MATH. SCAND. 86 (2000), 75–99

SPECTRAL REGULARIZATION INEQUALITIES

MIKHAIL LIFSHITS and MICHEL WEBER

Abstract

We develop a recent idea of spectral regularization introduced by M. Talagrand in the study of covering numbers of averages of contractions in a Hilbert space. We show that this idea can be concentrated in one inequality, which turns out to be a suitable tool for the study of other characteristics of the set of averages and yields many useful corollaries. In particular, we recover some recent results for the Littlewood-Paley square functions in ergodic theory due to R. Jones, I. Ostrovskii and J. Rosenblatt. We also easily deduce original Talagrand's estimate of covering numbers and provide better estimates for geometric subsequences of the averages. Similar results for the bilateral Hilbert transform as well as for some non Hilbertian case are also obtained. Developing more the idea of spectral regularization towards oscillations of averages, we obtain a second inequality which allows us to recover some recent results of R. Jones *et al*, related to oscillations functions in ergodic theory. In the last section, we prove a new criterion of the a.s. convergence of random sequences under suitable incremental conditions. Combining this criterion with our inequalities, we obtain as a corollary the classical theorem of Rademacher-Menshov on orthogonal series and the famous spectral criterion for the strong law of large numbers due to V. F. Gaposhkin.

1.1. *Main inequality and its consequences.*

Let $U: \ H \to H$ be a contraction in a Hilbert space $(H, \|\cdot\|)$. Put for any $f \in H$ and $n \geq 1$,

$$A_n(f) = \frac{1}{n}\sum_{j=0}^{n-1} U^j(f), \qquad A(f) = \{A_n(f), n \geq 1\}.$$

The spectral lemma ([8], p.94) reduces the study of sequences of polynomial operators of U like A_n in many problems from ergodic theory, probability theory or harmonic analysis to the study of some suitable set of functions in a space $L^2([-\pi, \pi), \mu)$. In our case, the characteristics of the set $A(f)$ are related to those of the set of functions

$$V = \left\{ V_n = \frac{e^{in\theta} - 1}{n(e^{i\theta} - 1)} \right\},$$

First author supported by International Science Foundation and Russian Foundation for Basic Research.

Received June 18, 1997

since we know by the spectral lemma that there exists a measure μ on $[-\pi, \pi)$ such that $\mu[-\pi, \pi) = \|f\|^2$ and

$$\|A_m(f) - A_n(f)\| \leq \|V_m - V_n\|_{2,\mu}, \tag{1}$$

with equality when U is an isometry.

Recently, M. Talagrand ([14], cf. also [15]) introduced in the study of the metric entropy of the set $A(f)$, an idea of spectral regularization. The main purpose of our paper is to develop this idea and to show that it can be concentrated in one inequality which turns out to be a suitable tool for the study of various characteristics of the set of averages.

We introduce a new measure $\hat{\mu}$, a regularization of μ by means of an appropriate kernel Q. Namely let the correspondent Lebesgue density be

$$\frac{d\hat{\mu}}{dx}(x) = \int_{-\pi}^{\pi} Q(\theta, x)\mu(d\theta) = \int_{|\theta|<|x|} |x|^{-3}\theta^2\mu(d\theta) + \int_{|x|<|\theta|\leq\pi} |\theta|^{-1}\mu(d\theta), \quad 0 < |x| \leq \pi.$$

Our main inequality is stated in the next theorem.

THEOREM 1. *Let $m \geq n$ be two positive integers. Then*

$$\|V_m - V_n\|_{2,\mu}^2 \leq 4\pi\ \hat{\mu}\left(\frac{1}{m}, \frac{1}{n}\right]. \tag{2}$$

The following corollaries show the remarkable efficiency of that elementary inequality.

COROLLARY 2 (Estimation of the Littlewood-Paley square function [6]). *Let n_p be an increasing sequence of positive integers. Then*

$$\sum_{p=1}^{\infty} \|A_{n_{p+1}}(f) - A_{n_p}(f)\|^2 \leq 6\pi\|f\|^2.$$

PROOF. Inequality (2) of Theorem 1 implies

$$\sum_{p=1}^{\infty} \|V_{n_{p+1}} - V_{n_p}\|_{2,\mu}^2 \leq 4\pi \sum_{p=1}^{\infty} \hat{\mu}\left(\frac{1}{n_{p+1}}, \frac{1}{n_p}\right] \leq 4\pi\hat{\mu}(0, \pi) = 2\pi\hat{\mu}[-\pi, \pi).$$

Besides,

$$\hat{\mu}[-\pi, \pi) = \int\int Q(\theta, x)dx\mu(d\theta) =$$

$$\int\left(\int_{|\theta|<|x|} |x|^{-3}dx\ \theta^2 + \int_{|x|<|\theta|} dx|\theta|^{-1}\right)\mu(d\theta) \leq 3\,\mu[-\pi, \pi) = 3\|f\|^2.$$

REMARK. By applying the above inequalities to the measure $\mu = \delta_\theta$, we also get

$$\sum_{p=1}^{\infty} \left| V_{n_{p+1}}(\theta) - V_{n_p}(\theta) \right|^2 \leq 6\pi.$$

That inequality was proved in [6], theorem 1.1, p.268 (see also the proof of theorem 1.2, p.269) with the constant 25^2.

Theorem 1 also allows to recover in a very simple way an entropy estimate due to M. Talagrand. Recall that the entropy number $N(V, \varepsilon)$ of a set V in a metric space (L, d) is defined for $\varepsilon > 0$ as a minimal possible number of sets $C_1, \ldots C_n$ such that $V \subset \cup_{j \leq n} C_j$ and $d - \operatorname{diam}(C_j) \leq \varepsilon$ for all j.

COROLLARY 3 (Estimation of the entropy numbers, [14]). *Let $N(V, \varepsilon)$ be the entropy number of order ε of the set V in $L^2([-\pi, \pi), \mu)$. Then,*

$$N(V, \varepsilon) \leq K\varepsilon^{-2} + 1$$

with $K = 6\pi\mu[-\pi, \pi)$.

PROOF. Let $h = (4\pi)^{-1}, \varepsilon > 0$ and $J \geq K\varepsilon^{-2} - 1$. Put

$$V^{(j)} = \{V_n : \; jh\varepsilon^2 \leq \hat{\mu}[0, \frac{1}{n}] \leq (j+1)h\varepsilon^2\}, \qquad 0 \leq j \leq J.$$

By Theorem 1, the diameter of the sets $V^{(j)}$ does not exceed $(4\pi h\varepsilon^2)^{1/2} = \varepsilon$. Moreover,

$$\bigcup_{j=0}^{J} V^{(j)} = \{V_n \; : \; 0 < \hat{\mu}[0, \frac{1}{n}] \leq (J+1)h\varepsilon^2\}.$$

We notice that $\hat{\mu}[0, 1] \leq \hat{\mu}[-\pi, \pi)/2 \leq 3\mu[-\pi, \pi)/2$. In order to cover V, it is thus sufficient to have

$$(J+1)h\varepsilon^2 \geq 3\mu[-\pi, \pi)/2$$

or

$$J + 1 \geq 6\pi\mu[-\pi, \pi)\varepsilon^{-2}.$$

This observation achieves the proof.

REMARK. The spectral inequality (1) yields the similar bound for entropy numbers in H:

$$N(A(f), \varepsilon) \leq 6\pi\|f\|^2\varepsilon^{-2} + 1.$$

That estimate is also opimal. This can be seen by considering rotations. Take $X = [-\pi, \pi)$ provided with the normalized Lebesgue measure m. Let also

$\theta \in X$ be irrational and consider the unitary operator U on $L^2(X,m)$ associated with the rotation θ: $\tau x = x + \theta \bmod(2\pi)$, $x \in X$ and defined by $Uf = f \circ \tau$. Let $f \in L^2(X,m)$, $f = \sum_{n\in Z} a_n e_n$ where we denote $e_n(x) = e^{inx}$. Then $\|A_N(f) - A_M(f)\|_2^2 = \sum_{n\in Z} |a_n|^2 |V_N(n\theta) - V_M(n\theta)|^2$.
By virtue of Weyl's criterion, we can build inductively two increasing sequences of positive integers $N_1 < N_2 < \cdots$ and $l_1 < l_2 < \cdots$ such that

$$\forall j = 1, \cdots \ \forall i < j \qquad |V_{N_j}(l_j\theta)| > \frac{1}{2}, \qquad |V_{N_j}(l_i\theta)| < \frac{1}{4}.$$

