Clearly Proposition 9 and Corollary 3 are both false in the case $p=1$, $q=\infty$. This may be seen by considering the dual pairing $\langle l^1, m\rangle$ and taking an example as in [B_2]. Notice that Corollary 3 again becomes false if we omit the condition of monotonicity on E. Take e.g. the subspace E of l^p $(p>1)$ consisting of all vectors x having $x_{2n+1}=x_{2n}$ eventually. Then $E^\alpha = l^q$, but E is not barrelled (since the β-dual is not l^q).

References

[B_1] G. Bennett, *Some inclusion theorems for sequence spaces*, Pacific J. Math. 46 (1973), 17–30.

[B_2] —, *Sequence spaces with small β-duals*, Math. Z. 194 (1987), 321–329.

[K] N. J. Kalton, *On summability domains*, Proc. Cambridge Philos. Soc. 73 (1973), 327–338.

[S] W. Stadler, *Zu einer Frage von Wilansky*, Arch. Math. (Basel) 48 (1987), 149–152.

[Sw] J. Swetits, *A characterization of a class of barrelled sequence spaces*, Glasgow Math. J. 19 (1978), 27–31.

[W_1] A. Wilansky, *Modern Methods in Topological Vector Spaces*, McGraw-Hill, New York 1978.

[W_2] —, *Summability through Functional Analysis*, North-Holland, New York 1984.

[ZB] K. Zeller und W. Beekmann, *Theorie der Limitierungsverfahren*, 2. Auflage, Springer, New York 1970.

Dominikus Noll
UNIVERSITÄT STUTTGART
MATHEMATISCHES INSTITUT B
Pfaffenwaldring 57, 7000 Stuttgart 80, F.R.G.

Wolfgang Stadler
UNIVERSITÄT STUTTGART
MATHEMATISCHES INSTITUT A
Pfaffenwaldring 57, 7000 Stuttgart 80, F.R.G.

Received July 7, 1987 (2335)

STUDIA MATHEMATICA, T. XCIV (1989)

Weighted integral inequalities for the ergodic maximal operator and other sublinear operators. Convergence of the averages and the ergodic Hilbert transform

by

DIEGO GALLARDO (Málaga)

Abstract. In this paper we study those pairs of weights (u, w) for which the ergodic maximal operator and other sublinear operators, associated to an automorphism of a σ-finite measure space $(X, \mathcal{M}, \mu)$, are of weak or strong type (ϕ, ϕ) with respect to the measures $w\,d\mu$ and $u\,d\mu$, where ϕ is an N-function. As consequences, we get the μ-almost everywhere convergence and the norm convergence of certain weighted averages (which include the Cesàro averages), as well as of the ergodic Hilbert transform, for every function in the Orlicz space $L_\phi(w\,d\mu)$, where the weight w belongs to a certain class.

1. Introduction and preliminaries. The $A_\phi(T, \mu)$-condition. Let $(X, \mathcal{M}, \mu)$ be a σ-finite measure space and T an invertible μ-measure-preserving transformation from X into itself. Such a transformation will be called an *automorphism* of $(X, \mathcal{M}, \mu)$. Let $\mathfrak{M} \equiv \mathfrak{M}(X, \mathcal{M}, \mu)$ be the space of $\mathcal{M}$-measurable and μ-almost everywhere (μ-a.e.) finite functions from X to $\boldsymbol{R}$ or to $\boldsymbol{C}$.

For each pair of nonnegative integers, m and n, and every $f \in \mathfrak{M}$ we consider the average

$$R_{m,n} f(x) = (m+n+1)^{-1} \sum_{i=-m}^{n} f(T^i x) \tag{1.1}$$

and let M_T be the *ergodic maximal operator*, acting on $\mathfrak{M}$, defined by

$$M_T f = \sup_{m,n \geqslant 0} R_{m,n} |f|. \tag{1.2}$$

The role of the operator M_T in the study of the convergence of the

---

1980 *Mathematics Subject Classification*: Primary 47A35; Secondary 46E30.

Key words and phrases: almost everywhere convergence, automorphism, Besicovitch sequences, Cesàro averages, ergodic Hilbert transform, ergodic maximal Hilbert transform, ergodic maximal operator, extrapolation theorem, N-function, norm convergence, Orlicz spaces, weighted averages, weighted integral inequalities, weights.

This paper contains some of the results of the author's Doctoral Thesis written under the direction of Professor Alberto de la Torre at the University of Málaga.

Cesàro averages $R_{0,n-1}f$ and $R_{n,n}f$ is well known. On the other hand, it is also known that for every $f \in L_p(\mu)$, with $1 \leqslant p < \infty$, the averages $R_{0,n}f$ converge almost everywhere and in norm (see [13]). The convergence of the averages corresponding to the functions $f \in L_p(v)$, $1 \leqslant p < \infty$, where v is another measure absolutely continuous with respect to μ, was studied in [14] by Martín-Reyes; more precisely, he studies the $\mathcal{M}$-measurable functions w: $X \to [0, \infty)$ such that for each f in $L_p(w\,d\mu)$ the *Hopf averages* converge μ-a.e. and in norm.

By a *weight* on X we shall mean an $\mathcal{M}$-measurable function with values in $[0, \infty)$.

A part of our aim, in this paper, is to study the weights w on X for which we have the μ-a.e. convergence and the norm convergence of certain weighted averages (which include the Cesàro averages) of every function $f \in L_\phi(wd\mu)$, where ϕ is an N-function and $L_\phi(wd\mu)$ is the Orlicz space associated to it. This is studied in Section 4.

First, in Section 2, we characterize the pairs of weights (u, w) which satisfy a weighted integral inequality of the form

$$u\{x \in X:\ M_T f(x) > \lambda\} \leqslant \frac{C}{\phi(\lambda)} \int_X \phi(|f|)\, wd\mu, \tag{1.3}$$

with C independent of $f \in \mathfrak{M}$ and $\lambda > 0$, where M_T is the ergodic maximal operator defined by (1.2). (We write $u(E)$ for $\int_E u\,d\mu$ when $E \in \mathcal{M}$.) When (1.3) holds we shall say that M_T is of *weak type* (ϕ, ϕ) *with respect to the measures* $wd\mu$ *and* $ud\mu$ (or simply with respect to (u, w)). Moreover, we shall prove that the weak type (ϕ, ϕ) of M_T is equivalent to the *uniform weak type* (ϕ, ϕ) *of the averages*, i.e.

$$\sup_{m,n \geqslant 0} u\{x \in X:\ |R_{m,n} f(x)| > \lambda\} \leqslant \frac{C}{\phi(\lambda)} \int_X \phi(|f|)\, wd\mu, \tag{1.4}$$

where C is independent of f and λ, and also to the *uniform strong type* (ϕ, ϕ) *of the averages*, i.e.

$$\sup_{m,n \geqslant 0} \int_X \phi(|R_{m,n} f|)\, ud\mu \leqslant C \int_X \phi(|f|)\, wd\mu, \tag{1.5}$$

with C independent of f.

The above results extend partly Theorems 2.7 and 2.26 in [14], given for the L_p-case, $1 < p < \infty$. Finally, in the case $\mu(X) < \infty$, we determine those weights w for which there exists a positive weight u such that M_T is of weak type (ϕ, ϕ) with respect to (u, w).

In Section 3 we study weighted strong type integral inequalities for M_T. More precisely, we characterize the weights w which satisfy

$$\int_X \phi(M_T f)\, wd\mu \leqslant C \int_X \phi(|f|)\, wd\mu, \tag{1.6}$$

with C independent of $f \in \mathfrak{M}$.

In Section 4, as we mentioned at the beginning, we obtain the almost everywhere convergence and the norm convergence of the *weighted averages*

$$n^{-1} \sum_{k=0}^{n-1} b_k f \circ T^k, \qquad (2n+1)^{-1} \sum_{k=-n}^{n} b_k f \circ T^k$$

for every $f \in L_\phi(wd\mu)$, where w belongs to the class characterized by (1.6) and $\{b_k\}$ is a *bounded Besicovitch sequence*. The Besicovitch sequences as weights in the averages were used by Olsen in [17].

In Section 5, we obtain an *extrapolation theorem* in the theory of weights for sublinear operators S defined on $\mathfrak{M}$, which permits us to get weighted integral inequalities of the type (1.6), with S instead of M_T, by obtaining weighted integral inequalities for a particular case of the form $\phi(s) = s^p$, with $1 \leqslant p < \infty$. In this way, some of the results obtained in the theory of weights for the L_p-case can be trivially extended to Orlicz spaces. This occurs, for example, with the *ergodic square function* associated to T, studied in [15].

Finally, we obtain weighted integral inequalities for the *ergodic maximal Hilbert transform* associated to T, and thus we get the existence, in the pointwise and the norm sense, of the *ergodic Hilbert transform*

$$H_T f(x) = \lim_{n \to \infty} \sum_{i=-n}^{n} f(T^i x)/i$$

of each function $f \in L_\phi(wd\mu)$, where w is a weight of the class characterized by (1.6).

Now, we shall present the basic definitions and results concerning N-functions and Orlicz spaces which will be used in this paper.

An N-*function* is a continuous and convex function ϕ: $[0, \infty) \to \boldsymbol{R}$ such that $\phi(s) > 0$, $s > 0$, $s^{-1}\phi(s) \to 0$ as $s \to 0$ and $s^{-1}\phi(s) \to \infty$ as $s \to \infty$.

The function ϕ is an N-function if and only if it has the representation $\phi(s) = \int_0^s \varphi$ where φ: $[0, \infty) \to \boldsymbol{R}$ is continuous from the right, nondecreasing and such that $\varphi(s) > 0$ for $s > 0$, $\varphi(0) = 0$ and $\varphi(s) \to \infty$ as $s \to \infty$. More precisely, φ is the right derivative of ϕ and will be called the *density function* of ϕ.

Associated to φ we have the function ϱ: $[0, \infty) \to \boldsymbol{R}$ defined by $\varrho(t) = \sup\{s:\ \varphi(s) \leqslant t\}$ which has the same properties as φ. We will call ϱ the *generalized inverse* of φ.

We have $\varphi(\varrho(t)) \geqslant t$, $t \geqslant 0$, and $\varphi(\varrho(t) - \varepsilon) \leqslant t$ for any positive reals t and ε such that $\varrho(t) - \varepsilon \geqslant 0$.

The N-function ψ defined by $\psi(t) = \int_0^t \varrho$ is called the *complementary* N-*function* of ϕ. Thus, if $\phi(s) = p^{-1}s^p$, $p > 1$, then $\psi(t) = q^{-1}t^q$ where $pq = p+q$.

Young's inequality asserts that $st \leqslant \phi(s) + \psi(t)$ for $s, t \geqslant 0$, equality holding if and only if $\varphi(s-) \leqslant t \leqslant \varphi(s)$ or else $\varrho(t-) \leqslant s \leqslant \varrho(t)$.

