac{1}{2\pi} \iint_{\mathbb{D}} \Delta\big(W(z)f(z), f(z)\big) \log\frac{1}{|z|} dxdy =$$
$$\frac{2}{\pi} \iint_{\mathbb{D}} \big(\overline{\partial} W(z) f'(z), f(z)\big) \log\frac{1}{|z|} dxdy + \frac{2}{\pi} \iint_{\mathbb{D}} \big(\partial W(z) f(z), f'(z)\big) \log\frac{1}{|z|} dxdy +$$
$$+ \frac{2}{\pi} \iint_{\mathbb{D}} \big(W(z) f'(z), f'(z)\big) \log\frac{1}{|z|} dxdy = \frac{2}{\pi}(\mathcal{I}_1 + \mathcal{I}_2 + \mathcal{I}_3)$$

The last integral $\mathcal{I}_3$ is exactly the integral we want to estimate. Let us denote $A^2 := \int_{\mathbb{T}} \big(Wf, f\big) dm$, $B^2 := \mathcal{I}_3$. We want to show that $A \asymp B$ in a sense of two sided estimate. Let us estimate $\mathcal{I}_1$:

$$|\mathcal{I}_1| = \left| \iint_{\mathbb{D}} \big(W(z)^{-1/2}\overline{\partial} W(z) W(z)^{-1/2} W(z)^{1/2} f'(z), W(z)^{1/2} f(z)\big) \log\frac{1}{|z|} dxdy \right| \leq$$
$$\leq \left| \iint_{\mathbb{D}} \big\|W(z)^{-1/2}\overline{\partial} W(z) W(z)^{-1/2}\big\| \cdot \big\|W(z)^{1/2} f'(z)\big\| \cdot \big\|W(z)^{1/2} f(z)\big\| \cdot \log\frac{1}{|z|} dxdy \right| \leq$$
$$\leq \left(\iint_{\mathbb{D}} \big\|W(z)^{-1/2}\overline{\partial} W(z) W(z)^{-1/2}\big\|^2 \big(W(z)f(z), f(z)\big)_{\mathbb{C}^d} \log\frac{1}{|z|} dxdy \right)^{1/2} \times$$
$$\times \left(\iint_{\mathbb{D}} \big(W(z)f'(z), f'(z)\big)_{\mathbb{C}^d} \log\frac{1}{|z|} dxdy \right)^{1/2} .$$

The measure $\big\|W(z)^{-1/2}\overline{\partial} W(z) W(z)^{-1/2}\big\|^2 \log\frac{1}{|z|} dxdy$ is Carleson, so by Lemma 6.1 the first term in the product is estimated by KA (K is a constant). The second term is just B so $|\mathcal{I}_1| \leq KAB$. Similarly $|\mathcal{I}_2| \leq KAB$. So

$$A^2 = B^2 + \mathcal{I}_1 + \mathcal{I}_2\,,$$

where

$$|\mathcal{I}_1|,\ |\mathcal{I}_2| \leq KAB\,.$$

This immediately implies

$$\frac{1}{C}A \leq B \leq CA$$

for an appropriate choice of C. □

7. Proof of the implication $6 \Longrightarrow 1$

To prove the implication $6 \Longrightarrow 1$ we need to estimate $\int_{\mathbb{T}} (Wz^n f, g) dm$, $f \in X^0$, $g \in X_0$, $\|f\|_{L^2(W)} = \|g\|_{L^2(W)} = 1$.

Using the Green's formula and taking into account that $g(0)=0$ and $\Delta = 4\partial\overline{\partial} = 4\overline{\partial}\partial$ we get

$$\begin{aligned}\int_{\mathbb{T}}(Wz^n f, g)dm &= \frac{1}{2\pi}\iint_{\mathbb{D}} \Delta\big(W(z)z^n f(z), g(z)\big)_{\mathbb{C}^d} \log\frac{1}{|z|}dxdy = \\ &= \frac{2}{\pi}\iint_{\mathbb{D}} \big(\overline{\partial}W(z)\partial(z^n f(z)), g(z)\big)_{\mathbb{C}^d} \log\frac{1}{|z|}dxdy + \\ &\qquad + \frac{2}{\pi}\iint_{\mathbb{D}} \big(\overline{\partial}W(z)(z^n f(z)), \overline{\partial}g(z)\big)_{\mathbb{C}^d} \log\frac{1}{|z|}dxdy = \frac{2}{\pi}(\mathcal{I}_1+\mathcal{I}_2)\end{aligned}$$