Let now (r_k) be some increasing sequence of integers, and put $R_k = \sum_{j<k} r_j$. Define $f_k = \frac{1}{\sqrt{r_k}} \sum_{R_k \le s < R_{k+1}} e_{l_s}$ and $f = \sum_{k\ge 1} \frac{c}{k} f_k$ where $c = \sqrt{\frac{6}{\pi^2}}$. Then, the system (f_k) is orthonormal and $\|f\|_{2,\mu} = 1$. Moreover for each $R_k \le i < j < R_{k+1}$

$$\left\|A_{N_j}(f_k) - A_{N_i}(f_k)\right\|_2^2 \ge \frac{1}{r_k} \sum_{s=R_k}^{R_{k+1}-1} \left|V_{N_j}(l_s\theta) - V_{N_i}(l_s\theta)\right|^2$$

$$\ge \frac{1}{r_k} \left|V_{N_j}(l_i\theta) - V_{N_i}(l_i\theta)\right|^2 \ge \frac{1}{16 r_k}.$$

Hence, $\left\|A_{N_j}(f) - A_{N_i}(f)\right\|_2 \ge \frac{c}{k} \left\|A_{N_j}(f_k) - A_{N_i}(f_k)\right\|_2 \ge \frac{c}{k} \frac{1}{4\sqrt{r_k}}$ for any $R_k \le i < j < R_{k+1}$; which proves that

$$N\left(f_k, \frac{1}{4\sqrt{r_k}}\right) \ge r_k, \quad \text{and} \quad N\left(f, \frac{c}{k} \frac{1}{4\sqrt{r_k}}\right) \ge r_k$$

The first inequality shows the optimality of Talagrand's estimate, by taking into account its homogeneity properties. The second inequality shows this: whatever $\varphi : \mathsf{R}^+ \to \mathsf{R}^+$, with $\lim_{x\to 0} \varphi(x) = 0$, there exists $f \in L^2(X,m)$ such that $\limsup_{\varepsilon\to 0} \frac{N(f,\varepsilon)}{\varepsilon^{-2}\varphi(\varepsilon)} > 0$.
Relatively surprisingly, entropy numbers attached to i.i.d. sequences behave more smoothly. To see this, let H be some $L^2(\mu)$, μ a probability measure, and choose U and $f \in L^2(\mu)$ with $<f, 1> = 0$, $\|f\| = 1$, such that f, Uf, $U^2 f, \ldots$ is a sequence of i.i.d. r.v.'s. If we write more simply $A_n = A_n^U(f)$, then $\|A_n - A_m\|^2 = 1/n - 1/m$, for any integers $n < m$. Let $0 < \varepsilon \le 1$ be fixed. Thus $\|A_n\| \le \varepsilon$ if $n > \varepsilon^{-2}$. For each $1 \le n \le \varepsilon^{-1}$, we cover A_n with one ball of radius ε. Finally if $\varepsilon^{-1} \le n \le \varepsilon^{-2}$, let $m_k = [1/((k+1)\varepsilon)]$, $2 \le k \le \varepsilon^{-1}$, where $[x]$ stands for the integer part of x. Let $m_{k-1} \le n \le m_k$. Then, $\|A_n - A_{m_k}\|^2 \le 1/n - 1/m_k \le c_2 \varepsilon^2$ for some absolute constant $c_2 > 0$. Therefore,

$$c_2^{-1} \varepsilon^{-1} \le N((A_n), \varepsilon) \le c_2 \varepsilon^{-1}.$$

These plain computations would also show, when combined with Rosenthal's inequalities, that this estimate continues to hold in L^p, $2 < p < \infty$, with a constant c_p depending on p only. We conclude these remarks by pointing out a Cauchy type uniform estimate of averages A_n, easy to draw from the above estimates and Pisier's theorem ([11], Theorem 2.1), and which is apparently new, at least to us:

$$\mathbf{E} \sup_{N \neq M \geq 1} \frac{|A_N(f) - A_M(f)|}{\Phi\left(\left|\frac{1}{M} - \frac{1}{N}\right|^{\frac{1}{2}}\right)} < \infty$$

whenever $\Phi : \mathsf{R}^+ \to \mathsf{R}^+$ is an increasing map such that $\int_0^1 \frac{du}{\sqrt{u}\Phi(u)} < \infty$.

The next statement shows that for specific subsets of V one can get better estimates than in Corollary 3.

COROLLARY 4 (Estimation of the entropy numbers for a lacunary subsequence, [15].) *Let $\{n_p\}$ be a sequence of positive numbers satisfying for each p, $n_{p+1} \geq 2n_p$. Set $V' = \{V_{n_p},\ p \geq 1\}$. Then, for each $\varepsilon > 0$,*

$$N(V', \varepsilon) \leq \inf_{u \in [0,1]} \left\{ \frac{4\pi\hat{\mu}[0,u]}{\varepsilon^2} + \frac{\ln(1/u)}{\ln 2} + 2 \right\} \leq$$

$$\inf_{v \in [0,1]} \left\{ \frac{6\pi\mu[-v,v] + 4\pi\mu[-\pi,\pi)\, v}{\varepsilon^2} + \frac{2\ln(1/v)}{\ln 2} + 2 \right\}.$$

PROOF. We notice that $n_p \geq 2^{p-1}$. We fix $u \in [0,1]$ and observe that

$$\#\left\{ p :\ \hat{\mu}\left[0, \frac{1}{n_p}\right] > \hat{\mu}[0,u] \right\} \leq$$

$$\#\left\{ p :\ \frac{1}{n_p} \geq u \right\} \leq \frac{\ln(1/u)}{\ln 2} + 1.$$

In order to cover the remaining part (large p), we can use the sets $V^{(j)}$ introduced in the preceding proof. We still choose $h = (4\pi)^{-1}$, but this time J must satisfy

$$(J+1)h\varepsilon^2 \geq \hat{\mu}[0,u].$$

To cover this part, we only need $\frac{4\pi}{\varepsilon^2}\hat{\mu}[0,u] + 1$ sets of diameter bounded by ε; hence, the first estimate follows. The second one follows from the following comparison inequalities for $\hat{\mu}$ and μ.

$$\hat{\mu}[0,u] = \int \left(\int_0^u Q(\theta,x)dx \right) \mu(d\theta) \leq$$

$$\int_{|\theta| < u^{1/2}} \left(\int_0^\pi Q(\theta,x)dx \right) \mu(d\theta) + \int_{|\theta| > u^{1/2}} |\theta|^{-1} u \ \mu(d\theta) \leq$$

$$3\mu[-u^{1/2}, u^{1/2}]/2 + \mu[-\pi,\pi)u^{1/2}.$$

Thus, letting $v = u^{1/2}$ we get

$$\frac{4\pi\hat{\mu}[0,u]}{\varepsilon^2} + \frac{\ln(1/u)}{\ln 2} \le \frac{6\pi\mu[-v,v] + 4\pi\mu[-\pi,\pi)\,v}{\varepsilon^2} + \frac{2\ln(1/v)}{\ln 2}.$$

REMARK. The estimate of Corollary 4 is optimal when we choose $u = u_*(\varepsilon)$ as the solution of equation

$$\frac{\hat{\mu}[0,u]}{\varepsilon^2} = \ln(1/u).$$

Then,

$$N(V',\varepsilon) \le \left(4\pi + \frac{1}{\ln 2}\right)\ln(1/u_*) + 2 = \left(4\pi + \frac{1}{\ln 2}\right)\frac{\hat{\mu}[0,u_*]}{\varepsilon^2} + 2.$$

The size of the entropy numbers thus strongly depends on the behavior near 0 of the measure μ (or $\hat{\mu}$).

We conclude this section by raising the natural question concerning the possibility to extend all above results in L^p-spaces, $p > 2$. A first step in that direction was carried out in [16] where entropy estimates are obtained, but for rotations only. It seems reasonable to hope in getting full extensions by using Berkson-Gillespie's spectral theory.

1.2 *Proof of Theorem 1.*

The proof relies upon the following pair of elementary propositions.