Another inequality we shall use is $s \leqslant \phi^{-1}(s)\psi^{-1}(s) \leqslant 2s$, $s \geqslant 0$.

The so-called Δ_2-condition for an N-function is a very important condition, which plays a fundamental role in many questions. An N-function ϕ is said to satisfy the *Δ_2-condition in* $[0, \infty)$ (or simply the *Δ_2-condition*) if $\sup_{s>0} \phi(2s)/\phi(s) < \infty$. If φ is the density function of ϕ, then ϕ satisfies Δ_2 if and only if there exists a constant $\alpha > 1$ such that $s\varphi(s) < \alpha\phi(s)$, $s > 0$.

The Δ_2-condition for ϕ does not necessarily transfer to the complementary N-function. The latter satisfies the Δ_2-condition if and only if there exists a constant $\beta > 1$ such that $\beta\phi(s) < s\varphi(s)$, $s > 0$. As examples of N-functions which, together with their complementary N-functions, satisfy the Δ_2-condition we have $\phi_1(s) = s^p$, $p > 1$; $\phi_2(s) = s^p(1+\log(1+s))$, $p > 1$; $\phi_3(s) = s^p(1+\log^+ s)$, $p > 1$; $\phi_4(s) = \int_0^s \varrho$ where $\varrho\colon [0, \infty) \to [0, \infty)$ is defined by $\varrho(0) = 0$, $\varrho(t) = 2^{-n}$ if $t \in [2^{-n}, 2^{-n+1})$ and $\varrho(t) = 2^{n-1}$ if $t \in [2^{n-1}, 2^n)$, n a positive integer.

If $(X, \mathcal{M}, \mu)$ is a σ-finite measure space the *Orlicz spaces* $L_\phi \equiv L_\phi(X, \mathcal{M}, \mu)$ and $L_\phi^* \equiv L_\phi^*(X, \mathcal{M}, \mu)$ are defined by

$$L_\phi = \{f \in \mathfrak{M}\colon \int_X \phi(|f|)\,d\mu < \infty\}, \qquad L_\phi^* = \{f \in \mathfrak{M}\colon fg \in L_1 \text{ for all } g \in L_\psi\},$$

where ψ is the complementary N-function of ϕ. We have $L_\phi \subset L_\phi^*$ and if ϕ satisfies Δ_2, then $L_\phi = L_\phi^*$.

L_ϕ^* is a Banach space with the norms

$$\|f\|_\phi = \sup\{\int_X |fg|\,d\mu\colon g \in S_\psi\}, \text{ where } S_\psi = \{g \in L_\psi\colon \int_X \psi(|g|)\,d\mu \leqslant 1\}, \quad \text{and}$$

$$\|f\|_{(\phi)} = \inf\{\lambda > 0\colon \int_X \phi(\lambda^{-1}|f|)\,d\mu \leqslant 1\},$$

which are called the *Orlicz norm* and the *Luxemburg norm* respectively. Both norms are equivalent, actually $\|f\|_{(\phi)} \leqslant \|f\|_\phi \leqslant 2\|f\|_{(\phi)}$.

Hölder's inequality asserts that for every $f \in L_\phi^*$ and every $g \in L_\phi^*$ we have $\|fg\|_1 \leqslant \|f\|_{(\phi)}\|g\|_\psi$, where ϕ and ψ are complementary N-functions.

If $\phi(s) = s^p$ with $p > 1$ then $L_\phi^* = L_\phi = L_p$, $\|f\|_{(\phi)} = \|f\|_p$ and $\|g\|_\psi = \|g\|_q$ where $pq = p+q$.

The convergence $f_n \to f$ in $[L_\phi^*, \|\ \|_\phi]$ implies the mean convergence $\lim_{n\to\infty} \int_X (|f_n - f|)\,d\mu = 0$ but, in general, mean convergence only implies norm convergence when ϕ satisfies Δ_2. Then the set $\mathcal{S}$ of simple functions (with support of finite measure) is dense in $[L_\phi, \|\ \|_\phi]$ if ϕ satisfies Δ_2.

If ϕ and ψ satisfy Δ_2, then $[L_\phi, \|\ \|_{(\phi)}]$ is *reflexive*.

The proofs of most of the above-mentioned results can be found in [12] or in IV-13 of [16].

We shall also use in this paper an *interpolation theorem*; namely, given an N-function ϕ, let α_ϕ and β_ϕ be defined by

$$\alpha_\phi = \lim_{s\to 0^+} -\log h_\phi(s)/\log s = \inf_{0<s<1} -\log h_\phi(s)/\log s,$$

$$\beta_\phi = \lim_{s\to\infty} -\log h_\phi(s)/\log s = \sup_{s>1} -\log h_\phi(s)/\log s,$$

where $h_\phi(s) = \sup_{t>0} \phi^{-1}(t)/\phi^{-1}(st)$. We have $0 \leqslant \beta_\phi \leqslant \alpha_\phi \leqslant 1$; $\beta_\phi > 0$ if ϕ satisfies the Δ_2-condition, and if the complementary N-function of ϕ satisfies Δ_2, then $\alpha_\phi < 1$. We call $q_\phi = \alpha_\phi^{-1}$ and $p_\phi = \beta_\phi^{-1}$ the *lower* and *upper exponents* of ϕ respectively. Then we have

THEOREM 1.7. *Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{F}, \nu)$ be two σ-finite measure spaces, ϕ an N-function satisfying, together with its complementary N-function, the Δ_2-condition and q_ϕ, p_ϕ the lower and upper exponents of ϕ. Let $T\colon L_r + L_s \to \mathfrak{M}(Y)$ be a quasi-additive operator which is simultaneously of weak type (r, r) and (s, s) where $1 \leqslant r < q_\phi$, $p_\phi < s \leqslant \infty$. Then T maps $L_\phi(\mu)$ into $L_\phi(\nu)$ and there exists a constant C such that*

$$\int_Y \phi(|Tf|)\,d\nu \leqslant C\int_X \phi(|f|)\,d\mu \quad \textit{for every } f \in L_\phi(\mu).$$

A direct proof of Theorem 1.7 can be found in [6].

A condition similar to the A_ϕ-condition, given in [6] for the Hardy–Littlewood maximal operator, plays a fundamental role in this paper. We shall call this condition the *$A_\phi(T, \mu)$-condition.*

DEFINITION 1.8. Let φ be the density function of the N-function ϕ, ϱ the generalized inverse of φ, and let u and w be weights on X. We shall say that the pair (u, w) satisfies the *$A_\phi(T, \mu)$-condition*, or that it belongs to the *$A_\phi(T, \mu)$-class*, if there exists a positive constant C such that for every positive integer k, every positive real ε and μ-a.e. x

$$\text{(1.9)} \qquad \Big(k^{-1}\sum_{i=0}^{k-1} \varepsilon u(T^i x)\Big)\varphi\Big(k^{-1}\sum_{i=0}^{k-1} \varrho\big(1/\varepsilon w(T^i x)\big)\Big) \leqslant C.$$

Since $\varrho(s) \to \infty$ and $\varphi(s) \to \infty$ as $s \to \infty$, if for some i with $0 \leqslant i \leqslant k-1$, $w(T^i x) = 0$, we assume that (1.9) holds if $\sum_{i=0}^{k-1} u(T^i x) = 0$, i.e. the pair (u, w) satisfies $A_\phi(T, \mu)$ if (1.9) is satisfied for a.e. x in

$$\text{(1.10)} \qquad Y = X - \bigcup_{i=-\infty}^{\infty} T^i\{x \in X\colon w(x) = 0\}$$

and $\mu(E) = 0$ where

$$\text{(1.11)} \qquad E = \{x \in X\colon w(x) = 0\} - \bigcap_{i=-\infty}^{\infty} T^i\{x \in X\colon u(x) = 0\}.$$

When (w, w) belongs to the $A_\phi(T, \mu)$-class we shall simply say that w *satisfies* $A_\phi(T, \mu)$.

If $\phi(s) = p^{-1} s^p$, $p > 1$, ε does not take part in (1.9) since ϱ is multiplicative and moreover it is the inverse of φ. Thus, in this case, (1.9) reduces to saying

$$(1.12)\qquad \Big(k^{-1}\sum_{i=0}^{k-1} u(T^i x)\Big)\Big(k^{-1}\sum_{i=0}^{k-1} w(T^i x)^{-q/p}\Big)^{p/q} \leqslant C$$

for every positive integer k and μ-a.e. x, where q is the conjugate exponent of p, i.e. $1/p+1/q = 1$. In this way, we obtain the $A_p'(g, T, \mu)$-condition, for $g = 1$, given in [14] (see also [3]). When (1.12) holds we shall say that (u, w) satisfies $A_p(T, \mu)$.

In [2] the $A_1(T, \mu)$-condition, similar to the well-known Muckenhoupt A_1-condition, is defined; namely, the pair (u, w) satisfies $A_1(T, \mu)$ if there exists a constant $C > 0$ such that for μ-a.e. x we have $M_T u(x) \leqslant Cw(x)$. It is easy to see that *if ϕ satisfies the Δ_2-condition then every pair (u, w) of the $A_1(T, \mu)$-class is in the $A_\phi(T, \mu)$-class* (see a similar proof in [6] for the real case).

In what follows, unless otherwise stated, we shall always assume that ϕ, together with its complementary N-function, satisfy the Δ_2-condition, $(X, \mathcal{M}, \mu)$ is a σ-finite measure space, T an automorphism of $(X, \mathcal{M}, \mu)$, and $R_{m,n}$, M_T are defined by (1.1) and (1.2) respectively. Sometimes we shall assume that T is *ergodic*, i.e. if $E \in \mathcal{M}$ is T-invariant then $\mu(E) = 0$ or $\mu(X - E) = 0$. On the other hand, we shall consider two sets "equal" if they agree up to a set of measure zero and we will use the convention that C denotes an absolute positive constant which may change from line to line.

2. Weighted weak type integral inequalities for the ergodic maximal operator and the boundedness of the averages. The characterization of the weights u and w for which the ergodic maximal operator M_T is of weak type (ϕ, ϕ) with respect to the measures $wd\mu$ and $ud\mu$, i.e. M_T satisfies (1.3), can be given via the $A_\phi(T, \mu)$-condition. Moreover, the same condition characterizes the uniform strong type (ϕ, ϕ) of the averages $R_{m,n}$; more precisely:

THEOREM 2.1. *Let u and w be weights on X. The following conditions are equivalent*:

(a) *There exists C such that for every $f \in \mathfrak{M}$ and every $\lambda > 0$*

$$u\{x \in X:\ M_T f(x) > \lambda\} \leqslant \frac{C}{\phi(\lambda)} \int_X \phi(|f|)\, wd\mu.$$

(b) *There exists C such that for every $f \in \mathfrak{M}$ and every $\lambda > 0$*

$$\sup_{k \geqslant 0} u\{x \in X:\ |R_{k,k} f(x)| > \lambda\} \leqslant \frac{C}{\phi(\lambda)} \int_X \phi(|f|)\, wd\mu.$$

(c) *There exists C such that for every $f \in \mathfrak{M}$*

$$\sup_{k \geqslant 0} \int_X \phi(|R_{k,k} f|)\, ud\mu \leqslant C \int_X \phi(|f|)\, wd\mu.$$

(d) *The pair (u, w) satisfies $A_\phi(T, \mu)$.*

Notes. It is easy to see that conditions (b) and (c) in Theorem 2.1 are equivalent respectively to conditions (1.4) and (1.5). Observe that condition (1.5) implies the *uniform boundedness of the averages*; more precisely,

$$\sup_{m,n \geqslant 0} \|R_{m,n} f\|_{(\phi), ud\mu} \leqslant \max(1, C)\, \|f\|_{(\phi), wd\mu},$$

where the norms are those of Luxemburg with respect to the measures $ud\mu$ and $wd\mu$ respectively.

