Ark. Mat., **35** (1997), 363–386

Continuous frame decomposition and a vector Hunt–Muckenhoupt–Wheeden theorem

Sergei Treil and Alexander Volberg

Introduction

Statement of the problem. This paper deals with the weighted norm inequalities for the Hilbert transform with matrix-valued weights. The main problem can be formulated as follows. Let W be a $d\times d$ matrix weight, i.e. an L^1 function whose values are selfadjoint nonnegative $d\times d$ matrices. We suppose that the weight W is defined on the unit circle $\mathbf{T}=\{z\in\mathbf{C}:|z|=1\}$.

Let $L^2=L^2(\mathbf{C}^d)$ be the space of square summable functions on $\mathbf{T}$ with values in $\mathbf{C}^d$, let $H^2=H^2(\mathbf{C}^d)$ be the corresponding Hardy space of analytic functions, and let P_+ be the orthogonal projection in L^2 onto H^2. Let T denote the Hilbert transform, $T=-iP_++i(I-P_+)$.

The question we are interested in is under what conditions on W the following weighted norm inequality for T holds (say for all $f\in L^2\cap L^\infty$),

$$\int_{\mathbf{T}}(W(\xi)Tf(\xi),Tf(t))\,dt\le C\int_{\mathbf{T}}(W(\xi)f(\xi),f(\xi))\,dm(\xi),$$

where m is the normalized ($m(\mathbf{T})=1$) Lebesgue measure on $\mathbf{T}$. Clearly this inequality is equivalent to the same inequality for P_+ (with another constant).

If we define a weighted space $L^2(W)$ as the space of all measurable $\mathbf{C}^d$-valued functions on $\mathbf{T}$ satisfying

$$\|f\|^2_{L^2(W)}\stackrel{\text{def}}{=}\int_{\mathbf{T}}(W(\xi)f(\xi),f(\xi))\,dm(\xi)<\infty,$$

(of course we should factorize it over the subspace of functions of norm 0), then the last inequality means that T (or, equivalently P_+) is a bounded operator in $L^2(W)$.

(1) Partially supported by the grant DMS9622936 from the National Science Foundation.

If the dimension d equals 1 and everything is scalar-valued, the answer is given by the famous Hunt–Muckenhoupt–Wheeden theorem, which says that the Muckenhoupt condition

$$(A_2) \qquad \sup_I \left(\frac{1}{|I|}\int_I W\right)\left(\frac{1}{|I|}\int_I W^{-1}\right) < \infty$$

(supremum is taken over all intervals) is necessary and sufficient for boundedness of the Hilbert transform in $L^2(W)$.

In the matrix case it was conjectured by the first author in [T] and proved in our paper [TV1] that the vector Muckenhoupt condition

$$(\mathbf{A}_2) \qquad \sup_I \left\| \left(\frac{1}{|I|}\int_I W\,dm\right)^{1/2} \left(\frac{1}{|I|}\int_I W^{-1}\,dm\right)^{1/2} \right\| < \infty$$

is necessary and sufficient for boundedness of the Hilbert transform in $L^2(W)$ with matrix weight. In this paper we present an alternative proof of that result.

The main technical tool we are using here is a matrix version of a Littlewood–Paley type inequality that gives an equivalent norm in the weighted L^2 space in terms of a weighted L^2 norm of the derivative of the harmonic extension (see Theorem 3.2 below). The scalar version was developed by us earlier in [TV2].

This equivalent norm inequality can be viewed as a continuous analogue of the wavelet type decomposition (the Haar system is a Riesz basis in $L^2(W)$) that was used by us in [TV1]. But in this case a continuous "system of coefficients" (derivatives of harmonic extension) is over-determined, so it is more appropriate to call it a continuous frame decomposition (see [D] for a discussion of frames).

Although the main result about boundedness of the Riesz projection is already known, we feel that the technique we use is of independent interest and deserves separate consideration. In a sense the main result is Theorem 3.2 below about continuous frame decomposition, and boundedness of the Riesz projection serves as an illustration of the usefulness of this theorem.

The advantage of the proof presented here is its simplicity. The disadvantage is that it is applicable only to the Hilbert case $p=2$ as far as we can see. In a sense we are just presenting a complex analytic proof of the Hunt–Muckenhoupt–Wheeden theorem for the case $p=2$. The matrix nature of weights and the underlying non-commutativity of the problem make our task more difficult. On the other hand the proofs below are new even in the scalar case. The reader who does not like matrices may restore the scalar proof just by replacing (Wf,f) by $|f|^2W$ everywhere or may be referred to [TV2] where this has been done. One can obtain the preprint [TV2] from our homepages: `http://www.mth.msu.edu/~treil/` and `http://www.mth.msu.edu/~volberg/`.

The difficulties. We would like to mention that the problem cannot be reduced to the scalar case. The proof for the scalar case cannot be reproduced for matrix valued weights. The main reason is that the original proof by Hunt–Muckenhoupt–Wheeden [HMW] and all its modifications (see e.g. [GCRF] and [St]) make extensive use of maximal functions. And it is very difficult for us to imagine how one can introduce *working* maximal functions for matrix-valued weights. Indeed, for scalar weighted spaces we have a very simple but wonderful fact that a function f belongs $L^p(\mu)$ if and only if $|f|\in L^p(\mu)$. We do not have an analogue of this for matrix-valued weights, even for $p=2$.

As an illustration of what kind of difficulties one can encounter in the vector case, let us present several very simple examples. It is trivial in the scalar case that if we have an integral operator in $L^2(W)$ with positive (scalar) kernel, and we know that an operator with a larger kernel is bounded, then the original operator is bounded too. This statement (even for scalar kernels) does not hold for weighted L^2-spaces with matrix weights. Certainly if $W(t)$ can be diagonalized by *the same* basis for each $t\in\mathbf{T}$, we do not have any difficulties. But this is not the case we are interested in. Another difficulty stems from the fact that it is not true that every nonzero nonnegative operator on $\mathbf{C}^d$ is invertible when $d>1$, as it is when $d=1$. For example, suppose we meet the expression (we do meet such expressions while working with matrix weights) $(A(I+B)^{-1}x,(I+B)^{-1}x)$, where A and B are nonnegative operators, $B\le I$, and x is a vector in $\mathbf{C}^d$. If A, B, x were numbers, the estimate from below $\frac{1}{4}A|x|^2$ would follow. But no estimate $\delta\|x\|^2$ with $\delta>0$ exists for operators.

Carleson measures and (A_∞). The observation is that even though many scalar methods are not available now, there is at least one which still manages to survive. Let us remind the reader of the role of (A_∞) weights. A positive function cannot be much larger than its average over I on a large portion of I. If we change the word "larger" to "smaller", we get the class of (A_∞) functions (see various definitions of (A_∞) on pp. 196–197, and 218 of [St]). In the classical case (see [St, pp. 200–205]) the class (A_∞) comes into play as follows. The singular integral operator is "factorized" through the maximal operator M in the sense that the weighted estimate $\|Hf\|_W\le C\|f\|_W$ is split into estimates

$$\|Hf\|_W\le C\|Mf\|_W$$

and

$$\|Mf\|_W\le C\|f\|_W.$$

The main idea. In this paper we use a different idea. The idea is very simple, and it comes from the scalar situation treated in [TV2]. The (A_∞) condition, which will now become the matrix $(A_{2,\infty})$ condition, will be used to prove a certain estimate that connects the weighted norm of f on the circle with the weighted norm of the gradient of the harmonic extension of f to the disc.

This allows us to prove that there exists an equivalent norm in $L^2(W)$, namely that there exists a finite positive constant C such that

$$\begin{aligned}\|f\|^2_{L^2(W)} &\le \iint_{\mathbf{D}} [(W(z)\partial f(z), \partial f(z)) + (W(z)\bar\partial f(z), \bar\partial f(z))](1-|z|^2)\,dx\,dy \\ &\le C\|f\|^2_{L^2(W)}.\end{aligned}$$

Here $f(z)$ stands for the harmonic extension of the vector function f into the disc. The equivalence of this new norm is actually necessary and sufficient for $W \in (\mathbf{A}_2)$. The equivalence of norms is proved in Theorem 3.2. Now it will be trivial to prove the boundedness of P_+ and P_-. In fact, we diagonalize the operator P_+ by using the new norm; if $f = f_+ + f_- \stackrel{\text{def}}{=} P_+ f + P_- f$, then $\partial f(z) = \partial f_+(z)$ and $\bar\partial f(z) = \bar\partial f_-(z)$, which means that the new (equivalent) norm of f is just equal to the sum of the new norms of f_+ and f_-. See details in Section 4 below.