PROPOSITION 5. *Let $a, b \in (0,\pi)$ and $a < b$. Then*

$$\hat{\mu}(a,b] \ge \int \kappa(a,b,\theta)\mu(d\theta)$$

with

$$\kappa(a,b,\theta) = \begin{cases} \frac{b-a}{2a^2b}\theta^2 & |\theta| < a; \\ \frac{b-a}{2b} & |\theta| \in [a,b]; \\ \frac{b-a}{|\theta|} & |\theta| \in (b,\pi]. \end{cases}$$

PROOF. By Fubini theorem,

$$\hat{\mu}(a,b] = \int\left(\int_a^b Q(\theta,x)dx\right)\mu(d\theta).$$

For $|\theta| < a$, we have

$$\int_a^b Q(\theta,x)dx = \int_a^b x^{-3}\theta^2 dx = \frac{a^{-2}-b^{-2}}{2}\theta^2 = \frac{b^2-a^2}{2a^2b^2}\theta^2 \geq \frac{b-a}{2a^2b}\theta^2.$$

For $|\theta| \in [a,b]$, we also have

$$\int_a^b Q(\theta,x)dx = \int_a^{|\theta|} |\theta|^{-1}dx + \int_{|\theta|}^b x^{-3}\theta^2 dx =$$

$$\left(1-\frac{a}{|\theta|}\right) + \left(1-\frac{\theta^2}{b^2}\right)/2 = \frac{3}{2} - \left(\frac{a}{|\theta|} + \frac{\theta^2}{2b^2}\right).$$

The function $\theta \to \frac{a}{|\theta|} + \frac{\theta^2}{2b^2}$ reaches its maximum on the interval $[a,b]$ at point $\theta = a$, where we get the bound

$$\frac{3}{2} - \left(1+\frac{a^2}{2b^2}\right) = \frac{b^2-a^2}{2b^2} \geq \frac{b-a}{2b}.$$

For $|\theta| \in (b,\pi]$, we have

$$\int_a^b Q(\theta,x)dx = \int_a^b |\theta|^{-1}dx = (b-a)|\theta|^{-1}.$$

Apply now Proposition 5 with $a = \frac{1}{m},\ b = \frac{1}{n}$ for integers $m > n$. Then,

$$\kappa_{m,n}(\theta) = \kappa\left(\frac{1}{m},\frac{1}{n},\theta\right) = \begin{cases} \frac{(m-n)m}{2}\theta^2 & |\theta| < \frac{1}{m}; \\ \frac{m-n}{2m} & |\theta| \in [\frac{1}{m},\frac{1}{n}]; \\ \frac{m-n}{mn|\theta|} & |\theta| \in (\frac{1}{n},\pi]. \end{cases}$$

PROPOSITION 6. *Let m,n be positive integers with $m > n$. Then, for each $\theta \in [-\pi,\pi)$ we have*

$$|V_m(\theta) - V_n(\theta)|^2 \leq 4\pi\ \kappa_{m,n}(\theta).$$

PROOF. For $|\theta| < \frac{1}{m}$, we use the following inequality, which is valid for all $\theta \in [-\pi,\pi)$ and integers $m \geq n$

$$|V_m(\theta) - V_n(\theta)| \leq \frac{\pi}{4}(m-n)|\theta|, \tag{3}$$

see, e.g. [14], p.789. Hence

$$|V_m(\theta) - V_n(\theta)|^2 \leq (\pi/4)^2(m-n)^2|\theta|^2 \leq (\pi/4)^2(m-n)m\theta^2 \leq 2\kappa_{m,n}(\theta).$$

For the case $|\theta| \in [\frac{1}{m},\frac{1}{n}]$, we apply the following general estimate

$$|V_m(\theta) - V_n(\theta)|^2 \leq 2|V_m(\theta) - \frac{n}{m}V_n(\theta)|^2 + 2|\left(1-\frac{n}{m}\right)V_n(\theta)|^2 \leq$$

$$(4)\quad 2m^{-2}\Big|\sum_{j=n}^{m-1} e^{ij\theta}\Big|^2 + 2\Big(1-\frac{n}{m}\Big)^2 |V_n(\theta)|^2 \le 2\Big(1-\frac{n}{m}\Big)^2 + 2\Big(1-\frac{n}{m}\Big)^2 = 4\Big(1-\frac{n}{m}\Big)^2.$$

The latter estimate is valid for all $\theta \in [-\pi, \pi)$ and integers $m \ge n$. In the case $|\theta| \in [\frac{1}{m}, \frac{1}{n}]$, we have consequently,

$$|V_m(\theta) - V_n(\theta)|^2 \le 4\Big(1-\frac{n}{m}\Big)^2 \le 4\Big(1-\frac{n}{m}\Big) = 8\kappa_{m,n}(\theta).$$

When $|\theta| \in (\frac{1}{n}, \frac{m\pi}{n(m-n)}]$, the estimate (4) gives

$$|V_m(\theta) - V_n(\theta)|^2 \le 4\Big(1-\frac{n}{m}\Big)^2 \le 4\frac{(m-n)^2}{m^2}\left[\frac{m\pi}{n(m-n)}|\theta|^{-1}\right] =$$

$$\frac{4(m-n)\pi}{mn}|\theta|^{-1} = 4\pi\kappa_{m,n}(\theta).$$

Finally, in the case $|\theta| \in (\frac{m\pi}{n(m-n)}, \pi]$, we use the estimate

$$(5)\qquad |V_n(\theta)| \le \frac{\pi}{n|\theta|}, \qquad \theta \in [-\pi, \pi),\ n \ge 1.$$

We have

$$|V_m(\theta) - V_n(\theta)|^2 \le 2|V_m(\theta)|^2 + 2|V_n(\theta)|^2 \le 2\Big(\frac{\pi}{m|\theta|}\Big)^2 + 2\Big(\frac{\pi}{n|\theta|}\Big)^2 \le$$

$$\frac{4\pi^2}{n^2}|\theta|^{-2} \le \frac{4\pi^2}{n^2}\left[\frac{m\pi}{n(m-n)}\right]^{-1}|\theta|^{-1} = 4\pi\,\frac{m-n}{mn}\,|\theta|^{-1} = 4\pi\kappa_{m,n}(\theta).$$

The proof of Theorem 1 is now immediate, since Propositions 5 and 6 provide

$$\|V_m - V_n\|_{2,\mu}^2 = \int |V_m(\theta) - V_n(\theta)|^2 \mu(d\theta) \le$$

$$4\pi \int \kappa_{m,n}(\theta)\mu(d\theta) \le 4\pi\ \hat{\mu}\Big(\frac{1}{m}, \frac{1}{n}\Big].$$

1.3 *Extensions to Hilbert transform.*

Results of the previous section have extensions to discrete bilateral Hilbert transform

$$M_n(f) = \sum_{0<|j|\le n} U^j(f)/j,$$

where $U : H \to H$ is still a contraction in a Hilbert space $(H, ||.||)$. The properties of M_n were considered in the work of R. Jajte [5].

The associated sequence of spectral kernels is defined as follows

$$W_n(\theta) = \sum_{0<|j|\le n} \exp\{ij\theta\}/j, = 2i \sum_{0<j\le n} \sin(j\theta)/j, \qquad W = \{W_n\}.$$

We also introduce the auxiliary sequence of functions

$$\Sigma_n(\theta) = \sum_{j=n+1}^{\infty} \sin(j\theta)/j.$$

Then, we observe that for all $m \ge n$,

$$|W_m(\theta) - W_n(\theta)| = 2\ |\Sigma_n(\theta) - \Sigma_m(\theta)|.$$

By applying the Abel transform, we get

LEMMA 7. *For all $\theta \in [-\pi, \pi)$ the following inequalities hold.*

a) *For all $n \ge 1$,* $\qquad |\Sigma_n(\theta)| \le 4/(n|\theta|)$ *;*

b) *For all $m \ge n$,* $\qquad |\Sigma_n(\theta) - \Sigma_m(\theta)| \le (m-n)/m$*;*

c) *For all $m \ge n$,* $\qquad |\Sigma_n(\theta) - \Sigma_m(\theta)| \le (m-n)|\theta|$.

One can easily deduce from Lemma 7 the following analogue of Proposition 6. For all integers $m \ge n$ and each $\theta \in [-\pi, \pi)$, inequality

$$|W_m(\theta) - W_n(\theta)|^2 \le 32\ \kappa_{m,n}(\theta)$$

is valid. By combining this inequality with Proposition 5, we get

THEOREM 8. *Let $m \ge n$ be two positive integers. Then,*

$$\| W_m - W_n\|_{2,\mu}^2 \le 32\ \hat{\mu}\left(\frac{1}{m}, \frac{1}{n}\right].$$

This result yields corollaries similar to those of Theorem 1. In particular, for any increasing sequence of positive integers (n_p)

$$\sum_p \|M_{n_{p+1}}(f) - M_{n_p}(f)\|^2 \le 48\ \|f\|^2,$$

and for every $\varepsilon > 0$ the entropy number $N(W, \varepsilon)$ of W in the space $L^2([-\pi, \pi), \mu)$ satisfies

$$N(W, \epsilon) \le 48\ \mu[-\pi, \pi)\varepsilon^{-2}.$$

For the entropy number of the set $\{M_n\} \subset H$ we thus obtain the upper bound $48\|f\|^2\varepsilon^{-2}$.

The extensions to the more difficult case of the *unilateral* Hilbert transform $M'_n = \sum_{0<j\le n} U^j(f)/j$, will be considered elsewhere. More about the properties of M'_n see in [4].

