Sharp constants in inequalities admitting the Calderón transference principle

DARIUSZ KOSZ

Basque Center for Applied Mathematics, 48009 Bilbao, Spain
and
Wrocław University of Science and Technology, 50-370 Wrocław, Poland
(e-mail: dkosz@bcamath.org)

Abstract. The aim of this note is twofold. First, we prove an abstract version of the Calderón transference principle for inequalities of admissible type in the general commutative multilinear and multiparameter setting. Such an operation does not increase the constants in the transferred inequalities. Second, we use the last information to study a certain dichotomy arising in problems of finding the best constants in the weak type $(1, 1)$ and strong type (p, p) inequalities for one-parameter ergodic maximal operators.

1. *Introduction*
1.1. *Historical background.* The Calderón transference principle [7] is a powerful tool in ergodic theory which allows one to transfer various quantitative results, such as inequalities of strong or weak type, from one specific dynamical system—called *canonical*—to all systems of the same type. Somewhat paradoxically, its main application is related to a purely qualitative property. Indeed, this kind of transference is widely used in verifying almost everywhere convergence of ergodic averages as the time parameter goes to infinity.

The classical approach to the latter subject relies on the following two-step procedure:
- finding a dense class of functions for which pointwise convergence holds;
- proving an appropriate maximal inequality for the related averaging operators, which implies that the set of functions enjoying pointwise convergence is closed.

What is transferred between systems is this maximal inequality, with the most remarkable example being the Hardy–Littlewood maximal inequality for the one-sided averaging operators on $\mathbb{Z}$, which, when combined with the mean ergodic theorem of von Neumann [21], can be used to deduce the celebrated pointwise ergodic theorem of Birkhoff [2].

In the modern approach, however, maximal estimates are replaced by some stronger ones, where variation of a sequence of numbers rather than the maximum is controlled, thanks to which pointwise convergence can be established for all functions directly. This idea was crucial in a breakthrough series of articles by Bourgain [3–5] who obtained the pointwise ergodic theorem for operators averaging along orbits with times determined by polynomials. Very recently, Krause, Mirek, and Tao [14] were able to show a similar result for certain bilinear operators, also using the ideas of [7] to reduce the problem to studying canonical systems. These results are major steps toward confirming the general conjecture about pointwise convergence of multilinear averages taken along polynomial orbits, see Conjecture 1.2, promoted by Furstenberg and stated by Bergelson and Leibman in [1].

In this note, we shall show that the Calderón transference principle is valid for a broad spectrum of scenarios regarding both the types of underlying dynamical systems and the types of estimates to be transferred, see Theorem 1.6. In particular, all the settings considered in Conjecture 1.2 are captured. As we shall prove, the whole process can be carried out without increasing the optimal constants in the studied inequalities, see equation (1.7).

In view of equation (1.7), the following dichotomy arises naturally. For a given dynamical system, the best constant in the studied inequality is either equal or strictly smaller than the best constant in the same inequality for the associated canonical system. To illustrate the importance of this observation, we shall use equation (1.7) to classify for which systems equality or strict inequality regarding the two constants occurs, in the case of weak type $(1, 1)$ and strong type (p, p) maximal inequalities for centered, uncentered, or one-sided averages taken along linear orbits determined by a single ergodic transformation, see Theorem 1.9.

1.2. *Calderón transference.* Let $\mathbb{X}$ be a *measure-preserving dynamical system*, that is, a quadruple $(X, \mathcal{B}, \mu, \mathcal{T})$, where $(X, \mathcal{B}, \mu)$ is a non-trivial σ-finite measure space and $\mathcal{T} = (T_1, \ldots, T_d), d \in \mathbb{N}$, is a family of *measure-preserving transformations* on X. (Here, by 'non-trivial', we mean that there exists a subset of X with finite and strictly positive measure.) The latter means that $T_i : X \to X$ are measurable and $\mu(T_i^{-1}(E)) = \mu(E)$ for all $E \in \mathcal{B}$. Throughout the paper, we assume that T_i are invertible and commute with each other.

Next, for $m, k \in \mathbb{N}$, let $\mathcal{P} = (P_{1,1}, \ldots, P_{d,m})$ be a family of k-variate polynomials $P_{i,j}$ such that $P_{i,j}(\mathbb{Z}^k) \subseteq \mathbb{Z}^k$. The associated m-linear averaging operators are defined by

$$A_N^{\mathcal{P}} f(x) := \mathbb{E}_{n \in [N]^k} \prod_{j \in [m]} f_j(T_1^{P_{1,j}(n)} \cdots T_d^{P_{d,j}(n)} x), \quad x \in X. \tag{1.1}$$

Here $f = (f_1, \ldots, f_m)$ is an m-tuple of μ-measurable functions $f_j : X \to \mathbb{R}$, by $[l]$, we mean $\{1, \ldots, l\}$, and $\mathbb{E}_{y \in Y} g(y)$ is the expected value of g, where the expectation is taken with respect to the discrete uniform distribution over the indicated finite set Y.