Proof. It will suffice to prove (d)$\Rightarrow$(c), (d)$\Rightarrow$(a) and (b)$\Rightarrow$(d) since (a)$\Rightarrow$(b) and (c)$\Rightarrow$(b) are easy.

Proof of (d)$\Rightarrow$(c). Let $k \in \boldsymbol{Z}^+$, where $\boldsymbol{Z}$ is the set of the integers, $\varepsilon > 0$, $f \in \mathfrak{M}$ and let $x \in Y$, where Y is the set defined by (1.10). We consider the measure space $(\boldsymbol{Z}, \mathcal{F}, \delta_\varepsilon)$ where $\mathcal{F}$ is the σ-algebra whose elements are all the subsets of $\boldsymbol{Z}$ and δ_ε is the measure such that $\delta_\varepsilon(i) = \varepsilon w(T^i x)$ if $i \in \boldsymbol{Z}$.

Let g and h be defined by

$$g(i) = f(T^i x)\chi_{[-2k,2k]}(i), \qquad h(i) = \big(\varepsilon w(T^i x)\big)^{-1}\chi_{[-2k,2k]}(i) \qquad (i \in \boldsymbol{Z}),$$

where $[-2k, 2k]$ denotes the interval in the integers $\{i \in \boldsymbol{Z}:\ -2k \leqslant i \leqslant 2k\}$.

Since $g \in l_\phi(\delta_\varepsilon)$ and $h \in l_\psi(\delta_\varepsilon)$, using Hölder's inequality, we have

$$(2.2)\qquad \sum_{i=-2k}^{2k} |f(T^i x)| = \sum_{i \in \boldsymbol{Z}} |g(i)\, h(i)|\, \delta_\varepsilon(i) \leqslant 2\|g\|_{(\phi),\delta_\varepsilon} \|h\|_{(\psi),\delta_\varepsilon},$$

where ψ is the complementary N-function of ϕ and the norms used are those of Luxemburg with respect to the measure δ_ε.

It follows from the equality cases in Young's inequality and the Δ_2-condition for ψ that there is a constant $\beta > 1$ such that $\psi(s) \leqslant s\varrho(s) \leqslant \beta\psi(s)$, $s \geqslant 0$. Then the $A_\phi(T, \mu)$-condition for (u, w) shows that there is a constant $C > \beta$ such that for all $k \in \boldsymbol{Z}^+$, $\varepsilon > 0$, $\lambda > 0$ and for μ-a.e. x in Y

$$\sum_{i=-2k}^{2k} \psi\Big(\big(\lambda\varepsilon w(T^i x)\big)^{-1}\Big)\varepsilon w(T^i x) \leqslant \lambda^{-1} \sum_{i=-2k}^{2k} \varrho\Big(\big(\lambda\varepsilon w(T^i x)\big)^{-1}\Big)$$

$$\leqslant \lambda^{-1}(4k+1)\,\varrho\Big(C(4k+1)/\sum_{i=-2k}^{2k} \lambda\varepsilon u(T^i x)\Big).$$

Therefore, for $\lambda = C(4k+1)\,\phi^{-1}(1/m(x))$, where $m(x) = \varepsilon\sum_{|i| \leqslant 2k} u(T^i x)$, and taking into account that $s \leqslant \phi^{-1}(s)\psi^{-1}(s)$, $s \geqslant 0$, we have

$$\sum_{i=-2k}^{2k} \psi\left((\lambda\varepsilon w(T^i x))^{-1}\right)\varepsilon w(T^i x) \leqslant \frac{\varrho\left((m(x)\phi^{-1}(1/m(x)))^{-1}\right)}{C\phi^{-1}(1/m(x))}$$
$$\leqslant \beta C^{-1} m(x)\psi\left(\frac{1/m(x)}{\phi^{-1}(1/m(x))}\right) < 1$$

and consequently $\|h\|_{(\psi),\delta_\varepsilon} \leqslant C(4k+1)\phi^{-1}(1/m(x))$.

In this way, it follows from (2.2) that for all $k \in \boldsymbol{Z}^+$, $\varepsilon > 0$, $f \in \mathfrak{M}$ and for μ-a.e. $x \in Y$

$$(4k+1)^{-1}\sum_{i=-2k}^{2k} |f(T^i x)| \leqslant 2C\phi^{-1}\Big(1/\sum_{i=-2k}^{2k} \varepsilon u(T^i x)\Big)\|g\|_{(\phi),\delta_\varepsilon}.$$

Then taking $\varepsilon = \left(\sum_{|i|\leqslant 2k}\phi(|f(T^i x)|)w(T^i x)\right)^{-1}$ we have $\|g\|_{(\phi),\delta_\varepsilon} = 1$ and therefore, since ϕ satisfies Δ_2,

$$\phi\Big((4k+1)^{-1}\sum_{i=-2k}^{2k}|f(T^i x)|\Big)\sum_{i=-2k}^{2k} u(T^i x) \leqslant C\sum_{i=-2k}^{2k}\phi(|f(T^i x)|)w(T^i x),$$

where the constant C is independent of f, k and x. Hence, keeping in mind that T preserves the measure μ,

$$\sum_{i=-k}^{k}\int_X \phi\Big((4k+1)^{-1}\sum_{j=-2k}^{2k}|f\circ T^{j-i}|\Big)u\,d\mu \leqslant C(4k+1)\int_X \phi(|f|)\,w\,d\mu$$

and thus we get

$$\int_X \alpha\phi(\alpha|R_{k,k}f|)\,u\,d\mu \leqslant C\int_X \phi(|f|)\,w\,d\mu$$

with $\alpha = (4k+1)^{-1}(2k+1)$, whence (c) follows easily.

In order to obtain (d) $\Rightarrow$ (a) we shall use transference methods. We shall see that the general result follows from the case in which $X = \boldsymbol{Z}$, μ is the counting measure and T is the shift transformation, i.e. $T(k) = k+1$. On the other hand, we shall prove that the result in the integers case follows from Theorem 2.3 in [6], for the Hardy–Littlewood maximal operator.

For a given function F defined in $\boldsymbol{Z}$ and $k \in \boldsymbol{Z}$ let

$$mF(k) = \sup_{m,n\geqslant 0}(m+n+1)^{-1}\sum_{i=-m}^{n}|F(k+i)|. \tag{2.3}$$

For the maximal operator $\boldsymbol{m}$ we have the following lemma:

LEMMA 2.4. *Let u, w: $\boldsymbol{Z} \to \boldsymbol{R}$ be nonnegative mappings. The following conditions are equivalent:*

(i) *There exists C such that for all $\lambda > 0$ and every mapping F: $\boldsymbol{Z} \to \boldsymbol{R}$*

$$\sum_{\{k\in\boldsymbol{Z}:\, mF(k)>\lambda\}} u(k) \leqslant \frac{C}{\phi(\lambda)}\sum_{k\in\boldsymbol{Z}}\phi(|F(k)|)\,w(k).$$

(ii) *There exists C such that for every pair of positive integers, n and k, and every positive real ε*

$$\Big(n^{-1}\sum_{i=0}^{n-1}\varepsilon u(k+i)\Big)\varphi\Big(n^{-1}\sum_{i=0}^{n-1}\varrho\big(1/\varepsilon w(k+i)\big)\Big) \leqslant C.$$

Proof of Lemma 2.4. In [6] the author proved the following result for the Hardy–Littlewood maximal operator defined by

$$f^*(y) = \sup_{y\in Q}|Q|^{-1}\int_Q |f|\,dy \qquad (f \in L^1_{\mathrm{loc}}(\boldsymbol{R}^n)),$$

where the supremum is taken over all cubes Q containing y:

THEOREM 2.5. *Let u and w be weights on $\boldsymbol{R}^n$. The following conditions are equivalent:*

(a) *There exists C such that for every $\lambda > 0$ and every $f \in L^1_{\mathrm{loc}}(\boldsymbol{R}^n)$*

$$u\{y \in \boldsymbol{R}^n\colon f^*(y) > \lambda\} \leqslant \frac{C}{\phi(\lambda)}\int_{\boldsymbol{R}^n}\phi(|f|)\,w\,dy. \tag{2.6}$$

(b) *The pair (u, w) satisfies A_ϕ, i.e. there exists C such that for every cube Q and every positive real ε* .

$$\Big(|Q|^{-1}\int_Q \varepsilon u\,dy\Big)\varphi\Big(|Q|^{-1}\int_Q \varrho(1/\varepsilon w)\,dy\Big) \leqslant C. \tag{2.7}$$

(In condition (b) of Theorem 2.5 we assume that (2.7) holds if for some cube Q, $\int_Q \varrho(1/\varepsilon w)\,dx = \infty$ and $u(Q) = 0$ and, likewise, if $u(Q) = \infty$ and $\int_Q \varrho(1/\varepsilon w)\,dx = 0$.)

Assume that (u, w) satisfies condition (ii) of Lemma 2.4. It is easy to see that the pair of functions $\bar{u}, \bar{w}$: $\boldsymbol{R} \to \boldsymbol{R}^+$ defined by $\bar{u}(y) = u([y])$ and $\bar{w}(y) = w([y])$, where $[y]$ denotes the integer part of y, satisfies A_ϕ and therefore we have (2.6) for $(\bar{u}, \bar{w})$.

Given F: $\boldsymbol{Z} \to \boldsymbol{R}$ let f: $\boldsymbol{R} \to \boldsymbol{R}$ be defined by $f(y) = F([y])$. Fix $k \in \boldsymbol{Z}$ and $y \in I_k = \{x \in \boldsymbol{R}\colon k \leqslant x \leqslant k+1\}$; then for every average $(m+n+1)^{-1}\sum_{i=-m}^{n}|F(k+i)|$ there is an interval I containing y such that

$$|I|^{-1}\int_I |f|\,dy = (m+n+1)^{-1}\sum_{i=-m}^{n}|F(k+i)|$$

and therefore $\boldsymbol{m}F(k) \leqslant f^*(y)$. Consequently, for every $\lambda > 0$ we get

$$\sum_{\{k\in \mathbf{Z}:\ mF(k)>\lambda\}} u(k) \leqslant \bar{u}\{y\in \mathbf{R}:\ f^*(y)>\lambda\} \leqslant \frac{C}{\phi(\lambda)} \sum_{k\in \mathbf{Z}} \phi(|F(k)|)\, w(k)$$

and thus we obtain (i).