1.4 *An non-Hilbertian case.*

We end this section by indicating an extension to the Wiener space $\mathcal{S}$ of correlated sequences, namely the space consisting with sequences $a = \{a(n),\ n \in Z\}$ such that for any integer k, the limit

$$\gamma_a(k) = \lim_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} a(j)a(j+k)$$

exists. We provide $\mathcal{S}$ with the semi-norm

$$\Delta(a) = \limsup_{n\to\infty} \left(\frac{1}{n}\sum_{j=0}^{n-1} a(j)^2\right)^{\frac{1}{2}}.$$

Entropy numbers associated to any subset E of $(\mathcal{S}, \Delta)$ are noted $N(E, \Delta, .)$. For any $a = (a_j)_{j\in Z} \in l_2(Z)$, let us write $\varphi_a(\alpha) = \sum_{j\in Z} e^{-ij\alpha} a(j)$. Let also T be the right shift on the space of sequences: $T(b_n, n \in Z) = (b_{n+1}, n \in Z)$, and note

$$A_N = \frac{I + T + \cdots + T^{N-1}}{N}, \quad N = 1, 2, \cdots$$

Corollary 9. *For any $a \in \mathcal{S}$, there exists a constant $K(a)$ depending on a only, such that*

$$\forall 0 < \varepsilon \le K(a), \quad N(\{A_n^T(a), n \ge 1\}, \Delta, \varepsilon) \le \frac{K(a)}{\varepsilon^2}.$$

Proof. By Bessel-Parseval's equality,

$$\forall N, M \in \mathsf{N}, \qquad \sum_{n\in Z} \big|\big(A_N - A_M\big)(a)(n)\big|^2 = \frac{1}{2\pi}\int_0^{2\pi} |\varphi_a(\alpha)|^2 |V_N(\alpha) - V_M(\alpha)|^2 \, d\alpha.$$

Hence

$\forall J \ge 1,\ \forall N, M$ such that $N \vee M < J$,

$$\frac{1}{J}\sum_{0\le n<J-N\vee M} \Big|\big(A_N - A_M\big)(a)(n)\Big|^2 \le \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{1}{J}\Big|\sum_{0\le j<J} e^{-ij\alpha} a(j)\Big|^2 |V_N(\alpha) - V_M(\alpha)|^2 \, d\alpha.$$

We can view the right integral as an integration with respect to the measure

$$\Lambda_{J,a}(d\alpha) = \frac{1}{J}\Big|\sum_{0\le j<J} e^{-ij\alpha} a(j)\Big|^2 \, d\alpha$$

Assume now that $a \in \mathcal{S}$. By mean of Herglotz's theorem, there exists a unique nonnegative bounded measure Λ_a on $[-\pi, \pi[$, the *spectral measure* of the sequence a, such that

$$\forall m \in Z, \qquad \gamma_a(m) = \int_{-\pi}^{\pi} e^{im\alpha} \Lambda_a(d\alpha).$$

From Theorem 1 of [1], we know that the family of measures $\Lambda_{J,a}$ weakly converges to Λ_a. We thus deduce

$$\forall N, M \geq 1, \quad \limsup_{J\to\infty} \frac{1}{J} \sum_{0\leq n<J} |(A_N - A_M)(a)(n)|^2 \leq \frac{1}{2\pi} \int_{-\pi}^{\pi} |V_N(\alpha) - V_M(\alpha)|^2 \ \Lambda_a(d\alpha).$$

The result is now easily deduced from Corollary 2.

2. Spectral regularization related to oscillation functions

2.1 *Regularization inequality*

In this section, we show how to modify the spectral regularization in order to control the oscillation functions of ergodic averages. We assume throughout the section that $(\mathcal{X}, \nu)$ is a measure space with a σ-finite measure ν, $H = L^2(\mathcal{X}, \nu)$, and U is the unitary operator generated by a measure preserving transformation of $(\mathcal{X}, \nu)$. We write $\mathrm{Log}(u) = \max\{1, \log u\}$ for $u \geq 1$. We still denote μ the spectral measure of an element $f \in H$ and define the regularized spectral measure $\hat{\mu}$ by letting its Lebesgue density be

$$\frac{d\hat{\mu}}{dx}(x) = \int_{-\pi}^{\pi} Q(\theta, x)\mu(d\theta),$$

where this time

$$Q(\theta, x) = \begin{cases} |\theta|^{-1}\mathrm{Log}^2\left(|\frac{\theta}{x}|\right) & |x| < |\theta|; \\ \theta^2|x|^{-3} & |\theta| \leq |x| \leq \pi. \end{cases}$$

The following theorem provides the control of oscillation over arbitrary block of averages. In the sequel, we denote by K a numerical constant which may vary at each occurrence.

THEOREM 10. *Let n, n_+ be positive integers such that $n \leq n_+$. Then*

$$\left\| \sup_{n\leq m<n_+} |A_m(f) - A_n(f)| \right\|_{2,\nu}^2 \leq K \ \hat{\mu}(\frac{1}{n_+}, \frac{1}{n}].$$

REMARK. The result still holds true for A_m generated by arbitrary contraction of H (not necessarily related to a measure preserving transformation) under supplementary assumption $n_+ \le Rn$. In the latter case the constant K depends on R.

Theorem 10 can be applied to easily recover a recent result due to R. L. Jones, R. Kaufman, J. Rosenblatt, and M. Wierdl concerning oscillation functions of ergodic averages. Indeed, Fubini theorem and an elementary calculation yield

$$\hat{\mu}[0,1] \le \frac{5}{2}\mu[-\pi,\pi) \le \frac{5}{2}||f||^2.$$

and we obtain

Corollary 11. (Estimation of the oscillation function [7], Theorem A). *Let (n_p) be an increasing sequence of positive integers. Then,*

$$\sum_{p=1}^{\infty}\left\|\sup_{n_p\le m<n_{p+1}}|A_m(f)-A_{n_p}(f)|\right\|_{2,\nu}^2 \le K\|f\|_{2,\nu}^2 .$$

Remarks. 1. The statement of corollary still holds true for A_m generated by arbitrary contraction of H under supplementary assumption $\sup n_{p+1}/n_p < \infty$. The relevant constant K depends on this ratio supremum.

2. It is interesting to point out that Corollary 11 contains the Riesz's maximal inequality. Indeed, it suffices to apply it to the sequence $(n_p) = (1, M, M+1, ...)$, then let M tend to infinity. We thus obtain

$$\left\|\sup_m |A_m(f)|\right\|_{2,\nu} = \sup_M \left\|\sup_{m\le M}|A_m(f)|\right\|_{2,\nu} \le K\|f\|_{2,\nu}.$$

The proof of Theorem 10 relies upon the following version of Proposition 5.

Proposition 12. *Let $a, b \in (0,\pi)$ and $a < b$. Then*

$$\hat{\mu}(a,b] \ge \int \kappa(a,b,\theta)\mu(d\theta)$$

with

$$\kappa(a,b,\theta) = \begin{cases} \frac{b-a}{2a^2b}\theta^2 & |\theta| < a; \\ \frac{b-a}{4b} & |\theta| \in [a,b]; \\ \frac{b-a}{|\theta|}\mathrm{Log}^2\left(\frac{|\theta|}{b}\right) & |\theta| \in (b,\pi]. \end{cases}$$

Proof. By Fubini theorem

$$\hat{\mu}[a,b] = \int\left(\int_a^b Q(\theta,x)dx\right)\mu(d\theta).$$

We consider four cases

Case 1: ($|\theta| < a$)

$$\int_a^b Q(\theta,x)dx = \int_a^b \theta^2 x^{-3} dx = \frac{a^{-2}-b^{-2}}{2}\theta^2 = \frac{b^2-a^2}{2a^2b^2}\theta^2 \geq \frac{b-a}{2a^2b}\theta^2.$$

Case 2: ($a \leq |\theta| \leq \frac{a+b}{2}$)

$$\int_a^b Q(\theta,x)dx \geq \int_{|\theta|}^b Q(\theta,x)dx = \int_{|\theta|}^b \theta^2 x^{-3} dx = \frac{1-\left(|\theta|/b\right)^2}{2} =$$

$$\frac{b^2-\theta^2}{2b^2} \geq \frac{b-|\theta|}{2b} \geq \frac{b-a}{4b}.$$

Case 3: ($\frac{a+b}{2} < |\theta| \leq b$)

$$\int_a^b Q(\theta,x)dx \geq \int_a^{|\theta|} Q(\theta,x)dx = \int_a^{|\theta|} |\theta|^{-1}\mathrm{Log}^2\left(\frac{|\theta|}{x}\right)dx =$$

$$\int_{\frac{a}{|\theta|}}^1 \mathrm{Log}^2\left(\frac{1}{u}\right)du \geq \int_{\frac{a}{|\theta|}}^1 1\ du = 1-\frac{a}{|\theta|} \geq 1-\frac{2a}{a+b} \geq \frac{b-a}{2b}.$$