Regarding the operators in equation (1.1), the following conjecture was posed in [1].

Conjecture 1.2. Fix $d, m, k \in \mathbb{N}$ and let $\mathbb{X}, \mathcal{P}$ be as before. Then, for the operators in equation (1.1),

$$\lim_{N \to \infty} A_N^{\mathcal{P}} f(x)$$

exists μ-almost everywhere for each m-tuple $f = (f_1, \ldots, f_m)$ with $f_j \in L^{\infty}(\mathbb{X})$, $j \in [m]$.

As mentioned before, the conjecture was verified in some specific cases and the proofs usually required some quantitative knowledge about the behavior of $(A_N^{\mathcal{P}} f(x))_{N \in \mathbb{N}}$. It is very likely that the general case, if true, will be proved in a similar fashion.

With this in mind, we introduce an abstract map $\mathcal{O} \colon \mathbb{R}^{\mathbb{N}} \to [0, \infty]$ whose arguments are sequences of real numbers. The only assumption imposed on $\mathcal{O}$ will be that it is approximated from below by a sequence of measurable mappings $(\mathcal{O}_K)_{K \in \mathbb{N}}$ depending on finitely many coordinates. Precisely, for each $K \in \mathbb{N}$, let $\mathcal{O}_K \colon \mathbb{R}^K \to [0, \infty)$ be measurable with respect to the standard topology on $\mathbb{R}^K$. We assume that $\mathcal{O}_K$ can be chosen so that

$$\mathcal{O}(a_n : n \in \mathbb{N}) \geq \mathcal{O}_K(a_n : n \in [K]), \quad K \in \mathbb{N}, (a_n : n \in \mathbb{N}) \in \mathbb{R}^{\mathbb{N}} \tag{1.3}$$

and

$$\mathcal{O}(a_n : n \in \mathbb{N}) = \lim_{K \to \infty} \mathcal{O}_K(a_n : n \in [K]), \quad (a_n : n \in \mathbb{N}) \in \mathbb{R}^{\mathbb{N}}. \tag{1.4}$$

Given $p = (p_0, p_1, \ldots, p_m) \in (0, \infty)^{m+1}$ with $1/p_1 + \cdots + 1/p_m = 1/p_0$, we examine the inequalities

$$\|\mathcal{O}(A_N^{\mathcal{P}} f : N \in \mathbb{N})\|_{p_0} \leq C \prod_{j \in [m]} \|f_j\|_{p_j} \quad \text{and}$$

$$\|\mathcal{O}(A_N^{\mathcal{P}} f : N \in \mathbb{N})\|_{p_0, \infty} \leq C \prod_{j \in [m]} \|f_j\|_{p_j}, \tag{1.5}$$

where by $\mathcal{O}(g_n : n \in \mathbb{N})$, we mean the function $x \mapsto \mathcal{O}(g_n(x) : n \in \mathbb{N})$, while $\|g\|_q$ and $\|g\|_{q,\infty}$ stand for the usual Lebesgue and weak Lebesgue q-quasinorms of g with respect to μ. More precisely, given $g \colon X \to \mathbb{R}$ and $q \in (0, \infty)$, we have

$$\|g\|_q := \left(\int_X |g(x)|^q \, d\mu(x) \right)^{1/q} \quad \text{and} \quad \|g\|_{q,\infty} := \sup_{\lambda \in (0,\infty)} \lambda \mu(\{x \in X : |g(x)| \geq \lambda\})^{1/q}.$$

We write respectively $\mathcal{C}_{\mathcal{O}}^{\mathcal{P}}(\mathbb{X}, p, s)$ and $\mathcal{C}_{\mathcal{O}}^{\mathcal{P}}(\mathbb{X}, p, w)$ for the smallest constants $C \in [0, \infty]$ such that these inequalities hold true for all m-tuples $f \in L^{p_1}(\mathbb{X}) \times \cdots \times L^{p_m}(\mathbb{X})$.

While reading, one can think of a model case, where $\mathcal{O}$ is the supremum norm

$$\mathcal{M}(a_n : n \in \mathbb{N}) := \sup_{n \in \mathbb{N}} |a_n|.$$

Notice that $\mathcal{M}$ is approximated from below by $\mathcal{M}_K(a_n : n \in [K]) := \max_{n \in [K]} |a_n|$. Another important example is the r-variation seminorm, $r \in [1, \infty)$, given by

$$\mathcal{V}^r(a_n : n \in \mathbb{N}) := \sup_{J \in \mathbb{N}} \sup_{n_0 < n_1 < \cdots < n_J} \left(\sum_{j \in [J]} |a_{n_j} - a_{n_{j-1}}|^r \right)^{1/r}$$

approximated by its truncated versions $\mathcal{V}_K^r$, where we additionally demand $n_J \leq K$.

Let us observe that $\mathcal{M}$ and $\mathcal{V}^r$ control respectively the size and variability of the input values. This is why estimates for $\mathcal{V}^r$ can be used to prove pointwise convergence directly, while maximal inequalities are only effective in estimating error terms, therefore requiring a dense class of functions for which pointwise convergence is known *a priori*.