The motivation. Stochastic processes. Let us consider a multivariate random stationary process. For simplicity we consider the case of discrete time. Let $\mathcal{W}$ be the spectral measure of the process; in our case this is a measure whose values are $d \times d$ nonnegative selfadjoint matrices. The reader can think of this as of a matrix whose entries are complex measures $\mu_{i,j}$ such that for any Borel set E the matrix $\{\mu_{i,j}(E)\}_{i,j=1}^d$ is nonnegative.

The geometry of the process is described by the geometry of the sequence of subspaces $z^n \mathbf{C}^d$, $n \in \mathbf{Z}$, in a weighted space $L^2(\mathcal{W}) = L^2(\mathcal{W}, \mathbf{C}^d)$. The space $L^2(\mathcal{W})$ consists of all functions on $\mathbf{T}$ with values in $\mathbf{C}^d$ such that

$$\|f\|^2_{L^2(\mathcal{W})} \stackrel{\text{def}}{=} \int (d\mathcal{W}(t) f(t), f(t)) < \infty.$$

In this representation the past of the process is $\operatorname{span}\{z^n \mathbf{C}^d : n < 0\}$ and the future is $\operatorname{span}\{z^n \mathbf{C}^d : n \ge 0\}$; the angle between past and future is nonzero if and only if the Riesz projection P_+ is bounded in the weighted space $L^2(\mathcal{W})$. This property for stationary Gaussian processes (the angle between past and future being nonzero) is in probability literature called *uniform mixing of past and future* (see [R]).

If the angle between past and future is nonzero, then for any vector $\mathbf{e} \in \mathbf{C}^d$ the angle between the subspaces $\operatorname{span}\{z^n \mathbf{e} : n < 0\}$ and $\operatorname{span}\{z^n \mathbf{e} : n \ge 0\}$ in $L^2(\mathcal{W})$ is nonzero. If a measure μ is defined on Borel sets E by $\mu(E) = (\mathcal{W}(E)\mathbf{e}, \mathbf{e})$ the last condition means that the angle between antianalytic polynomials $\operatorname{span}\{z^n : n < 0\}$ and analytic polynomials $\operatorname{span}\{z^n : n \ge 0\}$ in the (scalar) weighted space $L^2(\mu)$ is positive. Equivalently, we can say that the Riesz projection P_+ (or Hilbert transform T) is bounded in the weighted space $L^2(\mu)$. It is well known that this is possible if and only if the measure μ is absolutely continuous and its density w satisfies the (scalar) Muckenhoupt condition (A_2).

Therefore, if the angle between past and future is positive, the spectral measure $\mathcal{W}$ of the process is necessarily absolutely continuous, and the question about the angle gives rise to our problem.

One can also consider completely regular multivariate stochastic processes and try to characterize complete regularity in terms of spectral measure. For scalar processes this has been done in the article of Helson and Sarason [HS]. For multivariate processes we did that in [TV3], answering the question of V. Peller.

Similarly, if we consider stationary processes with continuous time, we arrive at the problem about Hilbert transform on $\mathbf{R}$.

Operator theory motivation. There is a part of the theory of singular integrals that treats Hardy spaces in $\mathbf{R}^n$. The passage from $\mathbf{R}^1$ (or $\mathbf{T}$) to $\mathbf{R}^n$ makes the theory immensely richer. At the same time the theory of vector valued Hardy spaces on $\mathbf{T}$ was developed for the needs of spectral theory of operators (see [N]), because the dilation theory of linear contractions (see [N]) reduces questions about bounded operators in Hilbert space to function-theoretic questions in a vector Hardy space. Even the finite dimensional case is known to be much richer than the scalar case (see [N] again). So the increase in dimension in this direction also enriches the theory. The connection with singular integrals becomes manifest if one considers Hankel and Toeplitz operators, given by formulae $H_F f \stackrel{\text{def}}{=} P_-(Ff) = (I-P_+)(Ff)$; and $T_F f \stackrel{\text{def}}{=} P_+(Ff)$, where F is a $d\times d$ matrix function and $f\in H^p(\mathbf{C}^d)$. One such problem that is very difficult already in the scalar case $d=1$ was considered in [ACS], and then a similar problem was considered in [S1], [S2]. It is closely related to a two-weight estimate for the Hilbert transform that is still open (see e.g. [TVZ], and the literature cited there).

There is another classical problem for Toeplitz operators that leads to a weighted estimate (with one weight) for the Hilbert transform. In fact, the invertibility of a Toeplitz operator T_F on H^p is equivalent (see [Si]) to the factorization $F=G_1^* G_2$, where G_1, G_2 are $d\times d$ outer matrix functions such that the following estimate holds

$$(0.1)\qquad \int_{\mathbf{T}} (V(t)P_+ f(t), P_+ f(t))\, dm(t) \le C \int_{\mathbf{T}} (W(t)f(t), f(t))\, dm(t),$$

where $W=G_1G_1^*$, $V=(G_2^{-1})^*G_2^{-1}$.

At first glance the matrix weights $W=G_1G_1^*$, $V=(G_2^{-1})^*G_2^{-1}$ seem to be different, but the invertibility of T_F implies easily that $F^{-1}\in L^\infty_{d\times d}$, which means that the matrix weights V, W are equivalent in the sense that there exists a constant C such that for all $\mathbf{e}\in\mathbf{C}^d$ and for almost all $t\in\mathbf{T}$

$$\frac{1}{C}(V(t)\mathbf{e},\mathbf{e})\leq(W(t)\mathbf{e},\mathbf{e})\leq C(V(t)\mathbf{e},\mathbf{e}).$$

This is how one can come to the matrix weighted norm inequality considered in this paper.

As far as we know, the first results about the matrix weight inequality were obtained by Steven Bloom [B1], [B2], who noticed that if the matrix weight W is assumed to be appropriately "smooth", then it can be diagonalized by a "smooth" unitary matrix function; furthermore, the operator of multiplication on this unitary matrix function commutes with P_+ up to a compact term (because of "smoothness"). This approach leads to pointwise diagonalization of the estimate under consideration and so to the corresponding scalar problem. In the present work the matrix function W is a priori arbitrary. Rather than doing pointwise diagonalization (which is not available now) we prefer to come to global (almost) diagonalization of our operator in the weighted space $L^2(W)$.

Acknowledgements. We are grateful to the referee whose remarks considerably improved this work. We are also grateful to Sheldon Axler who kindly agreed to read the paper and who made many useful remarks.

1. Properties of matrix $(\mathbf{A_2})$ weights

Given a matrix weight W and a set $I\subset\mathbf{R}$ let us denote by $W(I)$ the integral $W(I)\stackrel{\text{def}}{=}\int_I W\,dm$ and by W_I the average value $W_I\stackrel{\text{def}}{=}W(I)/|I|$, where m and $|\cdot|$ denotes normalized $(m(\mathbf{T})=1)$ Lebesgue measure on $\mathbf{T}$. Also for $\lambda\in\mathbf{D}$ let $W(\lambda)$ denote the harmonic (Poisson) extension of W at λ,

$$W(\lambda):=\int_{\mathbf{T}}\frac{1-|\lambda|^2}{|1-\bar{\lambda}\xi|^2}W(\xi)\,dm(\xi)$$

It was shown in [TV1] that if the Riesz projection P_+, $P_+(\sum_{-\infty}^{\infty}\hat{f}(k)z^k)=\sum_0^{\infty}\hat{f}(k)z^k$ is bounded in the vector-valued weighted space $L^2(W)$ then the matrix weight W satisfies

$$\sup_{\lambda\in\mathbf{D}}\|W(\lambda)^{1/2}W^{-1}(\lambda)^{1/2}\|<\infty. \tag{1.2}$$

The last condition implies (see the second part of Lemma 2.2 of [TV1])

$$\sup_I\|[W_I]^{1/2}[(W^{-1})_I]^{1/2}\|<\infty, \tag{$\mathbf{A}_2$}$$

where the *supremum* is taken over all arcs $I\subset\mathbf{T}$. The last condition is called the vector Muckenhoupt condition $(\mathbf{A}_2)$, and the supremum is called the Muckenhoupt norm of W.