The proof of (i) $\Rightarrow$ (ii) of Lemma 2.4 is similar to the proof of (a) $\Rightarrow$ (b) of Theorem 2.5 in [6]; anyway, for our purpose the implication (ii) $\Rightarrow$ (i) is enough.

The following result for the operator $\boldsymbol{m}$ follows from (ii) $\Rightarrow$ (i) of Lemma 2.4.

COROLLARY 2.8. *If the pair (u, w) of weights on X satisfies $A_\phi(T, \mu)$ then there exists C such that for all $\lambda > 0$, for every mapping F: $\mathbf{Z} \to \mathbf{R}$ and μ-a.e. x in X*

$$\sum_{\{k\in \mathbf{Z}:\ mF(k)>\lambda\}} u(T^k x) \leqslant \frac{C}{\phi(\lambda)} \sum_{k\in \mathbf{Z}} \phi(|F(k)|)\, w(T^k x). \tag{2.9}$$

Proof of Corollary 2.8. It is enough to obtain (2.9) for a.e. x in Y, where Y is defined by (1.10), since if $w(T^k x) = 0$ for some integer k then it follows from $A_\phi(T, \mu)$ that $u(T^j x) = 0$ for every integer j.

For x in Y let u_x, w_x: $\mathbf{Z} \to \mathbf{R}$ be the mappings given by $u_x(k) = u(T^k x)$ and $w_x(k) = w(T^k x)$. The $A_\phi(T, \mu)$-condition for (u, w) shows that, for μ-a.e. x in Y, we have condition (ii) of Lemma 2.4 for (u_x, w_x) with a constant C also independent of x, and therefore we get (2.9).

Proof of (d) $\Rightarrow$ (a). For every integer $L \geqslant 1$ we consider the truncated maximal operator $M_{T,L}$ defined by

$$M_{T,L} f = \sup_{0\leqslant m,n\leqslant L} R_{m,n}|f| \quad (f \in \mathfrak{M}).$$

For a given mapping F defined in $\mathbf{Z}$ let

$$\boldsymbol{m}_L F(k) = \sup_{0\leqslant m,n\leqslant L} (m+n+1)^{-1} \sum_{i=-m}^{n} |F(k+i)|$$

and for given $g \in \mathfrak{M}$ and $x \in X$ we denote by g_x the mapping defined by $g_x(k) = g(T^k x)$.

Fix $\lambda > 0$ and $f \in \mathfrak{M}$. Since T preserves the measure μ we have

$$u\{x\in X:\ M_{T,L} f(x) > \lambda\} = (2L+1)^{-1} \int_X \sum_{|k|\leqslant L} u(T^k x)\, \chi_{A_k}(x)\, d\mu(x), \tag{2.10}$$

where $A_k = \{x \in X:\ \boldsymbol{m}_L f_x(k) > \lambda\}$.

If $1 \leqslant k \leqslant N$ then $\boldsymbol{m}_L f_x(k) \leqslant \boldsymbol{m}(f_x \chi_{[-2L,2L]})(k)$, where $\boldsymbol{m}$ is defined by (2.3), and therefore Corollary 2.8 shows that for μ-a.e. x in X

$$\sum_{|k|\leqslant L} u(T^k x)\, \chi_{A_k}(x) \leqslant \sum_{\{k\in \mathbf{Z}:\ m(f_x\chi_{[-2L,2L]})(k)>\lambda\}} u(T^k x)$$
$$\leqslant \frac{C}{\phi(\lambda)} \sum_{|k|\leqslant 2L} \phi(|f(T^k x)|)\, w(T^k x)$$

with C independent of x, λ and f. Thus, substituting these inequalities in (2.10), we get

$$u\{x\in X:\ M_T f(x) > \lambda\} = \lim_{L\to\infty} u\{x\in X:\ M_{T,L} f(x) > \lambda\} \leqslant \frac{2C}{\phi(\lambda)} \int_X \phi(|f|)\, d\mu,$$

which proves condition (a).

Proof of (b) $\Rightarrow$ (d). It suffices to prove that for every positive $k \in \mathbf{Z}$ and μ-a.e. x in Y, where Y is the set defined by (1.10), we have

$$\Big(k^{-1} \sum_{i=0}^{k-1} u(T^i x)\Big)\, \varphi\Big(k^{-1} \sum_{i=0}^{k-1} \varrho\big(1/w(T^i x)\big)\Big) \leqslant C, \tag{2.11}$$

with C depending only on ϕ and the constant of condition (b), and moreover $\mu(E) = 0$, where E is defined by (1.11).

$\mu(E) = 0$ follows easily from (b) by taking $f = \chi_E$ since, then, $u\{x \in X:\ R_{k,k}\chi_E(x) > \lambda\} = 0$ for every integer k and every $\lambda > 0$ and, thus, for all j we have $u(T^j E) \leqslant u\{x\in X:\ R_{j,j}\chi_E(x) > (2j+2)^{-1}\} = 0$. This shows that

$$\int_E \sum_{j=-\infty}^{\infty} u \circ T^j\, d\mu = 0$$

and therefore $\mu(E) = 0$ because $\sum_{j=-\infty}^{\infty} u(T^j x) \neq 0$ if $x \in E$.

In order to obtain (2.11) we follow the ideas given in [3] and (14) and, therefore, we first need the concept of ergodic rectangle, given in [3].

DEFINITION 2.12. Let l be an integer with $l \geqslant 1$ and let B be a measurable set with $\mu(B) > 0$ such that $T^i B \cap T^j B = \emptyset$, $i \neq j$, $0 \leqslant i, j \leqslant l-1$. Then the set $\mathfrak{R} = \bigcup_{i=0}^{k-1} T^i B$ will be called the *ergodic rectangle* with *base* B and *length* l.

The following lemma is proved in [14].

LEMMA 2.13. *Let Y be a measurable subset of X and let k be a positive integer. Then there exists a countable family $\{B_j:\ j \in \mathbf{Z}^+\}$ of sets of finite measure such that:*

(i) $Y = \bigcup_{j=0}^{\infty} B_j$.

(ii) $B_i \cap B_j = \emptyset$ *if* $i \neq j$.

(iii) *For every j, B_j is the base of an ergodic rectangle of length $l(j) \leqslant k$ and such that if $l(j) < k$, then $T^{l(j)} A = A$ for every measurable set $A \subset B_j$.*

Now, fix a positive integer k and let $\{B_j\}_{j=0}^{\infty}$ be the sequence given by Lemma 2.13 for the set Y defined in (1.10) and the integer k. It is enough to prove (2.11) for μ-a.e. x in B_j with the constant C independent of B_j and the integer k.

For B_j fixed and each integer r, we define

$$H_{j,r} = \{x \in B_j\colon\ 2^r < k^{-1} \sum_{i=0}^{k-1} \varrho\big(1/w(T^i x)\big) \leqslant 2^{r+1}\}.$$

For any $A \subset H_{j,r}$, with $\mu(A) > 0$, let $\mathfrak{R}$ be the rectangle with base A of length $l(j) \leqslant k$ and let $f = \varrho(w^{-1}\chi_{\mathfrak{R}})$. Then $\mathfrak{R} \subset \{x \in X\colon |R_{k,k} f(x)| > 3^{-1} 2^r\}$ and consequently

$$\phi(2^r) u(\mathfrak{R}) \leqslant C \int_{\mathfrak{R}} \phi\big(\varrho(1/w)\big) w\, d\mu \leqslant C \int_{\mathfrak{R}} \varrho(1/w)\, d\mu,$$

where C depends only on ϕ and the constant in condition (b). Therefore, there exists a constant C such that for every integer k and every measurable set $A \subset H_{j,r}$

$$\int_A \phi(2^r) k^{-1} \sum_{i=0}^{k-1} u(T^i x)\, d\mu(x) \leqslant C \int_A k^{-1} \sum_{i=0}^{k-1} \varrho\big(1/w(T^i x)\big)\, d\mu(x).$$

On the other hand, since ϕ satisfies Δ_2, there exists $\alpha > 1$ such that $\phi(2s) \leqslant \alpha\phi(s)$ and $s\varphi(s) \leqslant \alpha\phi(s)$, $s \geqslant 0$. Then for every k and μ-a.e. $x \in H_{j,r}$ we get (2.11) with constant $C\alpha^2$ and thus the same is obtained for μ-a.e. $x \in B_j$, which proves that (u, w) satisfies $A_\phi(T, \mu)$.

In this way, the proof of Theorem 2.1 is complete.

In the case $\mu(X) < \infty$, Theorem 2.1 allows us to characterize those weights w on X for which there exists a weight $u > 0$ such that M_T is of weak type (ϕ, ϕ) with respect to the measures $w\,d\mu$ and $u\,d\mu$.

THEOREM 2.14. *Let $(X, \mathcal{M}, \mu)$ be a finite measure space and w a weight on X. There exists a positive weight u on X such that M_T is of weak type (ϕ, ϕ) with respect to (u, w) if and only if*

$$\sup_{\varepsilon > 0} \varepsilon\varphi\big(M_T \varrho(1/\varepsilon w)(x)\big) < \infty \quad \textit{for } \mu\textit{-a.e. } x \in X. \tag{2.15}$$

Proof. Taking into account Theorem 2.1 it suffices to prove that condition (2.15) is equivalent to the existence of a weight $u > 0$ such that (u, w) satisfies $A_\phi(T, \mu)$.

If (2.15) holds, then the measurable function

$$u(x) = \sup_{\varepsilon > 0} \big(\varepsilon\varphi\big(M_T \varrho(1/\varepsilon w)(x)\big)\big)^{-1}$$

satisfies $0 < u(x) < \infty$ μ-a.e. and the pair (u, w) satisfies $A_\phi(T, \mu)$ with constant $C = 1$.

Now, assume that there is $u > 0$ such that (u, w) belongs to the $A_\phi(T, \mu)$-class. It is no restriction to assume that $u \in L_1(\mu)$. Then it follows from *Birkhoff's individual ergodic theorem* (see Theorem 1.2.3 in [13]) that there exists a T-invariant function $\hat{u} \in L_1(\mu)$, with $\|\hat{u}\|_1 \leqslant \|u\|_1$, such that

$$\lim_{k \to \infty} k^{-1} \sum_{i=0}^{k-1} u(T^i x) = \hat{u}(x) \quad \text{for } \mu\text{-a.e. } x \in X.$$

Moreover, for each T-invariant $A \in \mathcal{M}$ with $\mu(A) < \infty$, $\int_A \hat{u}\, d\mu = \int_A u\, d\mu$. Then if $A = \{x \in X\colon \hat{u}(x) = 0\}$ we have $\int_A u\, d\mu = 0$ and therefore $\hat{u}(x) > 0$ for μ-a.e. x in X. (Observe that if T is, in addition, an ergodic transformation, then $\hat{u}(x) = \mu(X)^{-1} \int_X u\, d\mu$ when $\mu(X) < \infty$ but $\hat{u}$ is the null function when $\mu(X) = \infty$.)