In practice, estimates for $\mathcal{V}^r$ are more difficult to prove than their counterparts for $\mathcal{M}$. Moreover, quite often they hold only in a limited range. For example, sequences of expected value operators coming from martingales enjoy variation estimates for $r \in (2, \infty)$ but not for $r \in [1, 2]$, see [15, 18]. For more details and other important examples related to oscillation or jump inequalities, we refer the reader to [12, 17, 19].

Regarding the underlying system $\mathbb{X}$, our reference point is the d-dimensional canonical system $\mathbb{X}_d = (\mathbb{Z}^d, 2^{\mathbb{Z}^d}, \#_d, \mathcal{T}_d)$, that is, $\mathbb{Z}^d$ equipped with the σ-algebra of all subsets, counting measure, and the family $\mathcal{T}_d = (\mathcal{T}_{d,1}, \ldots, \mathcal{T}_{d,d})$ of d independent shifts defined by

$$\mathcal{T}_{d,i}(l_1, \ldots, l_{i-1}, l_i, l_{i+1}, \ldots l_d) := (l_1, \ldots, l_{i-1}, l_i + 1, l_{i+1}, \ldots l_d).$$

We shall show the following result relating the actions of $\mathcal{O}$ on $\mathbb{X}$ and $\mathbb{X}_d$ to each other.

THEOREM 1.6. (Calderón transference) *Fix $d, m, k \in \mathbb{N}$ and let $\mathbb{X}, \mathcal{P}, p$ be as before. Then,*

$$\mathcal{C}_{\mathcal{O}}^{\mathcal{P}}(\mathbb{X}, p, s) \le \mathcal{C}_{\mathcal{O}}^{\mathcal{P}}(\mathbb{X}_d, p, s) \quad and \quad \mathcal{C}_{\mathcal{O}}^{\mathcal{P}}(\mathbb{X}, p, w) \le \mathcal{C}_{\mathcal{O}}^{\mathcal{P}}(\mathbb{X}_d, p, w) \qquad (1.7)$$

hold for all $\mathcal{O}$ admitting equations (1.3) and (1.4), where $\mathbb{X}_d$ is the d-dimensional canonical system.

In particular, regarding maximal, variation, oscillation, or jump inequalities, it is enough to deal with canonical systems.

1.3. *Dichotomy.* Theorem 1.6 motivates the following question.

When does the equality take place in equation (1.7)?

We shall show how to deal with such a kind of problem by solving it in a very particular relatively easy case. In this subsection, we take $d = m = k = 1$ and $P(n) = n$. Moreover, we assume that T is *ergodic*, which means that $T^{-1}(E) = E$ implies $\mu(E) = 0$ or $\mu(X \setminus E) = 0$.

In this context, we introduce the one-sided maximal operator T_*^{os} by using the formula

$$T_*^{\mathrm{os}} f(x) := \sup_{N \in \mathbb{N} \cup \{0\}} \left| \frac{1}{N+1} \sum_{n=0}^{N} f(T^n x) \right|, \quad x \in X.$$

Similarly, we define the two-sided maximal operators, centered T_*^{c} and uncentered T_*^{u}, by

$$T_*^{\mathrm{c}} f(x) := \sup_{N \in \mathbb{N} \cup \{0\}} \left| \frac{1}{2N+1} \sum_{n=-N}^{N} f(T^n x) \right| \quad \text{and}$$

$$T_*^{\mathrm{u}} f(x) := \sup_{\underline{r}, \overline{r} \in \mathbb{N} \cup \{0\}} \left| \frac{1}{\underline{r} + \overline{r} + 1} \sum_{n=-\underline{r}}^{\overline{r}} f(T^n x) \right|.$$

Regarding these operators, it is natural to ask about the weak type $(1, 1)$ and strong type (p, p) inequalities for $p \in (1, \infty]$. Thus, we put $\mathcal{C}^{\mathrm{os}}(\mathbb{X}, 1) := \mathcal{C}^{P}_{T^{\mathrm{os}}_{*}}(\mathbb{X}, (1, 1), \mathrm{w})$ and similarly $\mathcal{C}^{\mathrm{os}}(\mathbb{X}, p) := \mathcal{C}^{P}_{T^{\mathrm{os}}_{*}}(\mathbb{X}, (p, p), \mathrm{s})$ for each $p \in (1, \infty)$, while $\mathcal{C}^{\mathrm{os}}(\mathbb{X}, \infty)$ will stand for the best constant C in the inequality $\|T^{\mathrm{os}}_{*} f\|_{\infty} \leq C \|f\|_{\infty}$, where

$$\|g\|_{\infty} := \inf\{\lambda \in [0, \infty] : |g(x)| \leq \lambda \text{ for } \mu\text{-almost every } x \in X\}.$$

We define $\mathcal{C}^{\mathrm{c}}(\mathbb{X}, p)$ and $\mathcal{C}^{\mathrm{u}}(\mathbb{X}, p)$, $p \in [1, \infty]$, in the same manner, replacing T^{os}_{*} with T^{c}_{*} and T^{u}_{*}. In view of Theorem 1.6, properties of the Hardy–Littlewood maximal operators on $\mathbb{Z}$, and obvious bounds for $p = \infty$, all these quantities are finite regardless of $\mathbb{X}$.