We will need some properties of Muckenhoupt weights. One such property is that $(\mathbf{A}_2)$ implies (1.2). Of course it follows immediately from the fact that P_+ is bounded in the weighted space $L^2(W)$. We shall however present a direct and much simpler proof (see Corollary 1.4 below).

Lemma 1.1. *Let W be a nonnegative, measurable, $d\times d$ matrix function on a measure space $\mathcal{X},\mu$. Then*

$$\int_{\mathcal{X}} \|W(t)\|\,d\mu(t)\le d\left\|\int_{\mathcal{X}} W(t)\,d\mu(t)\right\|.$$

Proof. Since for a nonnegative $d\times d$ matrix A we have $\|A\|\le\operatorname{trace}A\le d\|A\|$, we can conclude that

$$\begin{aligned}\int_{\mathcal{X}} \|W(t)\|\,d\mu(t)&\le\int_{\mathcal{X}}\operatorname{trace}(W(t))\,d\mu(t)\\&=\operatorname{trace}\left(\int_{\mathcal{X}} W(t)\,d\mu(t)\right)\le d\left\|\int_{\mathcal{X}} W(t)\,d\mu(t)\right\|.\quad\square\end{aligned}$$

Lemma 1.2. *Let W be a $d\times d$ matrix Muckenhoupt weight, and A be a positive nonsingular $d\times d$ (constant) matrix. Then the weight $\mathcal{W}{:=}AWA$ also satisfies the Muckenhoupt condition $(\mathbf{A}_2)$ with the same Muckenhoupt norm.*

Proof. Clearly, $\mathcal{W}_I=AW_IA$ and $(\mathcal{W}^{-1})_I=A^{-1}(W^{-1})_IA^{-1}$. The operator $B=(W_I)^{1/2}A$ is a Hermitian square root of $\mathcal{W}_I$, which means $B^*B=\mathcal{W}_I$. Therefore we can write a polar decomposition for the operator $B=(W_I)^{1/2}A=U(\mathcal{W}_I)^{1/2}$, where U is a unitary matrix. Similarly, for an operator $C=((W^{-1})_I)^{1/2}A^{-1}$ one can write $C=V((\mathcal{W}^{-1})_I)^{1/2}$, where V again is a unitary matrix.

Therefore

$$BC^*=(W_I)^{1/2}((W^{-1})_I)^{1/2}=U(\mathcal{W}_I)^{1/2}((\mathcal{W}^{-1})_I)^{1/2}V^*,$$

and so

$$\|(W_I)^{1/2}((W^{-1})_I)^{1/2}\|=\|(\mathcal{W}_I)^{1/2}((\mathcal{W}^{-1})_I)^{1/2}\|.\quad\square$$

The following two lemmas about scalar (A_2) weights are well known and their proofs are presented only to make the paper self-contained.

For an arc $I\subset\mathbf{T}$ let λ_I be the center point of the top cover of the corresponding "Carleson square", i.e., $1-|\lambda|=|I|$ and $\lambda/|\lambda|$ is the center of I.

Lemma 1.3. *Suppose a scalar weight w satisfies the Muckenhoupt (A_2) condition*

$$(A_2) \qquad \sup_I \{w_I (w^{-1})_I\} =: C < \infty.$$

Then for any arc $I \subset \mathbf{T}$ the inequality $w(\lambda_I) \le K(C) w_I$ holds, where the constant K depends only on the Muckenhoupt norm C.

Proof. Given an arc I let kI, $k>0$ denote the arc with the same center and of length $k|I|$ ($kI = \mathbf{T}$ for $k \ge 1/|I|$). It is easy to see from the form of the Poisson kernel and the Muckenhoupt condition that

$$\begin{aligned} w(\lambda_I) &\le A \sum_{k=0}^{\infty} 2^{-k} w_{2^k I} \le A \sum_{k=0}^{\infty} 2^{-k} \frac{w_{2^k I}}{(w^{-1})_{2^k I}} (w^{-1})_{2^k I} \\ &\le AC \sum_{k=0}^{\infty} 2^{-k} ((w^{-1})_{2^k I})^{-1}. \end{aligned}$$

The Muckenhoupt (A_2) condition implies that the weight w^{-1} is doubling, and therefore for any arc I we have $(w^{-1})_{2I} \ge (2-\varepsilon)^{-1} w_I$, where the constant ε depends only on the Muckenhoupt norm of w (or w^{-1}, which is equivalent). Therefore we can continue our inequality to obtain

$$\begin{aligned} w(\lambda_I) &\le AC \sum_{k=0}^{\infty} 2^{-k} ((w^{-1})_{2^k I})^{-1} \le AC \sum_{k=0}^{\infty} 2^{-k} (2-\varepsilon)^k ((w^{-1})_I)^{-1} \\ &\le C' (w^{-1})_I \le CC' w_I. \quad \square \end{aligned}$$

Corollary 1.4. *If a $d \times d$ matrix weight W satisfies the vector Muckenhoupt condition $(\mathbf{A}_2)$, then*

$$(\mathbf{A}_2') \qquad \sup_{\lambda \in \mathbf{D}} \|W(\lambda)^{1/2} W^{-1}(\lambda)^{1/2}\| < \infty,$$

where the supremum depends on d and the Muckenhoupt norm of W.

We will call the condition $(\mathbf{A}_2')$ *the conformally invariant* (or simply *the invariant*) Muckenhoupt condition.

To prove Corollary 1.4 we need the following result (Lemma 2.1 from [TV1]).

Let k be a scalar-valued function in $L^2 \cap L^\infty$, $\|k\|_2 = 1$. Consider a subspace $\mathcal{K}$ of $L^2(\mathbf{C}^d)$ given by

$$\mathcal{K} \stackrel{\text{def}}{=} k\mathbf{C}^d = \{k\mathbf{e} : \mathbf{e} \in \mathbf{C}^d\},$$

and let P denote the orthogonal projection (in the non-weighted space $L^2 = L^2(\mathbf{C}^d)$) onto $\mathcal{K}$. It is easy to see that P is given by the formula

$$Pf = k \int_{\mathbf{T}} f(t) \overline{k(t)}\, dt.$$

Lemma 1.5. *Given a $d\times d$ matrix weight W, the projection P introduced above extends (from a dense set $L^2\cap L^2(W)$) to a bounded operator on $L^2(W)$ if and only if both weights $W|k|^2$ and $W^{-1}|k|^2$ belong to L^1. In this case the norm of P in $L^2(W)$ is exactly*

$$\left\| \left[\int W|k|^2\right]^{1/2} \left[\int W^{-1}|k|^2\right]^{1/2} \right\|.$$

Proof of Corollary 1.4. It follows from Lemma 1.5 (with $k=|I|^{-1/2}\chi_I$) that the Muckenhoupt condition $(\mathbf{A}_2)$ is equivalent to the uniform boundedness of all averaging operators

$$f \longmapsto \chi_I \frac{1}{|I|}\int_I f, \quad f\in L^2(\mathbf{C}^d).$$

Fix any direction (unit vector) $\mathbf{e}\in\mathbf{C}^d$, and consider the restrictions of these operators on the subspace $\{f\mathbf{e}: f\in L^2\}$. It follows that the Muckenhoupt norm of the scalar weight $w(t)=(W(t)\mathbf{e},\mathbf{e})$ is bounded by the Muckenhoupt norm of W for any $\mathbf{e}\in\mathbf{C}^d$, $\|\mathbf{e}\|=1$.

Lemma 1.3 implies that $W(\lambda_I)\le CW_I$, and similarly for W^{-1}, $W^{-1}(\lambda_I)\le C(W^{-1})_I$. (Here $\le$ means inequality between associated quadratic forms.) By Douglas's lemma (see [Do]) there exist contractions T_1 and T_2 such that

$$W(\lambda_I)^{1/2}=CT_1{W_I}^{1/2}, \quad W^{-1}(\lambda_I)^{1/2}=CT_2{(W^{-1})_I}^{1/2},$$

and therefore

$$\|W(\lambda_I)^{1/2}W^{-1}(\lambda_I)^{1/2}\|\le C^2\|{W_I}^{1/2}{(W^{-1})_I}^{1/2}\|. \quad \square$$

Lemma 1.6. *If a scalar weight w satisfies the Muckenhoupt (A_2) condition, then*

$$w(\lambda)^{1/2}\le Cw^{1/2}(\lambda) \quad \textit{for all } \lambda\in\mathbf{D},$$

where the constant C depends only on Muckenhoupt norm of w.