Case 4: ($b < |\theta|$)

$$\int_a^b Q(\theta,x)dx = \int_a^b |\theta|^{-1}\mathrm{Log}^2\left(\frac{|\theta|}{x}\right)dx = \int_{\frac{a}{|\theta|}}^{\frac{b}{|\theta|}} \mathrm{Log}^2\left(\frac{1}{u}\right)du \geq \mathrm{Log}^2\left(\frac{|\theta|}{b}\right)\frac{b-a}{|\theta|}.$$

We will apply this proposition with $a = \frac{1}{m},\ b = \frac{1}{n}$ for integers $m > n$. Then,

$$\kappa_{m,n}(\theta) = \kappa\left(\frac{1}{m},\frac{1}{n},\theta\right) = \begin{cases} \frac{(m-n)m}{2}\theta^2 & |\theta| < \frac{1}{m}; \\ \frac{m-n}{4m} & |\theta| \in [\frac{1}{m},\frac{1}{n}]; \\ \frac{m-n}{mn|\theta|}\mathrm{Log}^2(n|\theta|) & |\theta| \in (\frac{1}{n},\pi]. \end{cases}$$

2.2 *Proof of Theorem 10.*

At first we prove the theorem for *short dyadic* block. Namely, let us assume additionally that for some integer p

$$n_+ - n = 2^p \leq 2n. \tag{6}$$

We use the classical dyadic scheme and thus introduce the following binary increments

$$\Delta_{j,k}(f) = A_{n+(j+1)2^{p-k}}(f) - A_{n+j2^{p-k}}(f), \qquad 1 \leq k \leq p, \quad 0 \leq j < 2^k - 1.$$

Each integer $m \in [n, n+2^p)$ can be written as

$$m = n + \sum_{k=1}^{p} \varepsilon_k(m) 2^{p-k}, \qquad \varepsilon_k(m) = 0 \text{ or } 1.$$

Thus,

$$A_m(f) = A_n(f) + \sum_{k=1}^{p} \varepsilon_k(m) \Delta_{j(k,m),k}(f),$$

where the indexes $\{j(k,m)\}$ are easily defined by $\{\varepsilon_k(m)\}$. Thus, we have

$$\sup_{n \le m < n+2^p} |A_m(f) - A_n(f)| \le \sum_{k=1}^{p} \sup_j |\Delta_{j,k}(f)|.$$

Next, we apply Cauchy-Schwarz inequality,

$$\Big(\sup_{n \le m < n+2^p} |A_m(f) - A_n(f)| \Big)^2 \le \Big(\sum_{k=1}^{p} \sup_j |\Delta_{j,k}(f)| \Big)^2 \le$$

$$\le \sum_{k=1}^{p} \sup_j |\Delta_{j,k}(f)|^2 k^2 \cdot \sum_{k=1}^{p} \frac{1}{k^2} \le \frac{\pi^2}{6} \sum_{k=1}^{p} \sup_j |\Delta_{j,k}(f)|^2 k^2 \le$$

$$\le \frac{\pi^2}{6} \sum_{k=1}^{p} \sum_j |\Delta_{j,k}(f)|^2 k^2.$$

Integration of the latter pointwise inequality produces the following inequality for norms.

$$\Big\| \sup_{n \le m < n+2^p} |A_m(f) - A_n(f)| \Big\|_{2,\nu}^2 \le K \sum_{k=1}^{p} \sum_j \|\Delta_{j,k}(f)\|_{2,\nu}^2 k^2.$$

We now can leave the space $(\mathcal{X}, \nu)$ for the space $L^2([-\pi,\pi),\mu)$, μ being the spectral measure of f, by replacing $\|\Delta_{j,k}(f)\|_{2,\nu}^2$ by $\|\check{\Delta}_{j,k}\|_{2,\mu}^2$, where

$$\check{\Delta}_{j,k} = V_{n+(j+1)2^{p-k}}(\theta) - V_{n+j2^{p-k}}(\theta).$$

It thus remains to study

$$\sum_{k=1}^{p} \sum_j \|\check{\Delta}_{j,k}\|_{2,\mu}^2 k^2 = \int_{-\pi}^{\pi} D_{n,p}(\theta)\ \mu(d\theta),$$

where we put $D_{n,p}(\theta) = \sum_{k=1}^{p} \sum_j |\check{\Delta}_{j,k}(\theta)|^2 k^2$. To this aim, we integrate the basic estimates (3),(4), and (5) in

PROPOSITION 13. *Let m_*, n_* be positive integers such that $m_* > n_*$. Then, for each $\theta \in [-\pi,\pi)$*

$$|V_{m_*}(\theta) - V_{n_*}(\theta)|^2 \leq K\min\left\{(m_* - n_*)^2\theta^2, \frac{(m_* - n_*)^2}{m_*^2}, \frac{1}{(n_*|\theta|)^2}\right\}.$$

We apply this Proposition with $n_* = n + j2^{p-k}$, $m_* = n + (j+1)2^{p-k}$, $m_* - n_* = 2^{p-k}$, in order to get a uniform estimate of $\check{\Delta}_{j,k}(\theta)$ in j:

$$|\check{\Delta}_{j,k}(\theta)|^2 \leq K\min\left\{2^{2(p-k)}\theta^2, 2^{2(p-k)}\frac{1}{n^2}, \frac{1}{n^2|\theta|^2}\right\}.$$

Taking now into account the number of terms in each internal sum, we have

$$D_{n,p}(\theta) \leq K\sum_{k=1}^{p} 2^k k^2 \min\left\{2^{2(p-k)}\theta^2, 2^{2(p-k)}\frac{1}{n^2}, \frac{1}{n^2|\theta|^2}\right\}.$$

Now, we can pass directly to the proof of Theorem 10 in the case (6). According to the previous estimates, it is enough to prove that $D_{n,p}(\theta) \leq K\kappa_{n+2^p,n}(\theta)$. We distinguish four cases.

Case 1: ($|\theta| \in [0, \frac{1}{n+2^p})$). Then,

$$D_{n,p}(\theta) \leq K\sum_{k=1}^{p} 2^k.2^{2(p-k)}\theta^2 k^2 = K2^{2p}\theta^2\left(\sum_{k=1}^{p} 2^{-k}k^2\right) \leq$$

$$K2^p(2^p + n)\theta^2 = K\kappa_{n+2^p,n}(\theta).$$

Case 2: ($|\theta| \in [\frac{1}{n+2^p}, \frac{1}{n}]$). Then,

$$D_{n,p}(\theta) \leq K\sum_{k=1}^{p}\left(2^k.2^{2(p-k)}\frac{1}{n^2}k^2\right) \leq K.\frac{2^{2p}}{n^2}\left(\sum_{k=1}^{\infty} 2^{-k}k^2\right) \leq K\frac{2^{2p}}{n^2}.$$

On the other hand,

$$\frac{2^{2p}}{n^2} = \frac{2^p}{(n+2^p)}\frac{(n+2^p)}{n}.\frac{2^p}{n} \leq K\frac{2^p}{(n+2^p)} = K\kappa_{n+2^p,n}(\theta).$$

Case 3: ($n > 2^p$ and $|\theta| \in (\frac{1}{n}, \frac{1}{2^p}]$) The previous estimate $D_{n,p}(\theta) \leq K\frac{2^{2p}}{n^2}$, is still sufficient, since

$$\kappa_{n+2^p,n}(\theta) = \frac{2^p}{(2^p + n)n|\theta|}\mathrm{Log}^2(n|\theta|) \geq \frac{2^{2p}}{n^2}.\frac{n}{(2^p + n)}.\frac{1}{2^p|\theta|} \geq \frac{2^{2p-1}}{n^2}.$$

Hence, $D_{n,p}(\theta) \leq K\kappa_{n+2^p,n}(\theta)$.