Therefore, for almost every $x \in X$ there exists a real $\alpha(x) > 0$ such that for all integer $k > 0$ and every real $\varepsilon > 0$

$$\varepsilon\varphi\Big(k^{-1} \sum_{i=0}^{k-1} \varrho\big(1/\varepsilon w(T^i x)\big)\Big) \leqslant C/\alpha(x), \tag{2.16}$$

with C independent of x, k and ε.

It is obvious that we can obtain an inequality as (2.16) with T^{-1} instead of T. On the other hand, since ϕ satisfies Δ_2, it is easy to see that there is a constant $K > 0$ such that

$$\varphi(s+t) \leqslant K\big(\varphi(s) + \varphi(t)\big), \qquad \varphi\big(\varrho(t)\big) \leqslant Kt$$

for any $s, t \geqslant 0$. Consequently, for almost every x in X there is a real $\beta(x) > 0$ such that for every pair of nonnegative integers m, n and every $\varepsilon > 0$ we have $\varepsilon\varphi\big(R_{m,n} \varrho(1/\varepsilon w)(x)\big) \leqslant \beta(x)$ and therefore

$$\varphi\big(M_T \varrho(1/\varepsilon w)(x)\big) \leqslant \varphi\big(\varrho(\varepsilon^{-1}\beta(x))\big) \leqslant K\varepsilon^{-1}\beta(x),$$

which proves (2.15).

Remarks. 1) In the case $\phi(s) = s^p$ with $1 < p < \infty$ condition (2.15) of Theorem 2.14 reduces to the following condition:

$$M_T(w^{-q/p})(x) < \infty \quad \textit{for } \mu\textit{-a.e. } x \in X, \tag{2.17}$$

where q is the conjugate exponent of p.

2) In proving that condition (2.20) implies the existence of $u > 0$ such that M_T is of weak type (ϕ, ϕ) with respect to (u, w), we do not use the fact that $\mu(X) < \infty$.

3) It follows from Theorem 2.14 that if w satisfies (2.15) then $M_T f(x) < \infty$ for μ-a.e. $x \in X$ for every $f \in L_\phi(w\,d\mu)$, since $\lim_{s \to \infty} \phi(s) = \infty$. However, the finiteness of the maximal operator can be obtained with a condition less restrictive, in general, than (2.15), namely:

PROPOSITION 2.18. *Let w be a weight on X such that*

$$M_T \varrho(1/w)(x) < \infty \quad \text{for } \mu\text{-a.e. } x \in X. \tag{2.19}$$

Then $M_T f(x) < \infty$ for μ-a.e. $x \in X$ for every f in $L_\phi(wd\mu)$.

Indeed, it follows from Young's inequality that $M_T f \leqslant M_T(\phi(|f|)w) + M_T \varrho(1/w)$ for every $f \in \mathfrak{M}$. On the other hand, M_T is of weak type (1, 1) with respect to the measure μ and therefore if $f \in L_\phi(wd\mu)$ we get $M_T f(x) < \infty$ μ-a.e.

4) In the case $\phi(s) = s^p$, it follows from *Nikishin's theorem* [9] that the finiteness of the operator M_T, i.e. $M_T f(x) < \infty$ μ-a.e. for every $f \in L_p(wd\mu)$, implies the existence of a positive weight u such that M_T is of weak type (p, p) with respect to (u, w). If the above result could be extended to any Orlicz space, then we would conclude from Theorem 2.14 and Proposition 2.18 that conditions (2.15) and (2.19) are equivalent for any N-function ϕ. In the case $\phi(s) = s^p$, $1 < p < \infty$, this is true since both conditions reduce to (2.17), but in general we think that (2.19) does not imply (2.15), though we do not know any counterexample. Therefore, we think that the finiteness of the maximal operator for every $f \in L_\phi(wd\mu)$ does not, in general, imply the weak type (ϕ, ϕ) for some pair (u, w) with $u > 0$.

3. Weighted strong type integral inequalities for the ergodic maximal operator. In this section we study the pairs of weights (u, w) for which M_T satisfies a weighted integral inequality of the form

$$\int_X \phi(M_T f) u d\mu \leqslant C \int_X \phi(|f|) w d\mu \quad (f \in \mathfrak{M}), \tag{3.1}$$

with C independent of f; more precisely, we characterize the weights w for which we have (3.1) with $u = w$.

In general, it is obvious that inequality (3.1) implies inequality (1.3), i.e. the weak type (ϕ, ϕ) for M_T, and therefore when (3.1) holds we shall say that M_T is of *strong type (ϕ, ϕ) with respect to (u, w).*

It follows from Theorem 2.1 that a necessary condition for M_T to be of strong type is the $A_\phi(T, \mu)$-condition for (u, w), but, in general, it does not suffice, as example (3.1) in [14] reveals, in the case $\phi(s) = s^2$.

In the L_p-case, $1 < p < \infty$, the strong type (p, p) with respect to (u, w) for M_T is characterized in Theorem 3.5 of [14] by means of the $S_p(T, \mu)$-condition. In our case, we can extend this condition to a condition which is necessary to obtain (3.1). Namely, given two nonnegative integers r and s we consider the truncated maximal operator $M_{r,s}$ defined by

$$M_{r,s} f = \max_{0 \leqslant m \leqslant r; 0 \leqslant n \leqslant s} R_{m,n}|f| \quad (f \in \mathfrak{M}).$$

With this notation we give the following definition:

DEFINITION 3.2. Let u and w be two weights on X and ϕ an N-function. Let Y and E be the sets defined by (1.10) and (1.11) respectively. We shall say that the pair (u, w) satisfies the *$S_\phi(T, \mu)$-condition* if $\mu(E) = 0$ and there exists a positive constant C such that for every positive integer k, every positive real ε and μ-a.e. x in Y

$$\sum_{i=0}^{k-1} \phi(M_{i,k-i-1} \varrho(1/\varepsilon w)(T^i x)) \varepsilon u(T^i x) \leqslant C \sum_{i=0}^{k-1} \varrho(1/\varepsilon w)(T^i x). \tag{3.3}$$

When (w, w) satisfies $S_\phi(T, \mu)$ we shall say that w satisfies $S_\phi(T, \mu)$.

If $\phi(s) = p^{-1} s^p$, $p > 1$, the $S_\phi(T, \mu)$-condition reduces to saying that for every positive integer k and μ-a.e. x in Y

$$\sum_{i=0}^{k-1} (M_{i,k-i-1} w^{-q/p}(T^i x))^p u(T^i x) \leqslant C \sum_{i=0}^{k-1} w^{-q/p}(T^i x), \tag{3.4}$$

where q is the conjugate of p, which gives the $S_p'(g, T, \mu)$-condition for $g = 1$, given in [14]. We shall simply say that (u, w) satisfies $S_p(T, \mu)$.

The $S_\phi(T, \mu)$-condition implies the $A_\phi(T, \mu)$-condition, since the left-hand side of inequality (3.3) is greater than or equal to

$$\beta k^{-1} \sum_{i=0}^{k-1} \varrho(1/\varepsilon w)(T^i x) \varphi\left(k^{-1} \sum_{i=0}^{k-1} \varrho(1/\varepsilon w)(T^i x)\right) \sum_{i=0}^{k-1} \varepsilon u(T^i x),$$

where β is such that $\phi(s) \geqslant \beta s \varphi(s)$, $s \geqslant 0$.

The $S_\phi(T, \mu)$-condition is a necessary condition for the strong type (ϕ, ϕ) of the maximal ergodic operator, as the following result shows:

PROPOSITION 3.5. *Let u and w be weights on X and assume that* (3.1) *holds. Then the pair (u, w) satisfies the $S_\phi(T, \mu)$-condition.*

Proof. Since (3.1) implies the $A_\phi(T, \mu)$-condition for (u, w), we have $\mu(E) = 0$. It follows from (3.1) that in order to obtain (3.3) it is enough to prove that there exists a constant C, depending only on ϕ and the constant of condition (3.1), such that for every positive integer k and for μ-a.e. x in Y

$$\sum_{i=0}^{k-1} \phi(M_{i,k-i-1} \varrho(1/w)(T^i x)) u(T^i x) \leqslant C \sum_{i=0}^{k-1} \varrho(1/w)(T^i x). \tag{3.6}$$

Fix k; let $\{B_j\}$ be the sequence given by Lemma 2.13 for the set Y. For B_j fixed, let A be a measurable subset of B_j and $\mathfrak{R}$ the rectangle with base A of length $l(j) \leqslant k$. Then we have

$$\int_A \sum_{i=0}^{k-1} \phi(M_{i,k-i-1} \varrho(\chi_{\mathfrak{R}}/w)(T^i x)) u(T^i x) d\mu(x)$$

$$\leqslant \int_A \sum_{i=0}^{k-1} \phi\big(M_T \varrho(\chi_{\mathfrak{R}}/w)(T^i x)\big) u(T^i x)\, d\mu(x)$$

$$\leqslant 2kl(j)^{-1} \int_{\mathfrak{R}} \phi\big(M_T \varrho(\chi_{\mathfrak{R}}/w)\big) u d\mu \leqslant 2kl(j)^{-1}\, C \int_{\mathfrak{R}} \varrho(1/w)\, d\mu$$

$$\leqslant \int_A 4C \sum_{i=0}^{k-1} \varrho(1/w)(T^i x)\, d\mu(x)$$

and thus we obtain (3.6) for μ-a.e. x in B_j with a constant independent of B_j and k, which proves (3.6) for μ-a.e. x in Y.

It follows from Proposition 3.5 that *every pair* (u, w) *of the* $A_1(T, \mu)$-*class satisfies the* $S_\phi(T, \mu)$-*condition*, since Theorem 2.7 in [14] shows that if (u, w) satisfies $A_1(T, \mu)$ the maximal operator M_T is of weak type (1, 1) with respect to (u, w) and, then, it follows from the interpolation theorem 1.7 that M_T is of strong type (ϕ, ϕ) with respect to (u, w).