Proof. Lemma 1.3 implies that the (A_2) condition can be rewritten as

$$\sup_{\lambda\in\mathbf{D}} w(\lambda)w^{-1}(\lambda)<\infty,$$

and the *supremum* depends on the Muckenhoupt norm of w. One can write

$$w(\lambda)w^{-1}(\lambda)=[w(\lambda)/\exp\{(\log w)(\lambda)\}][w^{-1}(\lambda)/\exp\{(\log w^{-1})(\lambda)\}].$$

By the arithmetic-geometric mean inequality, the expressions in brackets are at least 1. Hence the Muckenhoupt condition is equivalent to the following two conditions:

$$\sup_{\lambda\in\mathbf{D}}[w(\lambda)/\exp\{(\log w)(\lambda)\}]<\infty, \quad \sup_{\lambda\in\mathbf{D}}[w^{-1}(\lambda)/\exp\{(\log w^{-1})(\lambda)\}]<\infty.$$

From the first inequality and from Jensen's inequality we have

$$w(\lambda)\le C\exp\{(\log w)(\lambda)\}\le C(w^{1/2}(\lambda))^2,$$

which completes the proof. □

Lemma 1.7. *Let W be a $d\times d$ matrix weight satisfying the vector Muckenhoupt condition $(\mathbf{A}_2)$. Then there is a constant C depending on d and on the Muckenhoupt norm of W such that*

$$\|W(\lambda)\|\le C(\|W\|^{1/2}(\lambda))^2 \quad \textit{for all } \lambda\in\mathbf{D},$$

where $\|W\|^{1/2}(\lambda)$ denotes the harmonic extension of $\|W\|^{1/2}$ at the point λ.

Proof. Fix a unit vector $\mathbf{e}\in\mathbf{C}^d$ ($\|\mathbf{e}\|=1$). As pointed out in the proof of Corollary 1.4 the Muckenhoupt norm of $w=(W\mathbf{e},\mathbf{e})$ is bounded by the Muckenhoupt norm of W. Therefore there exists $C<\infty$ such that

$$(W(\lambda)\mathbf{e},\mathbf{e})\le C[(W\mathbf{e},\mathbf{e})^{1/2}(\lambda)]^2 \quad \text{for all } \lambda\in\mathbf{D}$$

If we take $\mathbf{e}$, $\|\mathbf{e}\|=1$, such that $(W(\lambda)\mathbf{e},\mathbf{e})=\|W(\lambda)\|$ we get the conclusion. □

2. Carleson measures.

Let us recall that a positive Borel measure μ in the unit disc is called a Carleson measure if

$$\mu(Q_I)\le C|I|$$

for any arc I on the circle. Here $Q_I\overset{\text{def}}{=}\{z\in\mathbf{D}:z/|z|\in I,\ 1-|z|<|I|\}$. The best constant C is called the Carleson norm of μ.

For a scalar, vector, or matrix function f on the circle $\mathbf{T}$ let $f(z)$, $z\in\mathbf{D}$, always denote the harmonic extension of f at z.

Lemma 2.1. *Let W be a harmonic function of n variables with values in the space of strictly positive $d\times d$ matrices ($W(x)=W(x)^*>0$ for all x). Then*

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

Proof. The proof is a good exercise in multivariate calculus. We need first to compute $(\partial/\partial x_j)(\log\det W)$. We first notice that $\log\det W=\operatorname{trace}\log W$. This is nice, because differentiation and trace commute. But there is a more important advantage in having trace here: $\operatorname{trace}(AB)=\operatorname{trace}(BA)$. In general, it is difficult to differentiate a function of matrix W, the function $\log W$ being no exception. However, taking into account that for any matrices A, B we have $\operatorname{trace}(AB)=\operatorname{trace}(BA)$, we can immediately see that for any analytic polynomial

$$\frac{\partial}{\partial x_j}(\operatorname{trace}(p(W))) = \operatorname{trace}\left(p'(W)\frac{\partial W}{\partial x_j}\right),$$

and therefore

$$\frac{\partial}{\partial x_j}(\log\det W) = \operatorname{trace}\left(W^{-1}\frac{\partial W}{\partial x_j}\right).$$

So, using again the fact that $(\operatorname{trace} W)'=\operatorname{trace}(W')$ one gets

$$\frac{\partial^2}{\partial {x_j}^2}(\log\det W) = \operatorname{trace}\left(W^{-1}\frac{\partial^2 W}{\partial {x_j}^2}\right) - \operatorname{trace}\left(W^{-1}\frac{\partial W}{\partial x_j}W^{-1}\frac{\partial W}{\partial x_j}\right). \tag{2.1}$$

Here we have also used the fact that the inverse matrix (and the resolvent in general) is easy to differentiate,

$$\frac{\partial}{\partial x_j}W^{-1} = -W^{-1}\frac{\partial W}{\partial x_j}W^{-1}.$$

Since W is a harmonic matrix-valued function,

$$\sum_{j=1}^{n}\operatorname{trace}\left(W^{-1}\frac{\partial^2 W}{\partial {x_j}^2}\right) = 0.$$

To complete the proof it is enough to recall that the similarity transformation $A\mapsto TAT^{-1}$ does not change the trace, so

$$\begin{aligned}\Delta(\log(\det W)) &= -\sum_{j=1}^{n}\operatorname{trace}\left(W^{-1}\frac{\partial W}{\partial x_j}W^{-1}\frac{\partial W}{\partial x_j}\right)\\ &= -\sum_{j=1}^{n}\operatorname{trace}\left(W^{-1/2}\frac{\partial W}{\partial x_j}(W^{-1/2})^2\frac{\partial W}{\partial x_j}W^{-1/2}\right). \qquad \square\end{aligned}$$

Now we would like to introduce the matrix $(\mathbf{A}_{2,\infty})$ condition and the matrix *invariant* $(\mathbf{A}_{2,\infty})$ condition. In the first one we require the existence of a positive constant C such that the following estimate holds for any interval (arc) I on the circle $\mathbf{T}$,

$$\log(\det(W_I)) - (\log\det W)_I \le C. \tag{2.2}$$

Here $(\,\cdot\,)_I$ stands for the averaging over the interval I. The opposite inequality is true for *any* matrix function, as follows from the elementary observation that

$$\det\bigl(\tfrac{1}{2}(A+B)\bigr) \ge (\det(A)\det(B))^{1/2},$$

which amounts in its turn to the arithmetic-geometric mean inequality (see the proof in [TV1]). Another remark is that $(\mathbf{A}_{2,\infty})$ implies (but is not equivalent to) the fact that $(\det(W))^{1/d}$ satisfies the usual scalar (A_∞) condition. (Recall that W is a $d\times d$ matrix). In [NT] and [V] the whole spectrum of conditions $(\mathbf{A}_{p,\infty})$, $p\in(1,\infty)$ was introduced and widely used. In the scalar case all these classes become A_∞, but this is not so in the matrix case. We refer the reader to [NT] and [V] for details.

But when we consider the case $p=2$ it is much more convenient to use the class of *invariant* $(\mathbf{A}_{2,\infty})$ matrix weights. We warn the reader that it does not coincide with $(\mathbf{A}_{2,\infty})$ even in the scalar case $d=1$. The convenience of *invariant* $(\mathbf{A}_{2,\infty})$ becomes manifest in this section and stems from two facts: (1) $(\mathbf{A}_2)$ implies *invariant* $(\mathbf{A}_{2,\infty})$; (2) *invariant* $(\mathbf{A}_{2,\infty})$ is a notion from complex rather than real analysis.

Theorem 2.2. *If a matrix weight W satisfies the* invariant $(\mathbf{A}_{2,\infty})$ *condition then the measures*

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

are Carleson measures.