Given a non-trivial ergodic system $\mathbb{X}$, we say that X consists of finitely many atoms if it splits into disjoint measurable sets $X_0, X_1, \ldots, X_L$, $L \in \mathbb{N}$, such that $\mu(X_0) = 0$, $T(X_l) \subseteq X_{l+1}$ for $l \in [L-1]$, $T(X_L) \subseteq X_1$, and none of $X_l, l \in [L]$, can be split further into two disjoint sets of non-zero measure. Observe that then $\mu(X_1) = \cdots = \mu(X_L) \in (0, \infty)$.

The following well-known fact will be useful later on.

Fact 1.8. (Kakutani–Rokhlin lemma) Let $\mathbb{X}$ be a non-trivial ergodic system. Then exactly one of the following two possibilities holds:

(A) the set X consists of finitely many atoms;
(B) for each $L \in \mathbb{N}$, there exists $E_L \in \mathcal{B}$ such that $\mu(E_L) \in (0, \infty)$ and the sets $T^{-l}(E_L), l \in [L]$, are disjoint.

We are ready to formulate our second main result.

THEOREM 1.9. (Sharp constants dichotomy) *Let $\mathbb{X}$ be as before. Then for $p \in \{1, \infty\}$, we have $\mathcal{C}^{\mathrm{os}}(\mathbb{X}, p) = \mathcal{C}^{\mathrm{os}}(\mathbb{X}_1, p)$, while for $p \in (1, \infty)$, the following dichotomy occurs:*

- *if case (A) from Fact 1.8 holds, then $\mathcal{C}^{\mathrm{os}}(\mathbb{X}, p) < \mathcal{C}^{\mathrm{os}}(\mathbb{X}_1, p)$;*
- *if case (B) from Fact 1.8 holds, then $\mathcal{C}^{\mathrm{os}}(\mathbb{X}, p) = \mathcal{C}^{\mathrm{os}}(\mathbb{X}_1, p)$.*

Moreover, we have $\mathcal{C}^{\mathrm{c}}(\mathbb{X}, \infty) = \mathcal{C}^{\mathrm{c}}(\mathbb{X}_1, \infty)$ and $\mathcal{C}^{\mathrm{u}}(\mathbb{X}, \infty) = \mathcal{C}^{\mathrm{u}}(\mathbb{X}_1, \infty)$, while for $p \in [1, \infty)$, the following dichotomy occurs:

- *if case (A) from Fact 1.8 holds, then $\mathcal{C}^{\mathrm{c}}(\mathbb{X}, p) < \mathcal{C}^{\mathrm{c}}(\mathbb{X}_1, p)$ and $\mathcal{C}^{\mathrm{u}}(\mathbb{X}, p) < \mathcal{C}^{\mathrm{u}}(\mathbb{X}_1, p)$;*
- *if case (B) from Fact 1.8 holds, then $\mathcal{C}^{\mathrm{c}}(\mathbb{X}, p) = \mathcal{C}^{\mathrm{c}}(\mathbb{X}_1, p)$ and $\mathcal{C}^{\mathrm{u}}(\mathbb{X}, p) = \mathcal{C}^{\mathrm{u}}(\mathbb{X}_1, p)$.*

This result has interesting consequences because some of the constants on $\mathbb{X}_1$ are known.

COROLLARY 1.10. *Let $\mathbb{X}$ be as before and assume that case (B) from Fact 1.8 holds. Then:*

- *for $p = 1$, we have $\mathcal{C}^{\mathrm{os}}(\mathbb{X}, 1) = 1$, $\mathcal{C}^{\mathrm{u}}(\mathbb{X}, 1) = 2$, and $\mathcal{C}^{\mathrm{c}}(\mathbb{X}, 1) = (11 + \sqrt{61})/12$;*
- *for $p \in (1, \infty)$ and the uncentered operator, we have $\mathcal{C}^{\mathrm{u}}(\mathbb{X}, p) = c_p$, where c_p is the unique positive solution of the equation $(p-1)x^p - px^{p-1} - 1 = 0$;*
- *for $p = \infty$, we have $\mathcal{C}^{\mathrm{os}}(\mathbb{X}, \infty) = \mathcal{C}^{\mathrm{u}}(\mathbb{X}, \infty) = \mathcal{C}^{\mathrm{c}}(\mathbb{X}, \infty) = 1$.*