It is an open problem whether or not the $S_\phi(T, \mu)$-condition implies, in general, the strong type (ϕ, ϕ) for M_T. (In the case $\phi(s) = s^p$, $p > 1$, this is true.) We shall see that this is so when $u = w$; more exactly:

THEOREM 3.7. *Let* w *be a weight on* X *and let* $p = q_\phi$ *be the lower exponent of* ϕ. *Then the following conditions are equivalent*:

(a) *There exists* C *such that for every* $f \in \mathfrak{M}$

$$\int_X \phi(M_T f)\, w d\mu \leqslant C \int_X \phi(|f|)\, w d\mu.$$

(b) *There exists* C *such that for every* $f \in \mathfrak{M}$ *and every* $\lambda > 0$

$$w\{x \in X\colon\ M_T f(x) > \lambda\} \leqslant \frac{C}{\phi(\lambda)} \int_X \phi(|f|)\, w d\mu.$$

(c) *There exists* C *such that for every* $f \in \mathfrak{M}$

$$\sup_{m,n \geqslant 0} \int_X \phi(|R_{m,n} f|)\, w d\mu \leqslant C \int_X \phi(|f|)\, w d\mu.$$

(d) *There exists* C *such that for every* $f \in \mathfrak{M}$ *and every* $\lambda > 0$

$$\sup_{m,n \geqslant 0} w\{x \in X\colon\ |R_{m,n} f(x)| > \lambda\} \leqslant \frac{C}{\phi(\lambda)} \int_X \phi(|f|)\, w d\mu.$$

(e) w *satisfies* $A_\phi(T, \mu)$.
(f) w *satisfies* $A_p(T, \mu)$.
(g) w *satisfies* $S_\phi(T, \mu)$.
(h) w *satisfies* $S_p(T, \mu)$.

Proof. Conditions (b)–(e) are equivalent as a consequence of Theorem 2.1. We know that (a) implies (b), (g) implies (e) and Proposition 3.5 shows

that (g) follows from (a). On the other hand, (f) and (h) are equivalent as a consequence of Theorem 4.1 in [14]. Therefore, it will suffice to prove (e) $\Rightarrow$ (f) and (f) $\Rightarrow$ (a).

Proof of (f) $\Rightarrow$ (a). The proof of this implication follows easily from interpolation theorem 1.7 and the following results:

LEMMA 3.8. *Let* $1 < p < \infty$. *If* w *satisfies* $A_p(T, \mu)$ *then there exists* r, *with* $1 < r < p$, *such that* w *satisfies* $A_r(T, \mu)$.

LEMMA 3.9. *Let* $1 < r < \infty$. *If* w *satisfies* $A_r(T, \mu)$ *then* M_T *is bounded in* $L_r(w d\mu)$.

The proofs of Lemmas 3.8 and 3.9 can be found in [3] where T is ergodic and $(X, \mathcal{M}, \mu)$ is a nonatomic probability space. The proofs are valid in our setting because these properties are not used.

Proof of (e) $\Rightarrow$ (f). For x in Y let $w_x\colon \boldsymbol{Z} \to \boldsymbol{R}$ be the mapping defined by $w_x(k) = w(T^k x)$. The $A_\phi(T, \mu)$-condition for w implies that there exists a constant C such that for μ-a.e. $x \in Y$, for every $\varepsilon > 0$ and for every pair of positive integers n and k

$$\Big(n^{-1} \sum_{i=0}^{n-1} \varepsilon w_x(k+i)\Big)\, \varphi\Big(n^{-1} \sum_{i=0}^{n-1} \varrho\big(1/\varepsilon w_x(k+i)\big)\Big) \leqslant C.$$

Let $w_x^*\colon \boldsymbol{R} \to \boldsymbol{R}$ be given by $w_x^*(y) = w_x([y])$. Then for μ-a.e. $x \in Y$ we find that w_x^* satisfies A_ϕ with a constant independent of x, i.e. there exists C such that for μ-a.e. $x \in Y$, every interval I in $\boldsymbol{R}$ and every $\varepsilon > 0$

$$\Big(|I|^{-1} \int_I \varepsilon w_x^*\, dy\Big)\, \varphi\Big(|I|^{-1} \int_I \varrho(1/\varepsilon w_x^*)\, dy\Big) \leqslant C.$$

In [11], Kerman and Torchinsky prove that w^* is in the A_ϕ-class if and only if w^* is in the A_p-class, where $p = q_\phi$. It follows from the proof given in [11] that the constant in the A_p-condition depends only on the constant in the A_ϕ-condition and on the N-function ϕ. Consequently, we see that every w_x^* satisfies A_p with a constant independent of x. Therefore, for μ-a.e. $x \in Y$ and for every positive integer k we get, taking $I = [0, k]$,

$$\Big(k^{-1} \sum_{i=0}^{k-1} w(T^i x)\Big)\Big(k^{-1} \sum_{i=0}^{k-1} w(T^i x)^{-q/p}\Big)^{p/q}$$
$$= \Big(|I|^{-1} \int_I w_x^*\, dy\Big)\Big(|I|^{-1} \int_I (w_x^*)^{-q/p}\, dy\Big)^{p/q} \leqslant C$$

and thus we have the $A_p(T, \mu)$-condition for w.

We have finished the proof of the theorem.

Note. A direct proof of (e) $\Rightarrow$ (a) can be given using a similar method to the one given in the proof of (d)⇒(a) of Theorem 2.1 and taking into account the result in [11] for the Hardy–Littlewood maximal operator.

4. Convergence of the averages. First, for the Cesàro averages we have the following:

PROPOSITION 4.1. *Let w be a weight on X satisfying $M_T\varrho(1/w)(x) < \infty$ for μ-a.e. $x \in X$. Then the averages*

$$n^{-1}\sum_{i=0}^{n-1} f\circ T^i, \quad (2n+1)^{-1}\sum_{i=-n}^{n} f\circ T^i$$

converge μ-almost everywhere for every f in $L_\phi(wd\mu)$.

Proof. It follows from Proposition 2.18 and the Banach principle that it is enough to prove the almost everywhere convergence for all f in a dense subset of $V=[L_\phi(wd\mu), \|\ \|_{(\phi),wd\mu}]$. The set $L_1(\mu)\cap L_\phi(wd\mu)$ is dense in V since the simple functions with respect to the measure $wd\mu$ belong to $L_1(\mu)$ and ϕ satisfies Δ_2. On the other hand, it is well known that for the functions in $L_1(\mu)$ we have the a.e. convergence. Thus, we get the a.e. convergence for every $f\in L_\phi(wd\mu)$.

Under the hypothesis of Proposition 4.1 we cannot assert that for each f in $L_\phi(wd\mu)$ the pointwise limit functions of the averages $R_{0,n-1}f$ and $R_{n,n}f$ belong to $L_\phi(wd\mu)$; however, this can be ensured if w satisfies $A_\phi(T,\mu)$; in fact, *if w satisfies $A_\phi(T,\mu)$ the mean ergodic theorem holds, i.e. for every $f\in L_\phi(wd\mu)$ the sequence $R_{0,n-1}f$ converges in the norm topology*, since *for a positive linear operator S on a reflexive Banach lattice B, the Cesàro averages $((I+S+\ldots+S^{n-1})/n)f$ converge in norm for every $f\in B$ if, and only if, S is a Cesàro-mean-bounded operator, i.e. the norms of the Cesàro averages are uniformly bounded in B* (see Theorem 4.2 in [5]), and our operator is Cesàro-mean-bounded as follows from Theorem 3.7. It is clear that the same happens for the averages $R_{n-1,0}f$ and consequently for $R_{n,n}f$.

On the other hand, let us observe that if w satisfies $A_\phi(T,\mu)$ then w satisfies condition (2.19) when $\mu(X)<\infty$.

In the following result we consider certain weighted averages which include the Cesàro averages.

DEFINITION 4.2. A sequence of complex numbers $\{b_k\}$ is called a *Besicovitch sequence* if for every $\varepsilon>0$ there exists a trigonometric polynomial α_ε such that

$$\limsup_{n\to\infty} n^{-1}\sum_{k=0}^{n-1}|b_k-\alpha_\varepsilon(k)|<\varepsilon.$$

THEOREM 4.3 (Individual and mean weighted ergodic theorem). *If w*

satisfies the $A_\phi(T,\mu)$-condition, then for every $f\in L_\phi(wd\mu)$ and every bounded Besicovitch sequence $\{b_k\}$ there exists $\bar f\in L_\phi(wd\mu)$ such that

$$\lim_{n\to\infty} n^{-1}\sum_{k=0}^{n-1} b_k f(T^k x)=\bar f(x) \quad \mu\text{-a.e.},$$

$$\lim_{n\to\infty}\|n^{-1}\sum_{k=0}^{n-1} b_k f\circ T^k-\bar f\|_{(\phi),wd\mu}=0.$$

Moreover, the same happens for the averages $(2n+1)^{-1}\sum_{k=-n}^{n} b_k f\circ T^k$.

Proof. Fix a nonnegative integer m and let S be the linear operator defined by $Sf=e^{im}f\circ T$.

Since w satisfies $A_\phi(T,\mu)$ it follows from Theorem 3.7 that S is a Cesàro-mean-bounded operator. On the other hand, if $A\in\mathcal{M}$ with $w(A)<\infty$ the sequence $\phi(|S^n\chi_A/n|)$ is dominated by $\phi(2M_T\chi_A)$ which belongs to $L_1(wd\mu)$; therefore, Lebesgue's dominated convergence theorem shows that $\lim_{n\to\infty}\int_X\phi(|S^n\chi_A/n|)\,wd\mu=0$ and consequently $\lim_{n\to\infty}\|S^n\chi_A/n\|_{(\phi),wd\mu}=0$ since ϕ satisfies Δ_2. Since $V=[L_\phi(wd\mu), \|\ \|_{(\phi),wd\mu}]$ is reflexive and $\{\chi_A\colon w(A)<\infty\}$ is a fundamental set in V it follows from Corollary VIII.5.4 in [4] that for every $f\in L_\phi(wd\mu)$ the Cesàro averages

$$A_n f=n^{-1}\sum_{k=0}^{n-1} e^{imk} f\circ T^k$$

converge in V. Therefore, Corollary VIII.5.2 in [4] (see also 2.1 in [13]) shows that $L_\phi(wd\mu)$ is the closure of the direct sum of the set of fixed points of S and the space $(I-S)L_\phi(wd\mu)$.

On the other hand, it is easy to see that if $f=h+g-Sg$, with $Sh=h$ and g a bounded function, then $\{A_n f\}$ converges μ-a.e. Therefore, it follows from Theorem 3.7 and the Banach principle that for every $f\in L_\phi(wd\mu)$ the averages $A_n f$ converge μ-a.e. and consequently for every trigonometric polynomial α and $f\in L_\phi(wd\mu)$ the limit

$$\lim_{n\to\infty} n^{-1}\sum_{k=0}^{n-1}\alpha(k)f(T^k x)$$

exists and is finite μ-a.e.

Given a Besicovitch sequence $\{b_k\}$ and $f\in L_\phi(wd\mu)\cap L_\infty(wd\mu)$, we find that for every $\varepsilon>0$ there exists a trigonometric polynomial α_ε such that

$$\limsup_{n\to\infty}\Big|n^{-1}\sum_{k=0}^{n-1} b_k f(T^k x)-n^{-1}\sum_{k=0}^{n-1}\alpha_\varepsilon(k)f(T^k x)\Big|<\varepsilon\|f\|_\infty \quad \mu\text{-a.e.}$$

and therefore the sequence

$$T_n f=n^{-1}\sum_{k=0}^{n-1} b_k f\circ T^k$$

converges μ-a.e. Thus, if $\{b_k\}$ is bounded it follows again from Theorem 3.7 that for every $f \in L_\phi(wd\mu)$ the weighted averages $T_n f$ converge μ-a.e. to a function from $L_\phi(wd\mu)$. We conclude, taking into account Lebesgue's dominated theorem, that the limit function of $T_n f$ in the pointwise sense is also the limit in norm.