Proof. We are given that

$$\log(\det(W(z))) - (\log\det W)(z) \le C. \tag{2.3}$$

By Green's formula and Lemma 2.1

$$\log(\det(W(s))) - (\log\det W)(s) = -\frac{1}{2\pi}\iint_{\mathbf{D}} \log\left|\frac{1-\bar{s}z}{z-s}\right| \Delta \log(\det(W(z)))\,dx\,dy$$
$$= \frac{1}{4\pi}\iint_{\mathbf{D}} \left(\operatorname{trace}\left(W(z)^{-1/2}\frac{\partial W(z)}{\partial x}W(z)^{-1/2}\right)^2 + \operatorname{trace}\left(W(z)^{-1/2}\frac{\partial W(z)}{\partial y}W(z)^{-1/2}\right)^2\right) \times \log\left|\frac{1-\bar{s}z}{z-s}\right|^2 dx\,dy.$$

Using the elementary inequality $\log(1/a)\ge 1-a$ for $0<a\le 1$ and the fact that $\|A\|\le \operatorname{trace} A$ for a non-negative matrix A, we can estimate the last integral from below by

$$\frac{1}{4\pi}\iint_{\mathbf{D}} \left\|W(z)^{-1/2}\frac{\partial W(z)}{\partial x}W(z)^{-1/2}\right\|^2 \log\left|\frac{1-\bar{s}z}{z-s}\right|^2 dx\,dy$$
$$\ge \frac{1}{4\pi}\iint_{\mathbf{D}} \left\|W(z)^{-1/2}\frac{\partial W(z)}{\partial x}W(z)^{-1/2}\right\|^2 \left(1-\left|\frac{1-\bar{s}z}{z-s}\right|^2\right) dx\,dy$$
$$= \iint_{\mathbf{D}} \left\|W(z)^{-1/2}\frac{\partial W(z)}{\partial x}W(z)^{-1/2}\right\|^2 \frac{(1-|s|^2)(1-|z|^2)}{|1-\bar{s}z|^2}\,dx\,dy.$$

Together with (2.3) this implies

$$\iint_{\mathbf{D}} \frac{(1-|s|^2)}{|1-\bar{s}z|^2}\left\|W(z)^{-1/2}\frac{\partial W(z)}{\partial x}W(z)^{-1/2}\right\|^2 (1-|z|^2)\,dx\,dy \le C \quad \text{for all } s\in\mathbf{D},$$

which yields that the measure

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

is a Carleson measure.

The measure

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

is treated similarly. □

Now let us show that $(\mathbf{A}_2)$ implies *invariant* $(\mathbf{A}_{2,\infty})$. Let us recall that $(\mathbf{A}_2)$ is equivalent (see Corollary 1.4) to

$$\sup_{z\in\mathbf{D}} \|W(z)^{1/2}(W^{-1})(z)^{1/2}\| < \infty.$$

Lemma 2.3. *Let W be a matrix weight satisfying the* $(\mathbf{A}_2)$ *condition. Then W satisfies the* invariant $(\mathbf{A}_{2,\infty})$.

Proof. The vector Muckenhoupt condition (A_2) implies that $W, W^{-1}\in L^1(\mathbf{T})$, and so $\log(\det W)\in L^1(\mathbf{T})$. Therefore there exists a factorization $W=F^*F$ a.e. on $\mathbf{T}$, 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

$$|\det F(z)| = \exp((\log|\det F|)(z)) = \exp\bigl(\tfrac{1}{2}(\log\det W)(z)\bigr) \tag{2.4}$$

It is well known that $F^*(z)F(z)\le W(z)$ for every $z\in\mathbf{D}$. There are many proofs of this fact, for example it admits a very simple operator-theoretic interpretation. The explanation we present here is more function-theoretic. Direct computation shows that

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

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

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 $W=G^*G$ on $\mathbf{T}$. We should point out to the reader that in general G does not necessarily coincide with F^{-1}. However, applying (2.4) to G one can conclude that

$$|\det G(z)| = \exp\bigl(\tfrac{1}{2}(\log\det W^{-1})(z)\bigr) = |\det F(z)|^{-1}. \tag{2.5}$$

As we noticed already the Muckenhoupt condition can be rewritten as

$$\sup_{z\in\mathbf{D}} \|W(z)^{1/2}(W^{-1})(z)^{1/2}\| < \infty.$$

Therefore,

$$\sup_{z\in\mathbf{D}} |\det(W(z))\det((W^{-1})(z))| < \infty.$$

Using (2.5) one can rewrite the last inequality as

$$\sup_{z\in\mathbf{D}} ([\det W(z)/|\det F(z)|^2][\det W^{-1}(z)/|\det G(z)|^2]) < \infty.$$

Since

$$F(z)^*F(z) \le W(z)$$

and

$$G(z)^*G(z) \le W^{-1}(z)$$

the expressions in brackets are at least 1, so, taking (2.1) into account

$$\sup_{z\in\mathbf{D}}[\det W(z)/\exp((\log\det W)(z))]=C<\infty,$$

or equivalently

$$\log(\det(W(z)))-(\log\det W)(z)\le C, \tag{2.6}$$

where the constant C depends only on the dimension d and the Muckenhoupt norm of W.

3. Equivalence of norms

Lemma 3.1. *Suppose a matrix-valued weight W satisfies the matrix Muckenhoupt (A_2) condition, and let μ be a Carleson measure. Then*

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

for any vector-function $f\in L^2(\mathbf{T},W)$, where the constant C depends on the dimension d, the Muckenhoupt norm of W, and the Carleson norm of μ.

Proof. Consider an operator $\mathcal{J}\colon L^2(\mathbf{T})\to L^2(\mu)$ (both L^2 spaces are vector-valued),

$$(\mathcal{J}f)(z)=W(z)^{1/2}(W^{-1/2}f)(z).$$

To prove the theorem it is enough to show that the operator $\mathcal{J}$ is bounded. We are going to show that its formal adjoint $\mathcal{J}^*$

$$(\mathcal{J}^*f)(\xi)=W^{-1/2}(\xi)\int_{\mathbf{D}}\frac{1-|z|^2}{|1-\bar{\xi}z|^2}W(z)^{1/2}f(z)\,d\mu(z)$$

is bounded. Direct computation shows

$$\begin{aligned}\|\mathcal{J}^*f\|^2&=\int_{\mathbf{D}}\int_{\mathbf{D}}\left[\int_{\mathbf{T}}(W^{-1}(\xi)W(z)^{1/2}f(z),W(s)^{1/2}f(s))\frac{(1-|z|^2)(1-|s|^2)}{|1-\bar{z}\xi|^2|1-\bar{s}\xi|^2}\,dm(\xi)\right]\\&\quad\times d\mu(z)\,d\mu(s)\\&\le\int_{\mathbf{D}}\int_{\mathbf{D}}\left[\int_{\mathbf{T}}\|W(s)^{1/2}W^{-1}(\xi)W(z)^{1/2}\|\frac{(1-|z|^2)(1-|s|^2)}{|1-\bar{z}\xi|^2|1-\bar{s}\xi|^2}\,dm(\xi)\right]\\&\quad\times\|f(z)\|\,\|f(s)\|\,d\mu(z)\,d\mu(s).\end{aligned}$$

It is easy to estimate the inner integral. Namely,

$$|1-\bar{s}z| = |1-\bar{s}\xi+\bar{s}\xi-\bar{s}z| \le |1-\bar{s}\xi|+|s|\,|\xi-z| \le |1-\bar{s}\xi|+|1-\bar{z}\xi|,$$

so

$$|1-\bar{s}\xi|^{-1}|1-\bar{z}\xi|^{-1} \le |1-\bar{s}z|^{-1}(|1-\bar{s}\xi|^{-1}+|1-\bar{z}\xi|^{-1}),$$

and finally

$$|1-\bar{s}\xi|^{-2}|1-\bar{z}\xi|^{-2} \le 2|1-\bar{s}z|^{-2}(|1-\bar{s}\xi|^{-2}+|1-\bar{z}\xi|^{-2}). \tag{3.1}$$