The second integral is easy to estimate:

$$\begin{aligned}|\mathcal{I}_2| &= \left|\iint_{\mathbb{D}} \big(W(z)^{-1/2}\overline{\partial}W(z)W(z)^{-1/2}W(z)^{1/2}(z^n f(z)), W(z)^{1/2}\overline{\partial}g(z)\big)_{\mathbb{C}^d} \log\frac{1}{|z|}dxdy\right| \le \\ &\le \iint_{\mathbb{D}} \|W(z)^{-1/2}\overline{\partial}W(z)W(z)^{-1/2}\| \cdot \|W(z)^{1/2}(z^n f(z))\| \cdot \|W(z)^{1/2}\overline{\partial}g(z)\| \log\frac{1}{|z|}dxdy \le \\ &\le \left(\iint_{\mathbb{D}} |z|^{2n}\cdot \|W(z)^{-1/2}\overline{\partial}W(z)W(z)^{-1/2}\|^2 \cdot \big(W(z)f(z), f(z)\big)_{\mathbb{C}^d} \log\frac{1}{|z|}dxdy\right)^{1/2} \times \\ &\qquad \times \left(\iint_{\mathbb{D}} \big(W(z)\overline{\partial}g(z), \overline{\partial}g(z)\big)_{\mathbb{C}^d} \log\frac{1}{|z|}dxdy\right)^{1/2}\end{aligned}$$

The last term is equivalent to the norm $\|g\|_{L^2(W)}$ (see Lemma 6.2), so by Lemma 6.1

$$|\mathcal{I}_2| \le \|f\|_{L^2(W)} \cdot \|g\|_{L^2(W)} \cdot \left\| |z|^2 \cdot \|W(z)^{-1/2}\overline{\partial}W(z)W(z)^{-1/2}\| \log\frac{1}{|z|}dxdy \right\|_C^{1/2}$$

Since the measure $\|W(z)^{-1/2}\overline{\partial}W(z)W(z)^{-1/2}\| \log\frac{1}{|z|}dxdy$ is a vanishing Carleson measure, the Carleson norm $\left\| |z|^2 \cdot \|W(z)^{-1/2}\overline{\partial}W(z)W(z)^{-1/2}\| \log\frac{1}{|z|}dxdy \right\|_C^{1/2} \to 0$ as $n\to\infty$. So $|\mathcal{I}_2| \to 0$ as $n\to\infty$.

To estimate $\mathcal{I}_1$ we pick $r<1$ close to 1 and split the integral into two: $\mathcal{I}_1 = \iint_{r\mathbb{D}} \ldots + \iint_{\mathbb{D}\setminus r\mathbb{D}} \ldots$. Acting as with $\mathcal{I}_2$ we can estimate

$$\begin{aligned}&\left|\iint_X \ldots\right| \le \\ &\quad \le \left(\iint_X \cdot\|W(z)^{-1/2}\overline{\partial}W(z)W(z)^{-1/2}\|^2 \cdot \big(W(z)g(z), g(z)\big)_{\mathbb{C}^d} \log\frac{1}{|z|}dxdy\right)^{1/2} \times \\ &\qquad\qquad \times \left(\iint_X \big(W(z)\partial\big(z^n f(z)\big), \partial\big(z^n f(z)\big)\big)_{\mathbb{C}^d} \log\frac{1}{|z|}dxdy\right)^{1/2},\end{aligned}$$

where X is either $r\mathbb{D}$ or $\mathbb{D}\setminus r\mathbb{D}$. Note that both terms are uniformly bounded.

We can say even more. If $X = r\mathbb{D}$ the second term can be made as small as we wish by picking sufficiently large n.

Let now $X = \mathbb{D}\setminus r\mathbb{D}$. The measure $\|W(z)^{-1/2}\overline{\partial}W(z)W(z)^{-1/2}\| \log\frac{1}{|z|}dxdy$ is a vanishing Carleson measure, so for r sufficiently close to 1 its restriction onto $\mathbb{D}\setminus r\mathbb{D}$ has the Carleson norm as small as we want. So by Lemma 6.1 the first term is as small as we want if r is sufficiently close to 1.

8. A counterexample to Peller's conjecture.

In this section we are going to construct a weight W, such that $W^{-1} \in L^1$, $\log W \in \mathrm{VMO}$, but the corresponding stationary process is not completely regular (i.e., the weight W does not satisfy any of the conditions 1–6 of Theorem 1.3).