Case 4: ($|\theta| \in (\frac{1}{2^p}, \pi]$). Consider the intervals $(\frac{1}{2^{p-k_o}}, \frac{2}{2^{p-k_o}}]$, $0 \leq k_o \leq p+1$ covering $(\frac{1}{2^p}, \pi]$. Let $|\theta| \in (\frac{1}{2^{p-k_o}}, \frac{2}{2^{p-k_o}}]$; then $2^{k_0} < |\theta|2^p \leq 2^{k_o+1}$. We split our estimate for $D_{n,p}(\theta)$ in two parts, with $k \leq k_o$ and $k > k_o$, respectively. For the first sum, we have the bound

$$\sum_{k\le k_o} 2^k k^2 \frac{1}{n^2\theta^2} \le K\frac{2^p}{n^2|\theta|}k_o^2 \le K\frac{2^p}{n(n+2^p)|\theta|}k_o^2.$$

For the second sum, we have a bound of the same order,

$$\sum_{k=k_o+1}^{p+1} 2^{k+2(p-k)}\frac{k^2}{n^2} \le \frac{2^{2p}}{n^2}\sum_{k=k_o+1}^{p+1} 2^{-k}k^2 \le K\frac{2^{2p}}{n^2}(k_o+1)^2 2^{-k_o} \le$$

$$\le K\frac{2^{2p}}{n^2}\frac{(k_o+1)^2}{2^p|\theta|} = K\frac{2^p}{n^2|\theta|}(k_o+1)^2.$$

We also have

$$k_o\log 2 \le \log(2^p|\theta|) \le K\log(2n|\theta|) \le \mathrm{Log}(n|\theta|) + K\log 2 \le K\mathrm{Log}(n|\theta|).$$

Consequently, the total estimate is

$$D_{n,p}(\theta) \le K\frac{2^p}{n(n+2^p)|\theta|}\cdot \mathrm{Log}^2(n|\theta|) = K\kappa_{n+2^p,n}(\theta).$$

The proof of Theorem 10 is now finished for the case (6). Next, we consider the short but not necessarily dyadic block assuming only that

$$n_+ - n \le n. \tag{7}$$

Choose p such that $2^{p-1} < n_+ - n \le 2^p$. We can apply the result to the dyadic block $[n, n+2^p]$ for which we have proved that

$$\left\|\sup_{n\le m<n+2^p}|A_m(f)-A_n(f)|\right\|_{2,\nu}^2 \le K\int D_{n,p}(\theta)\mu(d\theta) \le K\int \kappa_{n+2^p,n}(\theta)\mu(d\theta).$$

Moreover, it follows from the definition of the kernel $\kappa_{m,n}$ that inequality $n_- \le m < n_+$ yields

$$\kappa_{n+2^p,n}(\theta) \le K\kappa_{n_+,n}(\theta).$$

By natural monotonity of oscillations and by the previous line and the case that has already been proved, we have

$$\left\|\sup_{n\le m<n_+}|A_m(f)-A_n(f)|\right\|_{2,\nu}^2 \le \left\|\sup_{n\le m<n+2^p}|A_m(f)-A_n(f)|\right\|_{2,\nu}^2 \le$$

$$K\int \kappa_{n_+,n}(\theta)\mu(d\theta) \le K\ \hat\mu\left(\frac{1}{n_+},\frac{1}{n}\right].$$

Now, Theorem 10 is proved under assumption (7), i.e. for short blocks. For the control of oscillations over long blocks, the following result will be useful.

LEMMA 14. *Let $k_1 \le k_2$ and put $B = \{2^k, k = k_1, \cdots, k_2\}$. Then*

$$\left\| \sup_{b \in B} |A_b(f) - A_{b_+}(f)| \right\|_{2,\nu}^2 \le K\hat{\mu}\left(\frac{1}{b_+}, \frac{1}{b_-}\right]$$

where $b_- = \min\{b \in B\}$ and $b_+ = \max\{b \in B\}$.

We postpone the proof of Lemma 14 to the end of the section and now finish the proof of Theorem 10 by considering the remainder case $n_+ > 2n$. The set $B = \{2^k \in [n, n_+]\}$ is not empty. Let $b_- = \min\{b \in B\}$ and $b_+ = \max\{b \in B\}$. For each $m \in [n, b_+)$ we have

$$|A_m - A_{n_+}| \le \inf_{b \in B} |A_m - A_b| + \sup_{b \in B} |A_b - A_{b_+}| + |A_{b_+} - A_{n_+}|.$$

Moreover,

$$\sup_{m \in [n, b_+)} \inf_{b \in B} |A_m - A_b| \le \sup_{m \in [n, b_-)} |A_m - A_{b_-}| + \sup_{b \in B \setminus \{b_+\}} \sup_{m \in [b, 2b)} |A_m - A_b|.$$

Therefore, we have

$$\sup_{m \in [n, n_+)} |A_m - A_{n_+}| \le \sup_{m \in [n, b_-)} |A_m - A_{b_-}| + \sup_{b \in B \setminus \{b_+\}} \sup_{m \in [b, 2b)} |A_m - A_b| +$$

$$\sup_{m \in [b_+, n_+)} |A_m - A_{n_+}| + \sup_{b \in B} |A_b - A_{b_+}|.$$

By the definition of B, each block mentioned in the right hand side is short in the sense of (7). By application of Theorem 10 to those blocks and using Lemma 14, we finally obtain

$$\left\| \sup_{m \in [n, n_+)} |A_m - A_n| \right\|_{2,\nu}^2 \le K\hat{\mu}\left(\frac{1}{b_-}, \frac{1}{n}\right] +$$

$$K \sum_{b \in B \setminus \{b_+\}} \hat{\mu}\left(\frac{1}{2b}, \frac{1}{b}\right] + K\hat{\mu}\left(\frac{1}{n_+}, \frac{1}{b_+}\right] + K\hat{\mu}\left(\frac{1}{b_+}, \frac{1}{b_-}\right] \le K\hat{\mu}\left(\frac{1}{n_+}, \frac{1}{n}\right].$$

Now, the proof of Theorem 10 is achieved completely.

We finish this section with the

PROOF OF LEMMA 14: Let $E(\theta)$ be the resolution of the unit; that is the increasing family $(E(\theta))$ of projectors in $L^2(\mathcal{X}, \nu)$ such that

$$U = \int_{-\pi}^{\pi} e^{i\theta}\, dE(\theta).$$

Then,

$$A_b = A_b^U = \int_{-\pi}^{\pi} V_b(\theta)\, dE(\theta).$$

We approximate A_b with

$$A_b^{'} = \int_{-\frac{1}{b}}^{\frac{1}{b}} dE(\theta) = E\left(\frac{1}{b}\right) - E\left(\frac{-1}{b}\right).$$

This operator appears as the rest of the orthogonal series in the seminal work of V. F. Gaposhkin [3]. Our strategy will take advantage of two facts:
– a maximal inequality for the averages $A_b(f)$,
– an internal commutation property for the operators $A_b^{'}$: $A_{b_1}^{'} A_{b_2}^{'} = A_{max(b_1,b_2)}^{'}$.

Therefore, we go on as follows: we pass from A_b to $A_b^{'}$ to use the commutation, and thus make appear a difference $A_{b_+}^{'} - A_{b_-}^{'}$ instead of oscillation. Next, we return to A_b in order to apply the maximal inequality. Namely,

$$|\,(A_b - A_{b_+})f\,|^2 = \left|\,\left(A_b^{'} - A_{b_+}^{'}\right)f + (A_b - A_b{'}) + \left(A_{b_+}^{'} - A_{b_+}\right)\right|^2 \leq$$

$$\leq 3\left|\,\left(A_b^{'} - A_{b_+}^{'}\right)f\,\right|^2 + 6\sup_{b\in B}\left|\,\left(A_b - A_b^{'}\right)f\,\right|^2.$$

Apply now the commutation property

$$\left(A_b^{'} - A_{b_+}^{'}\right)f = \left(A_b^{'}A_{b_-}^{'} - A_b^{'}A_{b_+}^{'}\right)f = A_b^{'}\left(A_{b_-}^{'} - A_{b_+}^{'}\right)f = A_b f_0 + \left(A_b^{'} - A_b\right)f_0,$$

where we put $f_0 = \left(A_{b_-}^{'} - A_{b_+}^{'}\right)f$. Thus,

$$\left|\,\left(A_b^{'} - A_{b_+}^{'}\right)f\,\right|^2 \leq 2|A_b f_0|^2 + 2\left|\left(A_b^{'} - A_b\right)f_0\right|^2.$$

Consequently,

$$\sup_{b\in B}|\,(A_b - A_{b_+})f\,|^2 \leq 6\sup_{b\in B}|A_b f_0|^2 + 6\sup_{b\in B}\left|\,\left(A_b^{'} - A_b\right)f\,\right|^2 + 6\sup_{b\in B}\left|\left(A_b^{'} - A_b\right)f_0\right|^2.$$

We control the first term by using the dominated ergodic theorem ([8], Theorem 6.3, p. 52):

$$\sup_{b\in B}|A_b f_0|^2 \leq \|f_0\|_{2,\nu}^2 = \mu\{\theta \;:\; |\theta| \in (\frac{1}{b_+}, \frac{1}{b_-}]\}.$$

It was the only point of the proof which does not go through for arbitrary contractions and appeals to the ergodic nature of the operator U.