Assume without loss of generality that $|s|\ge|z|$. Then

$$\begin{aligned}
\int_{\mathbf{T}} &\|W(s)^{1/2}W^{-1}(\xi)W(z)^{1/2}\| \frac{(1-|z|^2)(1-|s|^2)}{|1-\bar{z}\xi|^2|1-\bar{s}\xi|^2}\,dm(\xi) \\
&\le \frac{2}{|1-\bar{s}z|^2}\int_{\mathbf{T}} \|W(s)^{1/2}W^{-1}(\xi)W(z)^{1/2}\| \\
&\qquad\times\left(\frac{(1-|z|^2)(1-|s|^2)}{|1-\bar{s}\xi|^2}+\frac{(1-|s|^2)(1-|z|^2)}{|1-\bar{z}\xi|^2}\right)dm(\xi) \\
&\le \frac{2}{|1-\bar{s}z|^2}\biggl((1-|z|^2)\int_{\mathbf{T}} \|W(s)^{1/2}W^{-1}(\xi)W(s)^{1/2}\| \\
&\qquad\times\|W(s)^{-1/2}W(z)^{1/2}\|\,\frac{1-|s|^2}{|1-\bar{s}\xi|^2}\,dm(\xi) \\
&\qquad+(1-|s|^2)\int_{\mathbf{T}} \|W(s)^{1/2}W(z)^{-1/2}\| \\
&\qquad\times\|W(z)^{1/2}W^{-1}(\xi)W(z)^{1/2}\|\,\frac{1-|z|^2}{|1-\bar{z}\xi|^2}\,dm(\xi)\biggr).
\end{aligned}$$

The weight W satisfies the Muckenhoupt condition (A_2), hence

$$\int_{\mathbf{T}} \|W(s)^{1/2}W^{-1}(\xi)W(s)^{1/2}\|\frac{1-|s|^2}{|1-\bar{s}\xi|^2}\,dm(\xi) \le C < \infty$$

and

$$\int_{\mathbf{T}} \|W(z)^{1/2}W^{-1}(\xi)W(z)^{1/2}\|\frac{1-|z|^2}{|1-\bar{z}\xi|^2}\,dm(\xi) \le C < \infty,$$

so

$$\begin{aligned}
\int_{\mathbf{T}} &\|W(s)^{1/2}W^{-1}(\xi)W(z)^{1/2}\|\frac{(1-|z|^2)(1-|s|^2)}{|1-\bar{z}\xi|^2|1-\bar{s}\xi|^2}\,dm(\xi) \\
&\le C\left(\frac{1-|z|^2}{|1-\bar{z}s|^2}\|W(s)^{-1/2}W(z)^{1/2}\|+\frac{1-|s|^2}{|1-\bar{z}s|^2}\|W(s)^{1/2}W(z)^{-1/2}\|\right).
\end{aligned}$$

Therefore, to prove the embedding it is enough to prove that an integral operator with kernel

$$\frac{1-|z|^2}{|1-\bar{z}s|^2}\|W(s)^{-1/2}W(z)^{1/2}\|+\frac{1-|s|^2}{|1-\bar{z}s|^2}\|W(s)^{1/2}W(z)^{-1/2}\|$$

is bounded in $L^2(\mu)$. Since each summand can be obtained from the other by interchanging s and z, it is sufficient to estimate only one, i.e., to prove that an integral operator with kernel, say,

$$K(z,s)=\frac{1-|s|^2}{|1-\bar{z}s|^2}\|W(s)^{1/2}W(z)^{-1/2}\|$$

is bounded.

Now we are going to apply the Senichkin–Vinogradov test to this integral operator. This test is a powerful weapon in proving that a certain kernel gives a bounded operator in $L^2(\mu)$. The reader can find a 5-line proof in Lecture 7 of [N]. Let us also quote the work of Kolmogorov and Seliverstov [KS] pointed out to us by E. Stein, where the trick of doubling the kernel was probably used for the first time.

We need to estimate

$$\int_{\mathbf{D}} K(z,s)K(z,t)\,d\mu(z)$$
$$=\int_{\mathbf{D}}\frac{1-|s|^2}{|1-\bar{z}s|^2}\,\frac{1-|t|^2}{|1-\bar{z}t|^2}\|W(s)^{1/2}W(z)^{-1/2}\|\,\|W(t)^{1/2}W(z)^{-1/2}\|d\mu(z).$$

By (3.1) above

$$|1-\bar{z}s|^{-2}|1-\bar{z}t|^{-2}\le 2|1-\bar{s}t|^{-2}(|1-\bar{z}s|^{-2}+|1-\bar{z}t|^{-2}). \tag{3.2}$$

The product of norms can also be easily estimated:

$$\begin{aligned}
&\|W(s)^{1/2}W(z)^{-1/2}\|\,\|W(t)^{1/2}W(z)^{-1/2}\|\\
&\qquad=\|W(s)^{1/2}W(z)^{-1/2}\|\,\|W(t)^{1/2}W(s)^{-1/2}W(s)^{1/2}W(z)^{-1/2}\|\\
&\qquad\le\|W(s)^{1/2}W(z)^{-1/2}\|\,\|W(t)^{1/2}W(s)^{-1/2}\|\,\|W(s)^{1/2}W(z)^{-1/2}\|\\
&\qquad=\|W(s)^{1/2}W(z)^{-1}W(s)^{1/2}\|\,\|W(t)^{1/2}W(s)^{-1/2}\|\\
&\qquad\le\|W(s)^{1/2}W^{-1}(z)W(s)^{1/2}\|\,\|W(t)^{1/2}W(s)^{-1/2}\|,
\end{aligned}\tag{3.3}$$

and similarly

$$\begin{aligned}
&\|W(s)^{1/2}W(z)^{-1/2}\|\,\|W(t)^{1/2}W(z)^{-1/2}\|\\
&\qquad\le\|W(t)^{1/2}W^{-1}(z)W(t)^{1/2}\|\,\|W(s)^{1/2}W(t)^{-1/2}\|.
\end{aligned}\tag{3.4}$$

Combining (3.2), (3.3), and (3.4), we obtain

$$\int_{\mathbf{D}} K(z,s)K(z,t)\,d\mu(z)$$
$$\leq 2\biggl(\frac{1-|t|^2}{|1-\bar{s}t|^2}\|W(t)^{1/2}W(s)^{-1/2}\|\int_{\mathbf{D}}\frac{1-|s|^2}{|1-\bar{s}z|^2}\|W(s)^{1/2}W^{-1}(z)W(s)^{1/2}\|\,d\mu(z)$$
$$+\frac{1-|s|^2}{|1-\bar{s}t|^2}\|W(s)^{1/2}W(t)^{-1/2}\|\int_{\mathbf{D}}\frac{1-|t|^2}{|1-\bar{z}t|^2}\|W(t)^{1/2}W^{-1}(z)W(t)^{1/2}\|\,d\mu(z)\biggr).$$

If we show that

$$\int_{\mathbf{D}}\frac{1-|s|^2}{|1-\bar{s}z|^2}\|W(s)^{1/2}W^{-1}(z)W(s)^{1/2}\|\,d\mu(z)\leq C<\infty, \tag{3.5}$$

where C depends on d, the Muckenhoupt norm of W, and the Carleson norm of μ, then

$$\int_{\mathbf{D}} K(z,s)K(z,t)\,d\mu(z)\leq C(K(s,t)+K(t,s)),$$

and we are done.

To prove (3.5) we first notice that it is easy to see that for a fixed $s\in\mathbf{D}$ the weight $W(s)^{1/2}W^{-1}(\,\cdot\,)W(s)^{1/2}$ is a Muckenhoupt weight (see, e.g., Lemma 1.2 or Lemma 3.5 of [TV1]), and its Muckenhoupt norm is at most the Muckenhoupt norm of W^{-1} (or, equivalently, of W). Let us prove (3.5) for $s=0$. By Lemma 1.7 (see also Lemma 3.6 of [TV1])

$$\int_{\mathbf{D}}\|W(0)^{1/2}W^{-1}(z)W(0)^{1/2}\|\,d\mu(z)\leq\int_{\mathbf{D}}((\|(W(0)^{1/2}W^{-1}W(0)^{1/2}\|(z))^{1/2})^2\,d\mu(z)$$
$$\leq C\int_{\mathbf{D}}(\|(W(0)^{1/2}W^{-1}W(0)^{1/2}\|^{1/2}(z))^2\,d\mu(z),$$

and since the measure μ is a Carleson measure

$$\int_{\mathbf{D}}\|(W(0)^{1/2}W^{-1}(z)W(0)^{1/2})^{1/2}\|^2\,d\mu(z)$$
$$\leq C\int_{\mathbf{T}}\|(W(0)^{1/2}W^{-1}(\xi)W(0)^{1/2})\|\,dm(\xi)$$
$$=C\int_{\mathbf{T}}\|W(0)^{1/2}W^{-1}(\xi)W(0)^{1/2}\|\,dm(\xi)$$
$$\leq Cd\left\|W(0)^{1/2}\left(\int_{\mathbf{T}}W^{-1}(\xi)\,dm(\xi)\right)W(0)^{1/2}\right\|$$
$$=C'\left\|W(0)^{1/2}W^{-1}(0)W(0)^{1/2}\right\|\leq C''.$$

We use here that $\int \|W\| \le d\| \int W\|$, by Lemma 1.1 (see also Lemma 3.1 of [TV1]), for nonnegative $d\times d$ matrix functions W. The last inequality in the chain is a particular case of the matrix $(\mathbf{A}_2)$ condition.