Let

$$W = U^* \begin{pmatrix} 1 & 0 \\ 0 & \delta(z) \end{pmatrix} U \,, \qquad U = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} .$$

Here

$$\delta(e^{it}) = 1/\log(1/|t|), \qquad -1/4 \le t \le 1/4,$$

and δ is a continuous function bounded away from 0 and ∞ on the rest of the circle, and

$$\alpha(e^{it}) = (t/|t|)\delta(e^{it})^{1/4}, \qquad -1/4 \le t \le 1/4,$$

and again α is continuous on the rest of the circle.

Then

$$\log W = U^* \begin{pmatrix} 0 & 0 \\ 0 & \log\delta \end{pmatrix} U = \begin{pmatrix} \sin^2\alpha \log\delta & \sin\alpha\cos\alpha\log\delta \\ \sin\alpha\cos\alpha\log\delta & \cos^2\alpha\log\delta \end{pmatrix},$$

and this matrix clearly belongs to VMO: $\log\delta = \log\log 1/|t|$ (considered only in a neighborhood of 0) is a "typical" unbounded function in VMO, so $\cos^2\alpha\log\delta \in \mathrm{VMO}$, and all other entries of the matrix are continuous.

Let us now show that the weight W does not even satisfies the Muckenhoupt condition (A_2). Direct computations show that

$$W = \begin{pmatrix} \cos^2\alpha & -\sin\alpha\cos\alpha \\ -\sin\alpha\cos\alpha & \sin^2\alpha \end{pmatrix} + \delta \begin{pmatrix} \sin^2\alpha & \sin\alpha\cos\alpha \\ \sin\alpha\cos\alpha & \cos^2\alpha \end{pmatrix}$$

and

$$W^{-1} = \begin{pmatrix} \cos^2\alpha & -\sin\alpha\cos\alpha \\ -\sin\alpha\cos\alpha & \sin^2\alpha \end{pmatrix} + \delta^{-1} \begin{pmatrix} \sin^2\alpha & \sin\alpha\cos\alpha \\ \sin\alpha\cos\alpha & \cos^2\alpha \end{pmatrix}$$

If we pick I to be a symmetric arc $[e^{-i\varepsilon}, e^{i\varepsilon}]$ ($\varepsilon > 0$ is small), then off-diagonal entries of W_I and $(W^{-1})_I$ equal 0, and so we can estimate

$$W_I \ge C \begin{pmatrix} \cos^2\alpha(\varepsilon) & 0 \\ 0 & \sin^2\alpha(\varepsilon) \end{pmatrix} , \qquad (W^{-1})_I \ge C \begin{pmatrix} \delta(\varepsilon)^{-1}\sin^2\alpha(\varepsilon) & 0 \\ 0 & \delta(\varepsilon)^{-1}\cos^2\alpha(\varepsilon) \end{pmatrix} .$$

Therefore

$$\big\| [W_I]^{1/2} [(W_I^{-1}]^{1/2} \big\| \ge C\delta(\varepsilon)^{-1} \sin\alpha(\varepsilon)\cos\alpha(\varepsilon) \to \infty \qquad \text{as} \quad \varepsilon \to 0 \,.$$

References

[1] Yu. A. Rozanov, "Stationary stochastic processes," Holden-day, SF, 1967.

[2] I. A. Ibragimov, Completely regular multidimensional stationary processes with discrete time, *Proc. Steklov Inst. Math.*, **111** (1970), 269–301.

[3] H. Helson and D. Sarason Past and Future, *Math. Scand.,* **21** (1967), 5–16.

[4] V. Peller, Hankel operators and multivariate stationary processes, *Operator theory: operator algebras and applications, Part 1* (Durham, NH, 1988), 357–371, Proc. Sympos. Pure Math., 51, Part 1, Amer. Math. Soc., Providence, RI, 1990.

[5] J. B. Garnett, "Bounded analytic functions," Acad. Press, NY, 1981.

[6] M. Rosenblum and J. Rovnyak, "Hardy classes and operator theory," (Oxford Mathematical Monographs) Oxford Science Publications. The Clarendon Press, Oxford University Press, New York, 1985.

[7] D. Sarason An addendum to "Past and Future", *Math. Scand.,* **30** (1972), 62–64.

[8] S. Treil and A. Volberg Wavelets and the angle between past and future, *accepted by Journal of functional analysis.*

[9] S. Treil and A. Volberg Continuous frame decomposition and a vector Hunt – Muckenhoupt – Wheeden Theorem, *accepted by Arkiv för Matematik.*

[10] S. Treil, A. Volberg and D. Zheng, Hilbert transform, Toeplitz operators and Hankel operators, and invariant A_∞ weights, *to appear in Rev. Mat. Iberoamericana*