Indeed, the equality $\mathcal{C}^{\mathrm{os}}(\mathbb{X}_1, 1) = 1$ is well known, see Lemma 3.1 for an easy proof. Next, $\mathcal{C}^{\mathrm{u}}(\mathbb{X}_1, 1) = 2$ because $\mathcal{C}^{\mathrm{u}}(\mathbb{X}_1, 1) \leq 2$ follows from the classical covering lemma on $\mathbb{Z}$

with overlap 2, while $\mathcal{C}^{\mathrm{u}}(\mathbb{X}_1, 1) \geq 2$ can be seen by taking $\tilde{f} = \mathbb{1}_{\{0\}}$ and letting $\lambda \to 0$. However, $\mathcal{C}^{\mathrm{c}}(\mathbb{X}_1, 1) = (11 + \sqrt{61})/12$ follows by combining the famous result of Melas [16], where the best constant in the weak type $(1, 1)$ inequality for the centered maximal operator on $\mathbb{R}$ is obtained, with [13, Theorem 1], where the best constants in discrete and continuous inequalities are compared to each other (in our case, both constants are equal). For $p = \infty$, the claim is trivial, and for $p \in (1, \infty)$, just like before, we combine the result on $\mathbb{R}$ of Grafakos and Montgomery-Smith [9] with the following transference principle.

PROPOSITION 1.11. *Fix $p \in (1, \infty)$ and let $\mathbb{X}$ be as before. Then $\mathcal{C}^{\mathrm{u}}(\mathbb{X}, p) = c_p$, where c_p is the best constant in the strong type (p, p) inequality for $\mathcal{M}^{\mathrm{u}}_{\mathbb{R}}$, the uncentered Hardy–Littlewood maximal operator $\mathbb{R}$.*

Proof. The inequality $\mathcal{C}^{\mathrm{u}}(\mathbb{X}, p) \geq c_p$ follows by repeating the arguments used to prove [13, equation (1.3)]. We only sketch the proof. Let $f_{\mathrm{cont}} \colon \mathbb{R} \to [0, \infty)$ be a smooth function for which c_p is almost attained. Using dilations, we can assume that f_{cont} is very slowly varying. Define $f_{\mathrm{dis}} \colon \mathbb{Z} \to [0, \infty)$ by sampling $f_{\mathrm{dis}}(l) := f_{\mathrm{cont}}(l)$. Then the constant in the maximal inequality stated for f_{dis} will be almost the same as that corresponding to f_{cont}.

However, $\mathcal{C}^{\mathrm{u}}(\mathbb{X}, p) \leq c_p$ follows easily, since for any $f_{\mathrm{dis}} \colon \mathbb{Z} \to [0, \infty)$, we have

$$\frac{\|T^{\mathrm{u}}_* f_{\mathrm{dis}}\|_p}{\|f_{\mathrm{dis}}\|_p} \leq \frac{\|\mathcal{M}^{\mathrm{u}}_{\mathbb{R}} f_{\mathrm{cont}}\|_p}{\|f_{\mathrm{cont}}\|_p},$$

where $f_{\mathrm{cont}} \colon \mathbb{R} \to [0, \infty)$ is given by $f_{\mathrm{cont}}(x) := f_{\mathrm{dis}}(l)$ for $x \in [l, l+1), l \in \mathbb{N}$. Indeed, this inequality can be easily seen, since for each $x \in [l, l+1)$, one has $\mathcal{M}^{\mathrm{u}}_{\mathbb{R}} f_{\mathrm{cont}}(x) \geq T^{\mathrm{u}}_* f_{\mathrm{dis}}(l)$ (it is important here that we are working with operators of uncentered type). $\square$

Similar arguments give an analog of Proposition 1.11 for the one-sided operator. For the centered operator, the situation is different and only the lower bound for $\mathcal{C}^{\mathrm{c}}(\mathbb{X}, p)$ can be easily transferred from the continuous setting to the discrete one, cf. [13, Theorem 1].

1.4. *Ergodicity.* The canonical system $\mathbb{X}_d$ can be viewed as a topological group with Haar measure and transformations $\mathcal{T}_{d,i}$ being translations by group elements, say $g_{d,i}$. Since the subgroup G generated by $g_{d,i}$ is just $\mathbb{Z}^d$, the canonical system is ergodic, that is, there are no non-trivial G-invariant sets $E \subseteq \mathbb{Z}^d$. We take this opportunity to highlight the importance of ergodicity in this context.

Let us specify our measure space $(X, \mathcal{B}, \mu)$ to be $[1, 0)$ with the usual Borel sets and Lebesgue measure. We consider the group $\mathrm{Aut}(X, \mathcal{B}, \mu)$ of bimeasurable automorphisms of that space, that is, invertible maps $T \colon X \to X$ such that T, T^{-1} are measure preserving. In fact, elements of $\mathrm{Aut}(X, \mathcal{B}, \mu)$ are equivalence classes, with $T \sim T'$ if and only if $T(x) = T'(x)$ holds μ-almost everywhere, and we shall refer to their representatives.