To prove (3.5) for general s let us introduce a new variable $\tilde z=(z-s)/(1-\bar s z)$ (so $z=(\tilde z+s)/(1+\bar s\tilde z)$) and a new measure $\tilde\mu$,

$$\tilde\mu(\tilde z(\Omega)) = \int_\Omega \left| \left(\frac{z-s}{1-\bar s z}\right)' \right| d\mu(z) = \int_\Omega \frac{1-|s|^2}{|1-\bar s z|^2}\, d\mu(z).$$

Let us also introduce a new weight $\widetilde W$,

$$\widetilde W(\tilde\xi) = W(\xi) = W\left(\frac{\tilde\xi+s}{1+\bar s\tilde\xi}\right), \quad \tilde\xi\in\mathbf{T},$$

i.e., $\widetilde W$ is obtained from W by composing it with a Möbius transformation. Clearly

$$\widetilde W^{-1}(\tilde z) = W^{-1}(z) = W^{-1}\left(\frac{\tilde z+s}{1+\bar s\tilde z}\right).$$

Then the left-hand side in (3.5) can be rewritten as

$$\int_{\mathbf{D}} \|\widetilde W(0)^{1/2}\widetilde W^{-1}(\tilde z)\widetilde W(0)^{1/2}\|\, d\tilde\mu(\tilde z).$$

Since $\widetilde W$ is a Muckenhoupt weight and $\tilde\mu$ is a Carleson measure (here the conformal invariance of Carleson measures is used, see, e.g., [G, p. 239]), we have reduced the inquality (3.5) to the case $s=0$, which we have already proved. □

Theorem 3.2. *Let W be a matrix Muckenhoupt (A_2) weight, and let f be a vector-function in $L^2(W)$ such that $f(0)=0$. Then there exists a constant C such that*

$$\frac{1}{C}\int_{\mathbf{T}} (Wf,f)\,dm \le \iint_{\mathbf{D}} [(W(z)\partial f(z),\partial f(z)) + (W(z)\bar\partial f(z),\bar\partial f(z))](1-|z|^2)\,dx\,dy$$
$$\le C\int_{\mathbf{T}} (Wf,f)\,dm.$$

Proof. Applying Green's formula we get

$$\int_{\mathbf{T}} (Wf,f)\,dm = \frac{1}{2\pi}\iint_{\mathbf{D}} \Delta[(W(z)f(z),f(z))]\log\frac{1}{|z|}\,dx\,dy.$$

Since for $|z|\le r<1$ we have $|\Delta f(z)|\le C(r)\int_T |f|\,dm$, the logarithmic singularity $\log(1/|z|)$ above does not matter and

$$\iint_{\mathbf{D}} \Delta[(W(z)f(z),f(z)))]\log\frac{1}{|z|}\,dx\,dy \asymp \iint_{\mathbf{D}} \Delta[(W(z)f(z),f(z))](1-|z|^2)\,dx\,dy,$$

where $\asymp$ means equivalence in a sense of two-sided estimate.

Using the formula $\Delta=4\partial\bar\partial=4\bar\partial\partial$ we obtain

$$\begin{aligned}\tfrac14\Delta[(Wf,f))] &= \partial\bar\partial[(Wf,f)]\\ &= (\partial W\bar\partial f,f)+(\partial Wf,\partial f)+(\bar\partial W\partial f,f)\\ &\quad+(\bar\partial Wf,\bar\partial f)+(W\partial f,\partial f)+(W\bar\partial f,\bar\partial f)\\ &= 2\operatorname{Re}(\partial Wf,\partial f)+2\operatorname{Re}(\bar\partial Wf,\bar\partial f)+(W\partial f,\partial f)+(W\bar\partial f,\bar\partial f).\end{aligned}$$

Mutiplying by $1-|z|^2$ and integrating over $\mathbf{D}$ we get

$$\begin{aligned}A^2 &\overset{\text{def}}{=} \iint_{\mathbf{D}} \Delta[(W(z)f(z),f(z))](1-|z|^2)\,dx\,dy\\ &= 8\iint_{\mathbf{D}} [\operatorname{Re}(\partial Wf,\partial f)+\operatorname{Re}(\bar\partial Wf,\bar\partial f)](1-|z|^2)\,dx\,dy\\ &\quad+4\iint_{\mathbf{D}} [(W\partial f,\partial f)+(W\bar\partial f,\bar\partial f)](1-|z|^2)\,dx\,dy.\end{aligned}\tag{3.6}$$

Let us denote the second term on the right side by $4B^2$. Note that B^2 is exactly the integral from the statement of the theorem. We already know that $\int_{\mathbf{T}}(Wf,f)\,dm\asymp A^2$, so to prove the theorem we need to show that $A\asymp B$.

Let us estimate

$$\begin{aligned}&\left|\iint_{\mathbf{D}} \operatorname{Re}(\partial Wf,\partial f)(1-|z|^2)\,dx\,dy\right|\\ &\quad=\left|\iint_{\mathbf{D}} \operatorname{Re}(W(z)^{-1/2}\partial W(z)W(z)^{-1/2}W(z)^{1/2}f,W(z)^{1/2}\partial f)(1-|z|^2)\,dx\,dy\right|\\ &\quad\le\iint_{\mathbf{D}} \|W(z)^{-1/2}\partial W(z)W(z)^{-1/2}\|\,\|W(z)^{1/2}f\|\,\|W(z)^{1/2}\partial f\|(1-|z|^2)\,dx\,dy\\ &\quad\le\left(\iint_{\mathbf{D}} \|W(z)^{-1/2}\partial W(z)W(z)^{-1/2}\|^2(W(z)f(z),f(z))(1-|z|^2)\,dx\,dy\right)^{1/2}\\ &\qquad\times\left(\iint_{\mathbf{D}} (W(z)\partial f(z),\partial f(z))(1-|z|^2)\,dx\,dy\right)^{1/2}.\end{aligned}$$

The second integral in the product is estimated above by B^2.

The first one is also easy to estimate. By Theorem 2.2 the measure

$$\|W(z)^{-1/2}\partial W(z)W(z)^{-1/2}\|(1-|z|^2)\,dx\,dy$$

is Carleson, so Lemma 3.1 implies

$$\iint_{\mathbf{D}} \|W(z)^{-1/2}\partial W(z)W(z)^{-1/2}\|^2(W(z)f(z),f(z))(1-|z|^2)\,dx\,dy \\ \leq C\int_{\mathbf{T}}(Wf,f)\,dm \leq C'A^2.$$

Therefore

$$\left|\iint_{\mathbf{D}} \mathrm{Re}(\partial Wf,\partial f)(1-|z|^2)\,dx\,dy\right| \leq CAB,$$

and similarly

$$\left|\iint_{\mathbf{D}} \mathrm{Re}(\bar\partial Wf,\bar\partial f)(1-|z|^2)\,dx\,dy\right| \leq CAB.$$

Summarizing the above, one can write

$$A^2 \leq CAB+4B^2$$

yielding $A\leq C'B$. Rewriting (3.6) as

$$4B^2=A^2-8\iint_{\mathbf{D}}[\mathrm{Re}(\partial Wf,\partial f)+\mathrm{Re}(\bar\partial Wf,\bar\partial f)](1-|z|^2)\,dx\,dy$$

and applying the same trick we obtain $4B^2\leq CAB+A^2$, so $B\leq C''A$. □

Let us make an important remark concerning Theorem 3.2. In [TV3] it is proved that if the weight W satisfies the *invariant* $(\mathbf{A}_{2,\infty})$, then the equivalence of norms stated in Theorem 3.2 holds for analytic f and also for antianalytic f. One does not need the full strength of $(\mathbf{A}_2)$ to do that—a "one sided condition" *invariant* $(\mathbf{A}_{2,\infty})$ suffices. This looks like a contradiction. After all, if we have the equivalence of norms for all analytic polynomials and for all antianalytic polynomials, why not add these relations to get the equivalence of norms for all harmonic polynomials? But there is certainly no contradiction. We cannot add, because to do that with a certain estimate we need that the angle between analytic and antianalytic spaces be positive. But this is not guaranteed by *invariant* $(\mathbf{A}_{2,\infty})$. Only $(\mathbf{A}_2)$ works, as we saw in Theorem 3.2.