Finally, it is clear that once we have the convergence of $T_n f$, we get the same for the averages $(2n+1)^{-1}\sum_{k=-n}^{n} b_k f\circ T^k$.

5. Weighted integral inequalities for other sublinear operators. The ergodic Hilbert transform. In [18] Rubio de Francia proves an extrapolation theorem in the theory of A_p-weights, for sublinear operators defined on a class of Lebesgue-measurable functions in $\mathbf{R}^n$. Another proof of this theorem is given by García-Cuerva in [7]. More exactly, the result is the following:

(5.1) *Let S be a sublinear operator defined on a class of measurable functions in $\mathbf{R}^n$. Let $1 \leqslant p^* < \infty$ and $1 < p < \infty$. Suppose that S is bounded in $L_{p^*}(w)$ (respectively of weak type (p^*, p^*) with respect to w) for every weight $w \in A_{p^*}$, with a norm that depends only upon the A_{p^*}-constant for w. Then, for every $w \in A_p$, S is bounded in $L_p(w)$ (respectively S is of weak type (p, p) with respect to w), with a norm that depends only upon the A_p-constant for w.* (As a corollary of this result it can be deduced that the boundedness of S in $L_p(w)$ can be obtained from the weak type (p^*, p^*).)

We observe that the proof given in [7] (see also Theorem IV.5.19 in [8]) can be easily adapted to weights $w \in A_p(T, \mu)$ and operators S defined on $\mathfrak{M}(X, \mathcal{M}, \mu)$ where $(X, \mathcal{M}, \mu)$ is a σ-finite measure space and, thus, we obtain a similar result to (5.1) in this case. Moreover, such a result can be strengthened to N-functions in the following way:

THEOREM 5.2 (Extrapolation theorem). *Let $(X, \mathcal{M}, \mu)$ be a σ-finite measure space, T an automorphism of $(X, \mathcal{M}, \mu)$ and S a sublinear operator defined on a subclass of $\mathfrak{M}(X, \mathcal{M}, \mu)$. Suppose that for some p^*, with $1 \leqslant p^* < \infty$, S satisfies*

$$w\{x \in X: |Sf(x)| > \lambda\} \leqslant C\lambda^{-p^*}\int_X |f|^{p^*} wd\mu \quad (f \in \mathfrak{M},\ \lambda > 0)$$

for every weight $w \in A_{p^*}(T, \mu)$, where C depends only on the $A_{p^*}(T, \mu)$-constant for w. Then for every N-function ϕ (which satisfies, together with its complementary N-function, the Δ_2-condition) we have

$$\int_X \phi(|Sf|)\, wd\mu \leqslant C\int_X \phi(|f|)\, wd\mu \quad (f \in \mathfrak{M}) \tag{5.3}$$

for every $w \in A_\phi(T, \mu)$ with the constant C depending only on the $A_\phi(T, \mu)$-constant for w.

Proof. Let $w \in A_\phi(T, \mu)$. It follows from Theorem 3.7 that $w \in A_p(T, \mu)$ where $p = q_\phi$ is the lower exponent of ϕ and therefore Lemma 3.8 implies that $w \in A_r(T, \mu)$ for some r such that $1 < r < p$ (r and the $A_r(T, \mu)$-constant for w depend only on the $A_\phi(T, \mu)$-constant for w). Also $w \in A_s(T, \mu)$ for s such that $p_\phi < s < \infty$, where p_ϕ is the upper exponent of ϕ.

Now, the result similar to (5.1) shows that S is simultaneously of weak type (r, r) and (s, s) with respect to $wd\mu$ with constants which depend only on the $A_\phi(T, \mu)$-constant for w, and therefore applying the interpolation theorem 1.7 we obtain (5.3) with a constant depending only upon the $A_\phi(T, \mu)$-constant for w, since the constant C in Theorem 1.7 depends only on ϕ and on the constants for the weak types (r, r) and (s, s).

Theorem 5.2 permits us to extend trivially to Orlicz spaces some of the results obtained in the theory of weights for the L_p-case, with $1 \leqslant p < \infty$. Thus, for example, let S_r and $\bar{S}_r$ be the operators defined by

$$S_r f = \Big[\sum_{k=-\infty}^{\infty} (|k|+1)^{-r} |f(T^k x)|^r\Big]^{1/r}, \tag{5.4}$$

$$\bar{S}_r f = \Big[\sum_{k=0}^{\infty} |R_{k+1,0} f - R_{k,0} f|^r + |R_{0,k+1} f - R_{0,k} f|^r\Big]^{1/r}, \tag{5.5}$$

where $1 < r < \infty$ and T is an automorphism of $(X, \mathcal{M}, \mu)$. For $r = 2$, $S_r f$ is the known *ergodic square function* associated to f, which appears in considering certain martingale inequalities.

In [15] it is proved that S_r *is of weak type* $(1, 1)$ *with respect to the measure $wd\mu$ if and only if $w \in A_1(T, \mu)$* (the same result holds for $\bar{S}_r$). Moreover, *if $1 < p < \infty$ and $w \in A_p(T, \mu)$ then S_r and $\bar{S}_r$ are of strong type (p, p) with respect to $wd\mu$.* (In every case, the constant of type (p, p) depends only on the $A_p(T, \mu)$-constant for w.) Consequently, the following corollary follows trivially from the above and from Theorem 5.2:

COROLLARY 5.6. *Let $1 < r < \infty$ and let S_r and $\bar{S}_r$ be the operators defined by* (5.4) *and* (5.5). *If w satisfies $A_\phi(T, \mu)$ then there exists C (depending only on the $A_\phi(T, \mu)$-constant for w) such that for every $f \in \mathfrak{M}$*

$$\int_X \phi(S_r f)\, wd\mu \leqslant C\int_X \phi(|f|)\, wd\mu, \qquad \int_X \phi(\bar{S}_r f)\, wd\mu \leqslant C\int_X \phi(|f|)\, wd\mu.$$

Now, our aim is to study the existence, in the pointwise and the norm sense, of the *ergodic Hilbert transform*

$$H_T f(x) = \lim_{n\to\infty} \sum_{i=-n}^{n}{}' f(T^i x)/i$$

of every $f \in L_\phi(wd\mu)$, where w is a weight of the $A_\phi(T, \mu)$-class ($'$ denotes omission of the 0th term). For this, we need to study weighted integral inequalities for the *ergodic maximal Hilbert transform*, i.e. for the operator S_T^*

defined by

$$S_T^* f = \sup_{n \geq 1} \Big| \sum_{i=-n}^{n} f(T^i x)/i \Big|$$

or else for the operator H_T^* given by

$$H_T^* f = \sup_{0 \leq m < n} \Big| \sum_{m < |i| < n} f(T^i x)/i \Big|. \tag{5.7}$$

More exactly, $S_T^* f \leq H_T^* f \leq 2 S_T^* f$.

In [1] and [2], Atencia and Martín-Reyes prove that, for a *nonatomic probability space* $(X, \mathcal{M}, \mu)$, an *ergodic* automorphism T and a *positive* weight w, the following results hold:

(5.8) *The operator H_T^* is of weak type* (1, 1) *with respect to $wd\mu$ if and only if $w \in A_1(T, \mu)$.*

(5.9) *Let $1 < p < \infty$; then H_T^* is bounded on $L_p(wd\mu)$ if and only if $w \in A_p(T, \mu)$.*

We shall prove that the hypothesis that $(X, \mathcal{M}, \mu)$ is a nonatomic probability space can be suppressed and that the ergodicity of T, and likewise the positivity of w, are not needed to obtain the boundedness of H_T^*.

First, in our setting, we get the following result:

PROPOSITION 5.10. *If w satisfies the $A_\phi(T, \mu)$-condition, then there exists C such that*

$$\int_X \phi(H_T^* f)\, w d\mu \leq C \int_X \phi(|f|)\, w d\mu \quad (f \in \mathfrak{M}), \tag{5.11}$$

where H_T^* is defined in (5.7).

Proof. Let $1 < p < \infty$. It follows from the extrapolation theorem 5.2 that in order to obtain Proposition 5.10 it is enough to prove that H_T^* is of weak type (p, p) with respect to $wd\mu$ for every $w \in A_p(T, \mu)$ with a constant depending only on the $A_p(T, \mu)$-constant for w. We shall prove directly the strong type (p, p).

For every $L \geq 1$ we consider the truncated operator $H_{T,L}^*$ defined by

$$H_{T,L}^* f(x) = \max_{0 \leq m < n \leq L} \Big| \sum_{m < |i| < n} f(T^i x)/i \Big|$$

and for a mapping F defined in $\mathbf{Z}$ and $k \in \mathbf{Z}$ let

$$h^* F(k) = \sup_{0 \leq m < n} \Big| \sum_{m < |i| < n} F(k+i)/i \Big|,$$

$$h_L^* F(k) = \max_{0 \leq m < n \leq L} \Big| \sum_{m < |i| < n} F(k+i)/i \Big|.$$

For a given $g \in \mathfrak{M}$ and $x \in X$ let g_x be the mapping defined by $g_x(k)$

$= g(T^k x)$. Then, since T preserves the measure μ, we get

$$\begin{aligned}\int_X (H_{T,L}^* f)^p\, w d\mu &= (2L+1)^{-1} \int_X \sum_{|k| \leq L} (H_{T,L}^* f)^p (T^k x)\, w(T^k x)\, d\mu(x) \\ &= (2L+1)^{-1} \int_X \sum_{|k| \leq L} (h_L^* f_x(k))^p\, w_x(k)\, d\mu(x) \\ &\leq (2L+1)^{-1} \int_X \sum_{k \in \mathbf{Z}} (h^*(f_x \chi_{[-2L, 2L]})(k))^p\, w_x(k)\, d\mu(x).\end{aligned} \tag{5.12}$$

On the other hand, there exists a constant C, depending only on the $A_p(T, \mu)$-constant for w, such that for μ-a.e. $x \in Y$, where Y is defined in (1.10), and for every pair of integers m, n with $m \leq n$

$$\sum_{k=m}^{n} w_x(k) \cdot \Big(\sum_{k=m}^{n} w_x(k)^{-1/(p-1)} \Big)^{p-1} \leq C(n+m+1)^p,$$

and therefore Theorem 10 in [10] shows that for μ-a.e. x in Y and every mapping F

$$\sum_{k \in \mathbf{Z}} |h^* F(k)|^p\, w_x(k) \leq C \sum_{k \in \mathbf{Z}} |F(k)|^p\, w_x(k), \tag{5.13}$$

where C depends only on the $A_p(T, \mu)$-constant for w. Moreover, if $w(T^k x) = 0$ for some integer k then $w(T^j x) = 0$ for every integer j and consequently (5.13) holds also for x in $X - Y$.