We say that $(T_n)_{n \in \mathbb{N}}$ *converges weakly* to T if $\lim_{n \to \infty} \mu(T_n(E) \Delta T(E)) = 0$ for all $E \in \mathcal{B}$ with Δ being the symmetric difference symbol. Also, if $\mu(\{x \in [0, 1) : T^n(x) = x\}) = 0$ for all $n \in \mathbb{N}$, then T is called *aperiodic*. Following Halmos [10], we notice that for aperiodic T, its conjugacy orbit $\{S^{-1} T S : S \in \mathrm{Aut}(X, \mathcal{B}, \mu)\}$ is weakly

dense. If T is ergodic instead, then so are all $S^{-1}TS$, while properties of μ imply that T is aperiodic as well. (Note that ergodic systems satisfying case (A) from Fact 1.8 lack aperiodicity, so referring to μ is necessary.) In view of ergodicity of $T_\pi(x) := x + \pi$ mod 1, say, ergodic automorphisms are weakly dense. This fact alone suggests that many problems are reducible to the case of ergodic transformations.

Yet another relevant result is the so-called Conze principle [8]. Its original formulation is as follows. Given a sequence $(i_n)_{n \in \mathbb{N}}$ of integers, for each $S \in \mathrm{Aut}(X, \mathcal{B}, \mu)$, the sharp constant in the weak type $(1, 1)$ inequality for $S_* f(x) := \sup_{N \in \mathbb{N}} |(1/N) \sum_{n=1}^{N} f(S^{i_n} x)|$ is not larger than its counterpart for T_* defined analogously for any aperiodic $T \in \mathrm{Aut}(X, \mathcal{B}, \mu)$. The proof uses weak density of the conjugacy orbit of T, and the principle can be generalized in many ways. Again, by properties of μ, the same is true for any ergodic T instead.

The last related issue we would like to discuss is the Stein maximal principle [20]. As we have seen, pointwise convergence results are often consequences of weak type inequalities, while proving the latter can be reduced to the ergodic case whenever the Conze principle is available. Stein's result, saying that in some cases, pointwise convergence is equivalent to the associated weak type inequality, makes these observations more fundamental.

Let this time X be a homogeneous space of a given compact group G. By this, we mean that, in particular, the structure of X is inherited from G, there is a unique normalized G-invariant measure μ on X, and G acts transitively on X. For example, X may be the n-dimensional sphere with G consisting of all rotations of X. Transitivity implies that there are no non-trivial G-invariant sets $E \subseteq X$, and ergodicity indeed is crucial here.

Now Stein's result reads as follows. For each $p \in [1, 2]$ and each given sequence $(T_n)_{n \in \mathbb{N}}$ of operators bounded on $L^p(X, \mu)$ and commuting with all transformations determined by the actions of $g \in G$ on X, if $\lim_{n \to \infty} T_n f(x)$ exists almost everywhere for all $f \in L^p(X, \mu)$, then $T_* f(x) := \sup_{n \in \mathbb{N}} |T_n f(x)|$ satisfies the weak type (p, p) inequality.

Let us check that ergodicity cannot be dropped. To see this, suppose that X consists of disjoint G-invariant sets $X_l, l \in \mathbb{N}$, admitting Stein's principle and such that $\mu(X_l) = 2^{-l}$. Then pointwise convergence gives weak type inequalities on all X_l separately but the same for X may be false, as the best constants may tend to infinity with l. This happens when $T_n f(x)$ equals $nf(x)$ for $x \in X_n$ and 0 otherwise.

For another instructive example, take $X_l := \{(l, 0), (l, 1), \ldots, (l, l)\}$ with equal masses $2^{-l}/(l+1)$, and set $G := \mathbb{Z}_2 \times \mathbb{Z}_3 \times \cdots$ with the action of $g = (g_1, g_2, \ldots) \in G$ on X given by $\tau_g(l, i) := (l, i + g_l \bmod l+1)$. Then for $g_* := (1, 1, \ldots) \in G$ and $T_n f(x) := (1/n) \sum_{i=1}^{n} f(\tau_{g_*}^{i^2} x)$, we see that $\lim_{n \to \infty} T_n f(x)$ exists everywhere for all $f \in L^1(X, \mu)$, while there is no weak type $(1, 1)$ inequality in view of the famous result by Buczolich and Mauldin [6].

2. *Proof of Theorem 1.6*

Let us now prove Theorem 1.6.

Proof of Theorem 1.6. We only prove the first inequality and the second one can proved by using very similar arguments.

Let $f = (f_1, \ldots, f_m)$ with $f_j \in L^{p_j}(\mathbb{X})$. For each $K \in \mathbb{N}$, we define

$$F_j^K(x, l) := f_j(T_1^{-KR_K+l_1} \cdots T_d^{-KR_K+l_d} x) \cdot \mathbb{1}_{\{l \in [KR_K]^d\}}, \quad x \in X, \ l \in \mathbb{Z}^d,$$

with $R_K \in \mathbb{N}$ so large that $|P_{i,j}(l)| \leq R_K$ for all i, j when $\|l\|_\infty \leq K$. Then,

$$\|F_j^K\|_{p_j} = (KR_K)^{d/p_j} \|f_j\|_{p_j}$$

by the Fubini–Tonelli theorem, since T_i are measure preserving. The first norm above is with respect to the product of μ and counting measure. Set $F^K = (F_1^K, \ldots, F_m^K)$. We have