To control the second and the third terms, we make use of the lacunary property of the set B.

$$\left\|\sup_{b\in B}\left|\left(A_b' - A_b\right)f\right|\right\|_{2,\nu}^2 \leq \sum_{b\in B}\left\|\left(A_b' - A_b\right)f\right\|_{2,\nu}^2 =$$

$$\int \sum_{b\in B}\left|V_b(\theta) - \mathbf{1}_{[-\frac{1}{b},\frac{1}{b}]}\right|^2 \mu(d\theta) = \int \kappa_0(\theta)\mu(d\theta),$$

where we put $\kappa_0(\theta) = \sum_{b\in B}\left|V_b(\theta) - \mathbf{1}_{[-\frac{1}{b},\frac{1}{b}]}\right|^2$. It remains to estimate $\kappa_0(\theta)$. Consider two cases

a) ($b|\theta| < 1$): Then, by (3), $|V_b(\theta) - 1| \leq \frac{\pi}{4}b|\theta|$.

b) ($b|\theta| \geq 1$): Then, by (5), $|V_b(\theta)| \leq \frac{\pi}{b|\theta|}$.

Thus, we have

$$\kappa_0(\theta) \leq \left(\sum_{\substack{b\in B\\ b|\theta|<1}} \frac{\pi^2}{16}b^2\right)\theta^2 + \left(\sum_{\substack{b\in B\\ b|\theta|\geq 1}} \frac{\pi^2}{b^2}\right)|\theta|^{-2}$$

or

$$\kappa_0(\theta) = \begin{cases} \frac{\pi^2}{12}b_+^2\theta^2 & |\theta| < \frac{1}{b_+}; \\ \frac{\pi^2}{12} + \frac{4\pi^2}{3} < 3\pi^2/2 & |\theta| \in [\frac{1}{b_+}, \frac{1}{b_-}]; \\ \frac{4\pi^2}{3b_-^2|\theta|^2} \leq \frac{4\pi^2}{3b_-|\theta|} & |\theta| \in (\frac{1}{b_-}, \pi]. \end{cases}$$

Comparing now $\kappa_0(\theta)$ with the function $\kappa_{b_+,b_-}(\theta)$, and taking into account the fact that $b_+ \geq 2b_-$ (otherwise $b_+ = b_-$ and there is nothing to prove), leads to

$$\kappa_0(\theta) \leq 6\pi^2\kappa_{b_+,b_-}(\theta).$$

Integration over $[-\pi,\pi)$ yields, via Proposition 5,

$$\left\|\sup_{b\in B}\left|\left(A_b' - A_b\right)f\right|\right\|_{2,\nu}^2 \leq 6\pi^2\hat{\mu}\left(\frac{1}{b_+},\frac{1}{b_-}\right].$$

Since the spectral measure of f_0 is bounded by μ, we also have

$$\left\|\sup_{b\in B}\left|\left(A_b' - A_b\right)f_0\right|\right\|_{2,\nu}^2 \leq 6\pi^2\hat{\mu}\left(\frac{1}{b_+},\frac{1}{b_-}\right].$$

Summarizing, what we have proved is

$$\left\|\sup_{b\in B}\left|\left(A_b - A_{b_+}\right)f\right|\right\|_{2,\nu}^2 \leq K\hat{\mu}\left(\frac{1}{b_+},\frac{1}{b_-}\right].$$

This achieves the proof of Lemma 14.

REMARK. Looking at the simple form of the function κ_0, attentive reader can observe that in the statement of Lemma 14 we can replace our regularized spectral measure by the smaller analogous measure from Section 1.1. Therefore, the control of oscillation along lacunary sequence proves to be easier than the same task for full short blocks.

3. Almost sure convergence

In what follows, we assume again that $H = L^2(\mathcal{X}, \nu)$, where ν is a non-negative finite measure. In Section 1 we studied (in the more general setting) contraction operators $U : H \to H$ and the families of averages $A_n(f)$ generated by U and by an element $f \in H$. The general principles of ergodic theory suggest to complete Theorem 1 with investigation of ν-almost sure convergence of the averages $A_n(f)$ on $\mathcal{X}$.

Though inspired by inequality (2) from Theorem 1, the following theorem embraces, however, more general framework. For example, it contains as immediate corollary the celebrated Rademacher-Menshov convergence criterion for orthogonal series.

THEOREM 15. *Let μ be a measure on* $[0,1]$ *such that*

$$\int_0^1 \log^2(1/u)\mu(du) < \infty.$$

Let $B = \{B_n,\ n \in \mathsf{N}\} \subset H$, be a sequence satisfying for all $m \geq n$,

$$\|B_m - B_n\|_{2,\nu}^2 \leq \mu(\frac{1}{m}, \frac{1}{n}].$$

Then, the sequence B converges ν-almost surely on $\mathcal{X}$.

PROOF. First, observe that B is Cauchy sequence in Banach space H. Hence, there exists an H-limit B_∞. Moreover, for each n we have $\|B_\infty - B_n\|_{2,\nu}^2 \leq \mu(0, \frac{1}{n}]$. In particular,

$$\sum_{p=1}^{\infty} \|B_\infty - B_{2^p}\|_{2,\nu}^2 \leq \sum_{p=1}^{\infty} \mu\left(0, \frac{1}{2^p}\right] =$$

$$\sum_{p=1}^{\infty} \int_{2^{-p-1}}^{2^{-p}} p\ \mu(du) \leq \frac{1}{\log 2} \int_0^{2^{-1}} \log(1/u)\mu(du) < \infty.$$

Thus we have $\sum_{p=1}^{\infty} |B_\infty - B_{2^p}|^2 < \infty$ and $B_\infty = \lim_{p\to\infty} B_{2^p}$ ν-almost surely.

Consider now the oscillations over dyadic blocks. Fix $p \geq 1$ and define the set $\mathcal{B}_p = \{B_n, 2^p \leq n < 2^{p+1}\} \subset H$. We have a simple bound for its entropy numbers,

$$N(\mathcal{B}_p,\varepsilon) \le \min\{2^p, \mu_p\varepsilon^{-2}+1\},$$

where $\mu_p = \mu(\frac{1}{2^{p+1}}, \frac{1}{2^p}]$. The first part of this bound is obvious. The second one appears when we cover $\mathcal{B}_p$ by the subsets

$$\mathcal{B}_{p,l} = \left\{ B_n : \mu(0,2^{-p-1}] + l\varepsilon^2 \le \mu\left(0,\frac{1}{n}\right] \le \mu(0,2^{-p-1}] + (l+1)\varepsilon^2, \right\}, \qquad l = 0,1,...$$

(Recall that we used such coverings in the proofs of Corollaries 3 and 4.) Next, we make use of G. Pisier entropy estimate ([11], Theorem 2.1) and thus obtain

$$\left\| \sup_{B_n\in\mathcal{B}_p} |B_n - B_{2^p}| \right\|_{2,\nu} \le K \int_0^{K\mu_p^{1/2}} \sqrt{N(\mathcal{B}_p,\varepsilon)}d\varepsilon \le$$

$$K\left(\int_0^{2^{-p/2}\mu_p^{1/2}} 2^{p/2}d\varepsilon + \int_{2^{-p/2}\mu_p^{1/2}}^{K\mu_p^{1/2}} \mu_p^{1/2}\varepsilon^{-1}d\varepsilon \right) \le$$

$$K\left(\mu_p^{1/2} + \mu_p^{1/2}\left(\log K + \frac{\log 2}{2}p \right)\right) \le K\mu_p^{1/2}(1+p), \tag{8}$$

with (different at each occurrence) numeric constants K.

It is interesting to notice that this key inequality follows not only from Pisier estimate but also by applying of a result of F. Móricz on the sums of dependent variables, which we recall here for convenience of the reader.

THEOREM 16. ([10]) *Let* (X_j) *be a sequence of random variables and* $S(n_1,n_2) = \sum_{n_1<j\le n_2} X_j$. *Assume that there exists an array of reals* $g(n_1,n_2)$ *such that two following conditions are satisfied.*

i) *For all integer* $n_1 < n_2$ *and some real* $r \ge 1$

$$\mathbf{E}|S(n_1,n_2)|^r \le g_{n_1,n_2};$$

ii) *If* $n_1 < n_2 < n_3$, *then*

$$g(n_1,n_3) \le g(n_1,n_2) + g(n_2,n_3).$$

Then, for all $n_1 < n_2$,

$$\mathbf{E} \sup_{n_1<n\le n_2} |S(n_1,n)|^r \le \left(\log 2(n_2-n_1)\right)^r g_{n_1,n_2}.$$

One may apply this theorem with $r=2$, $X_j = B_{2^p+j} - B_{2^p+j-1}$ $(1 \le j \le 2^p)$, $g(n_1,n_2) = \mu\left(\frac{1}{2^p+n_2}, \frac{1}{2^p+n_1}\right]$ and obtain

$$\left\| \sup_{B_n\in\mathcal{B}_p} |B_n - B_{2^p}| \right\|_{2,\nu}^2 = \left\| \sup_{0<n\le 2^p} |S(0,n)| \right\|_{2,\nu}^2 \le \left((p+1)\log 2\right)^2 \mu\left(\frac{1}{2^{p+1}}, \frac{1}{2^p}\right],$$

which is equivalent to (8).