On the other hand, we can still add the following inequalities from [TV3]

$$\int_{\mathbf{T}} (W(\xi)f_+(\xi), f_+(\xi))\, dm(\xi) \le C \iint_{\mathbf{D}} (W(z)\partial f_+(z), \partial f_+(z))(1-|z|^2)\, dx\, dy,$$

$$\int_{\mathbf{T}} (W(\xi)f_-(\xi), f_-(\xi))\, dm(\xi) \le C \iint_{\mathbf{D}} (W(z)\bar\partial f_-(z), \bar\partial f_-(z))(1-|z|^2)\, dx\, dy,$$

and to get from *invariant* $(\mathbf{A}_{2,\infty})$ the following one sided estimate involving any function $f=f_+ + f_- \stackrel{\text{def}}{=} P_+f + P_-f$ and its harmonic extension $f(z)$

$$\int_{\mathbf{T}} (W(\xi)f(\xi), f(\xi))\, dm(\xi) \le C \iint_{\mathbf{D}} (W(z)\nabla f(z), \nabla f(z))(1-|z|^2)\, dx\, dy.$$

You see that the angle estimate is not needed here. It appears only in the converse estimate.

4. Boundedness of the Hilbert transform

Now it is trivial to prove that if W satisfies the Muckenhoupt $(\mathbf{A}_2)$ condition then the Hilbert transform (or, equivalently, the Riesz projection P_+) is a bounded operator in $L^2(W)$. We present a formal proof for the sake of completeness.

Consider a weighted vector-valued L^2-space $\mathcal{H}=L^2(W(z)(1-|z|^2))$ of functions on $\mathbf{D}$ with values in $\mathbf{C}^d$,

$$\|f\|^2_{\mathcal{H}} = \iint_{\mathbf{D}} (W(z)f(z), f(z))(1-|z|^2)\, dx\, dy.$$

Let $L^2_0(W)=\{f\in L^2(W): f(0)=0\}$. Theorem 3.2 asserts that the mapping

$$f \longmapsto \partial f(z) \oplus \bar\partial f(z)$$

gives a representation of $L^2_0(W)$ as a subspace of $\mathcal{H}\oplus\mathcal{H}$ (with equivalent norm). Here again $f(z)$, $z\in\mathbf{D}$, denotes the harmonic extension of f.

Consider an operator $\mathbf{P}_+$ on $\mathcal{H}\oplus\mathcal{H}$,

$$\mathbf{P}_+(f\oplus g) = f\oplus 0$$

The operator $\mathbf{P}_+$ is obviously bounded, and its restriction to $L^2_0(W)$ is nothing but the Riesz projection P_+. Therefore P_+ is bounded on $L^2_0(W)$. To complete the proof note that for $f\in L^2(W)$

$$P_+f = P_+(f-f(0)) + f(0).$$

The operator

$$f \longmapsto 1\cdot f(0)$$

is a bounded operator in $L^2(W)$; it was shown above in Lemma 1.5 that its norm is exactly $\|W(0)^{1/2}(W^{-1}(0))^{-1/2}\|$. Since $f-f(0)\in L^2_0(W)$, the operator P_+ is bounded on $L^2(W)$.

References

[ACS] AXLER, S., CHANG, S.-Y. A. and SARASON, D., Products of Toeplitz operators, *Integral Equations Operator Theory* **1** (1978), 285–309.

[B1] BLOOM, S., A commutator theorem and weighted BMO, *Trans. Amer. Math. Soc.* **292** (1985), 103–122.

[B2] BLOOM, S., Applications of commutator theory to weighted BMO and matrix analogs of A_2, *Illinois J. Math.* **33** (1989), 464–487.

[D] DAUBECHIES, I., *Ten Lectures on Wavelets,* CBMS-NSF Regional Conf. Ser. in Appl. Math. **61**, Soc. Ind. Appl. Math., Philadelphia, Pa., 1992.

[Do] DOUGLAS, R. G., On majorization, factorization, and range inclusion of operators on Hilbert space, *Proc. Amer. Math. Soc.* **17** (1966), 413–415.

[GCRF] GARCÍA-CUERVA, J. and RUBIO DE FRANCIA, J. L., *Weighted Norm Inequalities and Related Topics*, North-Holland, Amsterdam–New York, 1985.

[G] GARNETT, J. B., *Bounded Analytic Functions*, Academic Press, Orlando, Fla., 1981.

[HS] HELSON, H. and SARASON, D., Past and future, *Math. Scand.* **21** (1967), 5–16.

[HMW] HUNT, R. A., MUCKENHOUPT, B. and WHEEDEN, R. L., Weighted norm inequalities for the conjugate function and the Hilbert transform, *Trans. Amer. Math. Soc.* **176** (1973), 227–251.

[KS] KOLMOGOROV, A. and SELIVERSTOV, G., Sur la convergence de series de Fourier, *C. R. Acad. Sci. Paris* **178** (1925), 303–305.

[NT] NAZAROV, F. and TREIL, S., The hunt for a Bellman function: applications to estimates of singular integral operators and to other classical problems in harmonic analysis, *Algebra i Analiz* **8**:5 (1996), 32–162.

[N] NIKOLSKII, N. K., *Treatise on the Shift Operator*, Springer-Verlag, Berlin–New York, 1986.

[R] ROZANOV, YU. A., *Stationary Stochastic Processes*, Holden-Day, San Francisco, Calif., 1967.

[S1] SARASON, D., Exposed points in H^1. II, in *Topics in Operator Theory*: *Ernst D. Hellinger Memorial Volume* (de Branges, L., Gohberg, I. and Rovnyak, J., eds.), Oper. Theory Adv. Appl. **48**, pp. 333–347, Birkhäuser, Basel, 1990.

[S2] SARASON, D., Products of Toeplitz operators, in *Linear and Complex Analysis Problem Book* 3, *Part* I (Havin, V. P. and Nikolski, N. K., eds.). Lecture Notes in Math. **1573**, pp. 318–319, Springer-Verlag, Berlin–Heidelberg, 1994.

[Si] SIMONENKO, I. B., Riemann's boundary value problem for n pairs of functions with measurable coefficients and its applications to the study of singular integrals in L^p spaces with weights, *Dokl Akad. Nauk SSSR* **141** (1961), 36–39 (Russian). English transl.: *Soviet Math. Dokl.* **2** (1961), 1391–1394.

[St] STEIN, E., *Harmonic Analysis*: *Real-Variable Methods, Orthogonality, and Oscillatory Integrals*, Princeton Univ. Press, Princeton, N. J., 1993.

[T] TREIL, S., Geometric methods in spectral theory of vector valued functions: Some recent results, in *Toeplitz Operators and Spectral Function Theory* (Nikolskii, N. K., ed.), Oper. Theory Adv. Appl. **42**, pp. 209–280, Birkhäuser, Basel, 1989.

[TV1] TREIL, S. and VOLBERG, A., Wavelets and the angle between past and future, *J. Funct. Anal.* **143** (1997), 269–308.

[TV2] TREIL, S. and VOLBERG, A., A simple proof of Hunt–Muckenhoupt–Wheeden theorem, *Preprint,* 1995.

[TV3] TREIL, S. and VOLBERG, A., Completely regular multivariate processes and matrix weighted estimates, *Preprint,* 1996.

[TVZ] TREIL, S., VOLBERG, A. and ZHENG, D., Hilbert transform, Toeplitz operators and Hankel operators, and invariant A_∞ weights, to appear in *Rev. Mat. Iberoamericana.*

[V] VOLBERG, A., Matrix A_p weights via S-functions, to appear in *J. Amer. Math. Soc.*

Received April 23, 1996
in revised form December 4, 1996

Sergei Treil
Department of Mathematics
Michigan State University
East Lansing, MI 48824
U.S.A.
email: treil@math.msu.edu

Alexander Volberg
Department of Mathematics
Michigan State University
East Lansing, MI 48824
U.S.A.
email: volberg@math.msu.edu