$$\mathcal{O}_K(A_N^{\mathcal{P}} F^K(x, l) : N \in [K]) = \mathcal{O}_K(A_N^{\mathcal{P}} f(T_1^{-KR_K+l_1} \cdots T_d^{-KR_K+l_d} x) : N \in [K])$$

for all $x \in X$ and $l \in \{R_K + 1, \ldots, (K-1)R_K\}^d$ when $K \geq 3$. Hence,

$$\|\mathcal{O}_K(A_N^{\mathcal{P}} F^K : N \in [K])\|_{p_0} \geq ((K-2)R_K)^{d/p_0} \|\mathcal{O}_K(A_N^{\mathcal{P}} f : N \in [K])\|_{p_0}$$

by the Fubini–Tonelli theorem, since T_i are measure preserving. Letting $K \to \infty$ gives

$$\|\mathcal{O}(A_N^{\mathcal{P}} f : N \in \mathbb{N})\|_{p_0} \leq \mathcal{C}_{\mathcal{O}}^{\mathcal{P}}(\mathbb{X}_d, p, s) \lim_{K \to \infty} ((K-2)R_K)^{-d/p_0} \prod_{j \in [m]} \|F_j^K\|_{p_j}$$

by Fatou's lemma, equation (1.3), equation (1.4), and the definition of $\mathcal{C}_{\mathcal{O}}^{\mathcal{P}}(\mathbb{X}_d, p, s)$. The limit is equal to

$$\lim_{K \to \infty} ((K-2)R_K)^{-d/p_0} (KR_K)^{d/p_1+\cdots+d/p_m} \prod_{j \in [m]} \|f_j\|_{p_j} = \prod_{j \in [m]} \|f_j\|_{p_j}$$

thanks to the Hölder exponent hypothesis. This completes the proof. □

Once we have seen the proof of Theorem 1.6, several remarks are in order.

- Apart from those specified in equation (1.5), other types of estimates could be considered in Theorem 1.6. One of the key factors is good scaling, which in our case was assured by assuming $1/p_1 + \cdots + 1/p_m = 1/p_0$.
- Properties of $\mathbb{X}_d$ resemble the structure of the group of transformations generated by T_i. Dropping commutativity is possible but it leads to more complicated canonical systems, see [11] for the case of nilpotent groups of step 2.
- Instead of $A_N^{\mathcal{P}}$, one can consider different operators, for example, the averaging operators used to define T_*^{os}, T_*^c, T_*^u from Theorem 1.9, see Proposition 2.1.

PROPOSITION 2.1. *Under the assumptions of Theorem 1.9, for each $p \in [1, \infty]$, we have $\mathcal{C}^{os}(\mathbb{X}, p) \leq \mathcal{C}^{os}(\mathbb{X}_1, p), \mathcal{C}^c(\mathbb{X}, p) \leq \mathcal{C}^c(\mathbb{X}_1, p),$ and $\mathcal{C}^u(\mathbb{X}, p) \leq \mathcal{C}^u(\mathbb{X}_1, p)$.*

Proof. We repeat the proof of Theorem 1.6. □

3. *Proof of Theorem 1.9*
The proof of Theorem 1.9 is divided into several lemmas. Throughout this section, $d = m = k = 1$, T is ergodic, and $P(n) = n$.

LEMMA 3.1. (Sharp constant equals 1) *We have* $\mathcal{C}^{\mathrm{os}}(\mathbb{X}_1, 1) = \mathcal{C}^{\mathrm{os}}(\mathbb{X}_1, \infty) = \mathcal{C}^{\mathrm{c}}(\mathbb{X}_1, \infty) = \mathcal{C}^{\mathrm{u}}(\mathbb{X}_1, \infty) = 1$. *Consequently,* $\mathcal{C}^{\mathrm{os}}(\mathbb{X}, 1) = \mathcal{C}^{\mathrm{os}}(\mathbb{X}_1, 1)$, $\mathcal{C}^{\mathrm{os}}(\mathbb{X}, \infty) = \mathcal{C}^{\mathrm{os}}(\mathbb{X}_1, \infty)$, $\mathcal{C}^{\mathrm{c}}(\mathbb{X}, \infty) = \mathcal{C}^{\mathrm{c}}(\mathbb{X}_1, \infty)$, *and* $\mathcal{C}^{\mathrm{u}}(\mathbb{X}, \infty) = \mathcal{C}^{\mathrm{u}}(\mathbb{X}_1, \infty)$.

Proof. We only need to prove the first part and the rest of the claim follows directly from Proposition 2.1 and the fact that $|f| \leq \min\{T_*^{\mathrm{os}} f, T_*^{\mathrm{c}} f, T_*^{\mathrm{u}} f\}$. The case $p = \infty$ is obvious. Also, $\mathcal{C}^{\mathrm{os}}(\mathbb{X}_1, 1) = 1$ is well known but we include the proof for the sake of completeness.