We continue the proof of Theorem 15 by writing

$$\sum_{p=1}^{\infty} \left\| \sup_{B_n \in \mathcal{B}_p} |B_n - B_{2^p}| \right\|_{2,\nu}^2 \leq \sum_{p=1}^{\infty} K\mu_p (1+p)^2 \leq K \int_0^1 \log^2(1/u)\mu(du) < \infty.$$

Finally we have

$$\sum_{p=1}^{\infty} \sup_{B_n \in \mathcal{B}_p} |B_n - B_{2^p}|^2 < \infty$$

and thus

$$\lim_{p\to\infty} \sup_{B_n \in \mathcal{B}_p} |B_n - B_{2^p}| = 0$$

ν-almost surely. Now, the convergence of B_{2^p}, which was proved earlier, yields the convergence of the whole sequence B_n.

Remark. There exists also a completely different proof of Theorem 15 (see [9]) based on Talagrand's technique of majorizing measures [13].

Corollary 17 (Rademacher-Menshov theorem, [2], Ch.4.4). *Let $\{\xi_k\} \subset H$ be an orthogonal sequence satisfying*

$$\sum \|\xi_k\|_{2,\nu}^2 \log^2 k < \infty.$$

Then, the series $\sum \xi_k$ converges ν-almost surely on $\mathcal{X}$.

Proof. It is enough to put $B_n = \sum_1^n \xi_k$ and define the measure

$$\mu = \sum_{k=1}^{\infty} \|\xi_k\|_{2,\nu}^2 \delta_{u_k}, \qquad u_k = \frac{1}{2}\left((k+1)^{-1} + k^{-1}\right).$$

Applying Theorem 15 to B and μ, we obtain the result.

Sometimes the following modification of Theorem 15 may be useful for the investigation of subsequences.

Theorem 18. *Let $J : \mathsf{R}_+ \to \mathsf{R}_+$ be an increasing map such that $J(\mathsf{N}) \subset \mathsf{N}$. Let $B = \{B_j,\ j \in J(\mathsf{N})\} \subset H$, be a sequence satisfying for all integer $m \geq n$,*

$$\|B_{J(m)} - B_{J(n)}\|_{2,\nu}^2 \leq \mu(\frac{1}{J(m)}, \frac{1}{J(n)}],$$

for a measure μ on $[0,1]$ such that

$$\int \left(\log J^{-1}(1/u)\right)^2 \mu(du) < \infty.$$

Then, the sequence B converges ν-almost surely on $\mathcal{X}$.

PROOF. Consider the transformation $S(u) = \frac{1}{J^{-1}(1/u)}$. We have $S(\frac{1}{J(m)}, \frac{1}{J(n)}] = (\frac{1}{m}, \frac{1}{n}]$. Therefore, Theorem 15 works for the sequence $B_n' = B_{J(n)}$ and measure $\mu' = \mu S^{-1}$.

COROLLARY 19. (Spectral criterion of V. F. Gaposhkin for the strong law of large numbers for weakly stationary processes, [3]). *Assume that the spectral measure μ of an element $f \in H$ with respect to unitary operator U satisfies*

$$\int \log^2 \log \frac{1}{|u|} \mu(du) < \infty.$$

Then the sequence of averages $(A_n(f))$ converges ν-almost surely on $\mathcal{X}$.

PROOF. Let $\hat{\mu}$ be the regularization of μ defined in Section 1. It is plain that also

$$\int (\log^2 \log(1/|u|)\hat{\mu}(du) < \infty.$$

Apply Theorem 18 with $J(p) = 2^p$, $B_{2^p} = A_{2^p}(f)$. Theorem 1 provides that the assumption of Theorem 18 is valid. Therefore, we obtain the convergence of $A_{2^p}(f), p \to \infty$. Moreover, Corollary 11 (with subsequent remark) yields that

$$\sum_{p=1}^{\infty} \sup_{2^p \le m < 2^{p+1}} |A_m(f) - A_{2^p}(f)|^2 < \infty$$

ν-almost surely. Hence,

$$\lim_{p \to \infty} \sup_{2^p \le m < 2^{p+1}} |A_m(f) - A_{2^p}(f)| = 0,$$

and Corollary 19 follows.

4. Continuous time

The results of the article remain valid, if we consider a semigroup of unitary operators $\{U_t,\ t \in \mathsf{R}_+\}$ in a Hilbert Space H and the correspondent averages

$$A_T(f) = \frac{1}{T} \int_0^T U_t(f) dt\ .$$

In this case, one must replace the space $L^2([-\pi, \pi), \mu)$ with $L^2(R^1, \mu)$, and the family of kernels V becomes

$$V' = \left\{ V_T'(\theta) = \frac{e^{iT\theta} - 1}{iT\theta} \right\}.$$

Since the basic elementary inequalities

$$|V'_M(\theta) - V'_T(\theta)| \le \min\left\{\frac{M-T}{2}|\theta|; \frac{2(M-T)}{M}\right\}, \qquad M \ge T,$$

and

$$|V'_T(\theta)| \le \frac{2}{T|\theta|}$$

hold true, we still have the analogue of Theorem 1,

$$\|V'_M - V'_T\|^2_{2,\mu} \le 8\hat{\mu}\left(\frac{1}{M}, \frac{1}{T}\right].$$

All corollaries about entropy numbers and oscillations follow straightforward.

Acknowledgement. The authors are much indebted to the referee for extremly careful reading of the manuscript and for his remarks which helped to clarify the exposition.

REFERENCES.

1. J. Coquet, K. Kamae, M. Mendes-France, *Sur la mesure spectrale de certaines suites arithmétiques*, Bull. Soc. Math. France 105 (1977), 369-384.
2. J. Doob, *Stochastic Processes*, Wiley Ed.,1953
3. V.F. Gaposhkin, *Criteria for the strong law of large numbers for some classes of second-order stationary processes and homogeneous random fields*, Theor. Probab. Appl. 22 (1977), 286-310.
4. V.F. Gaposhkin, *Spectral tests for existence of generalized ergodic transforms*, Theor. Probab. Appl. 41 (1996), 251-271 (in Russian edition).
5. R. Jajte, *On the existence of the ergodic Hilbert transform*, Ann. Probab. 15 (1987), 831-835.
6. R. Jones, Ostrovskii I., Rosenblatt J., *Square functions in ergodic theory*, Ergodic Theory Dynamical Systems 16 (1996), 267-305.
7. R. Jones, R. Kaufman, J. Rosenblatt, M. Wierdl, *Oscillation in ergodic theory*, Ergodic Theory Dynamical Systems 18 (1998), 889-935.
8. U. Krengel, *Ergodic Theorems*, de Gruyter, 1985.
9. M. Lifshits, M. Weber, *Régularisation spectrale en théorie ergodique et en probabilités*, C. R. Acad. Sci. Paris Sér. I Math. 324 (1996) , 99-103.
10. F. Móricz, 1976. *Moment inequalities and the strong law of large numbers*, Z. Wahrsch. Verw. Gebiete. 35 (1976), 299-314.
11. G. Pisier, *Conditions d'entropie assurant la continuité de certains processus et applications `a l'analyse harmonique*, Seminaire d'analyse fonctionnelle, Ecole Polytechnique, exposés 13-14 (1980).
12. J. Rosenblatt, M. Wierdl, *Pointwise ergodic theorems via harmonic analysis*, Ergodic Theory and its Connections with Harmonic Analysis, London Math. Soc. 205 (1995), 3-151.
13. M. Talagrand, *Sample boundedness of stochastic processes under increment conditions*, Ann. Probab. 18 (1990), p.1-49.
14. M. Talagrand, *Applying a theorem of Fernique*, Ann. Inst. H. Poincaré 32 (1996), 779-799.
15. M. Weber, Entropie métrique et accroissements des moyennes ergodiques, Ricerche Mat. 47 (1998), 13-27.

16. M. Weber, *An L^p-extension of Talagrand's covering numbers estimates for rotations*, J. Math. Kyoto Univ. 37 (1997), 689-700.
17. M. Weber, *Entropie métrique et convergence presque partout*, monography Coll. „Travaux en Cours'', Hermann, Paris, Vol. 58, 1998, 150 p.

KOMENDANTSKII PROSPECT, 22-2-49
197372, ST-PETERSBURG
RUSSIA
E-MAIL: LIFTS@MAIL.RCOM.RU

MATHÉMATIQUE
UNIVERSITÉ LOUIS-PASTEUR
7, RUE RENÉ DESCARTES
67084 STRASBOURG CEDEX
FRANCE
E-MAIL: WEBER@MATH.U-STRASBG.FR