Now, it follows from (5.12) and (5.13) that

$$\begin{aligned}\int_X (H_{T,L}^* f)^p\, w d\mu &\leq C(2L+1)^{-1} \int_X \sum_{|k| \leq 2L} |f(T^k x)|\, w(T^k x)\, d\mu(x) \\ &= C \frac{4L+1}{2L+1} \int_X |f|^p\, w d\mu\end{aligned}$$

and thus we get $\int_X (H_T^* f)^p\, w d\mu \leq 2C \int_X |f|^p\, w d\mu$, which proves the proposition.

COROLLARY 5.14. *If w satisfies the $A_\phi(T, \mu)$-condition then for every $f \in L_\phi(wd\mu)$ there exists a function $H_T f$ in $L_\phi(wd\mu)$ such that*

$$\lim_{n \to \infty} \sum_{i=-n}^{n}{}' f(T^i x)/i = H_T f(x) \quad \mu\text{-a.e.}$$

$$\lim_{n \to \infty} \Big\| \sum_{i=-n}^{n}{}' f \circ T^i/i - H_T f \Big\|_{(\phi), wd\mu} = 0.$$

Moreover, there exists C such that

$$\int_X \phi(|H_T f|)\, w d\mu \leq C \int_X \phi(|f|)\, w d\mu \quad (f \in \mathfrak{M}). \tag{5.15}$$

Proof. Since for every $f \in L_\phi(wd\mu)$ the Cesàro averages $R_{0,n-1} f$ converge in norm we find that $L_\phi(wd\mu)$ is the closure of the set $D = \{f\colon f = h + g - g\circ T$, with $h\circ T = h$ and g bounded$\}$. Let

$$S_n f = \sum_{i=-n}^{n}{}' f\circ T^i/i.$$

It is obvious that for every $f\in D$ the sequence $\{S_n f\}$ converges μ-a.e., since

$$S_n(g-g\circ T) = g+g\circ T - n^{-1}(g\circ T^{n+1}+g\circ T^{-n}) - \sum_{i=1}^{n-1}\frac{1}{i(i+1)}(g\circ T^{i+1}+g\circ T^{-i}).$$

Then the μ-almost everywhere convergence of $\{S_n f\}$, for every $f\in L_\phi(wd\mu)$, follows from Proposition 5.10.

Now, let $H_T f(x) = \lim_{n\to\infty} S_n f(x)$; it follows from (5.11) that $H_T f\in L_\phi(wd\mu)$ and the $\phi(|S_n f - H_T f|)$ are dominated by a function in $L_1(wd\mu)$. Thus, we get $\lim_{n\to\infty}\|S_n f - H_T f\|_{(\phi),wd\mu} = 0$.

Finally, the inequality (5.15) follows trivially from (5.11).

Now, if T is an *ergodic* automorphism, then the $A_\phi(T,\mu)$-condition is also necessary for the strong type of H_T^*; more exactly:

PROPOSITION 5.16. *Let T be an ergodic automorphism and w a positive weight. Then the following conditions are equivalent*:

(a) *There exists C such that for every $f\in\mathfrak{M}$*

$$\int_X \phi(H_T^* f)\,wd\mu \le C\int_X \phi(|f|)\,wd\mu.$$

(b) *There exists C such that for every $f\in\mathfrak{M}$ and every $\lambda>0$*

$$w\{x\in X\colon H_T^* f(x)>\lambda\} \le \frac{C}{\phi(\lambda)}\int_X \phi(|f|)\,wd\mu.$$

(c) *w satisfies $A_\phi(T,\mu)$.*

Proof. It follows from Proposition 5.10 that it will suffice to prove (b) $\Rightarrow$ (c).

First, notice that if $(X, \mathscr{M}, \mu)$ is a finite measure space which has atoms, then, since T is ergodic, every positive function in $\mathfrak{M}$ satisfies $A_\phi(T,\mu)$. In fact, let $A\subset X$ be an atom for $(X,\mathscr{M},\mu)$, i.e. $\mu(A)>0$ and for every $E\in\mathscr{M}$ with $E\subset A$ either $\mu(E)=0$ or $\mu(E)=\mu(A)$. The sets of the family $\mathscr{F} = \{T^k A\colon k\in\mathbf{Z}\}$ are also atoms and their union is T-invariant. Since T is ergodic and $\mu(X)<\infty$, X can be decomposed into a finite number of sets from $\mathscr{F}$. On the other hand, every measurable function is essentially constant over every atom, and therefore, since w is positive, we deduce that there exist constants C_1 and C_2 such that $0<C_1<w(x)<C_2$ μ-a.e. and consequently $w\in A_1(T,\mu)$, which shows that $w\in A_\phi(T,\mu)$.

Therefore, assume that the σ-finite measure space is such that either $\mu(X)=\infty$ or X has no atoms. For this case we use the following:

LEMMA 5.17. *Let $(X,\mathscr{M},\mu)$ be either a σ-finite nonatomic measure space or a σ-finite space with $\mu(X)=\infty$. Let T be an ergodic automorphism of $(X,\mathscr{M},\mu)$ and Y a measurable subset of X. Then for every positive integer k the set Y can be written as a countable union of bases of ergodic rectangles of length k.*

A proof of Lemma 5.17 can be found in [14]. (This lemma generalizes Lemma (2.2) in [3].)

Taking into account Lemma 5.17, the proof of (b) $\Rightarrow$ (c), in this case, is an adaptation of the proof given in [1], in the L_p-case, and of the proof of (b) $\Rightarrow$ (d) in Theorem 2.1. For that reason, we shall only point up some steps.

It suffices to prove that for every positive integer k and μ-a.e. x in X we have

$$(5.18)\qquad \Big(k^{-1}\sum_{i=0}^{k-1} w(T^i x)\Big)\varphi\Big(k^{-1}\sum_{i=0}^{k-1}\varrho(1/w(T^i x))\Big) \le C$$

with C depending only on the constant of condition (b).

For k fixed, there exists a countable family $\{B_j\colon j\in\mathbf{Z}\}$ of sets of finite measure such that

$$X = \bigcup_{j=0}^{\infty} B_j,\qquad B_i\cap B_j = \emptyset\quad \text{if } i\ne j,$$

and every B_j is the base of an ergodic rectangle of length $4k$.

For B_j fixed and each integer r, we define

$$D_{j,r} = \{x\in B_j\colon 2^r < (2k)^{-1}\sum_{i=0}^{k-1}\varrho(1/w(T^i x)) \le 2^{r+1}\},$$

and for $A\subset D_{j,r}$ let $\mathfrak{R}$ be the ergodic rectangle with base A of length k.

If f is a nonnegative function with support in $\mathfrak{R}$ we have

$$H_T^* f(T^n x) \ge (2k)^{-1}\sum_{i=0}^{k-1} f(T^i x)\qquad (x\in A,\ k\le n\le 2k-1).$$

Then for $f = \varrho(w^{-1}\chi_{\mathfrak{R}})$ we have

$$Q = \bigcup_{n=k}^{2k-1} T^n A \subset \{y\in X\colon H_T^*\varrho(1/w)(y) > 2^r\}$$

and therefore it follows from (b) that

$$(5.19)\qquad \phi(2^r)\,w(Q) \le C\int_{\mathfrak{R}} \varrho(1/w)\,d\mu.$$

On the other hand, if f is a nonnegative function with support in Q we get

$$H_T^* f(T^n x) \geqslant (2k)^{-1} \sum_{i=k}^{2k-1} f(T^i x) \quad (x \in A,\ 0 \leqslant n \leqslant k-1).$$

Then it follows from (b) and (5.19) that

$$\phi(2^r) w(\mathfrak{R}) \leqslant C \int_{\mathfrak{R}} \varrho(1/w)\, d\mu.$$

The last inequality was obtained in the proof of (b) ⇒ (c) of Theorem 2.1 and the argument used there can be used here to obtain (5.18). Thus, the proof of Proposition 5.16 is complete.

References

[1] E. Atencia and F. J. Martín-Reyes, *The maximal ergodic Hilbert transform with weights*, Pacific J. Math. 108 (1983), 257–263.
[2] —, —, *Weak type inequalities for the maximal ergodic function and the maximal ergodic Hilbert transform in weighted spaces*, Studia Math. 78 (1984), 231–244.
[3] E. Atencia and A. de la Torre, *A dominated ergodic estimate for L_p-spaces with weights*, ibid. 74 (1982), 35–47.
[4] N. Dunford and J. T. Schwartz, *Linear Operators I*, Wiley-Interscience, New York 1958.
[5] R. Emilion, *Mean-bounded operators and mean ergodic theorems*, J. Funct. Anal. 61 (1985), 1–14.
[6] D. Gallardo, *Weighted weak type integral inequalities for the Hardy–Littlewood maximal operator*, Israel J. Math. 66 (1) (1989), to appear.
[7] J. García-Cuerva, *An extrapolation theorem in the theory of A_p weights*, Proc. Amer. Math. Soc. 87 (1983), 422–426.
[8] J. García-Cuerva and J. L. Rubio de Francia, *Weighted Norm Inequalities and Related Topics*, North-Holland, 1985.
[9] J. E. Gilbert, *Nikishin–Stein theory and factorization with applications*, in: Proc. Sympos. Pure Math. 35 (1979), 233–267.
[10] R. Hunt, B. Muckenhoupt and R. Wheeden, *Weighted norm inequalities for the conjugate function and Hilbert transform*, Trans. Amer. Math. Soc. 176 (1973), 227–251.
[11] R. A. Kerman and A. Torchinsky, *Integral inequalities with weights for the Hardy maximal function*, Studia Math. 71 (1981), 277–284.
[12] M. A. Krasnosel'skii and Ya. B. Rutitskii, *Convex Functions and Orlicz Spaces*, Noordhoff, Groningen 1961.
[13] U. Krengel, *Ergodic Theorems*, de Gruyter Stud. Math. 6, 1985.
[14] F. J. Martín-Reyes, *Inequalities for the ergodic maximal function and convergence of the averages in weighted L^p-spaces*, Trans. Amer. Math. Soc. 296 (1986), 61–82.
[15] —, *Weights for ergodic square functions*, Ann. Inst. H. Poincaré 22 (1986), 333–345.
[16] J. Musielak, *Orlicz Spaces and Modular Spaces*, Springer, 1983.

[17] J. H. Olsen, *The individual weighted ergodic theorem for bounded Besicovitch sequences*, Canad. Math. Bull. 25 (1982), 468–471.
[18] J. L. Rubio de Francia, *Factorization theory and A_p-weights*, Amer. J. Math. 106 (1984), 533–547.

DEPARTAMENTO DE MATEMÁTICAS
FACULTAD DE CIENCIAS
UNIVERSIDAD DE MÁLAGA
Málaga, Spain

Received July 16, 1987 (2341)