We shall show that $\mathcal{C}^{\mathrm{os}}(\mathbb{X}_1, 1) \leq 1$ and the remaining inequality $\mathcal{C}^{\mathrm{os}}(\mathbb{X}_1, 1) \geq 1$ is obvious. Let $f \in L^1(\mathbb{X}_1)$ be non-negative and assume that $\lambda \in (0, \infty)$ is such that the set

$$E_\lambda^{\mathrm{os}}(f) := \{l \in \mathbb{Z} : |T_*^{\mathrm{os}} f(l)| \geq \lambda\}$$

is non-empty. Thus, $E_\lambda^{\mathrm{os}}(f) = \{l_1, \ldots, l_J\}$ for some integers $l_1 < \cdots < l_J$, $J \in \mathbb{N}$. For each $j \in [J]$, we let N_j be the smallest number $N \in \mathbb{N} \cup \{0\}$ such that $1/(N+1) \sum_{n=0}^{N} f(l_j + n) \geq \lambda$. We set $B_j := \{l_j, \ldots, l_j + N_j\}$ and observe that $E_\lambda^{\mathrm{os}}(f) \subseteq \bigcup_{j=1}^{J} B_j$. Next, define $D_1 := B_1$ and $D_j := B_j \setminus \bigcup_{i=1}^{j-1} B_i$ for $1 \neq j \in [J]$. Then D_j are disjoint, we have $\bigcup_{j=1}^{J} B_j = \bigcup_{j=1}^{J} D_j$, and for each j, either $D_j = \emptyset$ or $D_j = \{l_j + r_j, \ldots, l_j + N_j\}$ for some $r_j \in [N_j] \cup \{0\}$. In the latter case, by the definition of N_j, we have $\sum_{l \in D_j} f(l) \geq \lambda |D_j|$. Consequently,

$$\lambda |E_\lambda^{\mathrm{os}}(f)| \leq \sum_{j \in [J]} \lambda |D_j| \leq \sum_{j \in [J]} \sum_{l \in D_j} f(l) \leq \|f\|_1,$$

which justifies $\mathcal{C}^{\mathrm{os}}(\mathbb{X}_1, 1) \leq 1$. $\qquad\square$

Let us now consider the case $p \in (1, \infty)$ for atomic systems.

LEMMA 3.2. (Case (A), $p \in (1, \infty)$) *Assume that case* (A) *from Fact 1.8 holds. Then for each* $p \in (1, \infty)$, *we have* $\mathcal{C}^{\mathrm{os}}(\mathbb{X}, p) < \mathcal{C}^{\mathrm{os}}(\mathbb{X}_1, p)$, $\mathcal{C}^{\mathrm{c}}(\mathbb{X}, p) < \mathcal{C}^{\mathrm{c}}(\mathbb{X}_1, p)$, *and* $\mathcal{C}^{\mathrm{u}}(\mathbb{X}, p) < \mathcal{C}^{\mathrm{u}}(\mathbb{X}_1, p)$.

Proof. We only prove $\mathcal{C}^{\mathrm{os}}(\mathbb{X}, p) < \mathcal{C}^{\mathrm{os}}(\mathbb{X}_1, p)$ and the remaining inequalities can be proved by using very similar arguments.

It suffices to show that for each fixed $L \in \mathbb{N}$ and $p \in (1, \infty)$, we have $\mathcal{C}^{\mathrm{os}}(\mathbb{X}_{[L]}, p) < \mathcal{C}^{\mathrm{os}}(\mathbb{X}_1, p)$, where $\mathbb{X}_{[L]} = ([L], 2^{[L]}, \#_{[L]}, T_{[L]})$ is the finite shift system with L elements, that is, $T_{[L]}(l) := l + 1$ for $l \in [L-1]$ and $T_{[L]}(L) := 1$.

Given $\epsilon \in (0, 1)$, let $f : [L] \to [0, \infty)$ be such that $\|T_*^{\mathrm{os}} f\|_p \geq (1 - \epsilon) \mathcal{C}^{\mathrm{os}}(\mathbb{X}_{[L]}, p) \|f\|_p$. Then for a large parameter $R \in \mathbb{N}$, we define $F : \mathbb{Z} \to [0, \infty)$ by

$$F(l) := \begin{cases} f(l \bmod L) & \text{for } l \in [RL], \\ 0 & \text{otherwise,} \end{cases}$$

where we identify $f(0)$ with $f(L)$. Note that $\|F\|_p = R^{1/p} \|f\|_p$. Moreover, if R is sufficiently large (with respect to L, p, and ϵ), then

$$\|T_*^{\mathrm{os}} F \cdot \mathbb{1}_{[RL]}\|_p \geq (1 - \epsilon)^2 \mathcal{C}^{\mathrm{os}}(\mathbb{X}_{[L]}, p) \|F\|_p.$$

For each $l \in \{-RL + 1, \ldots, 0\}$, by using Hölder's inequality, we obtain

$$T_*^{\mathrm{os}} F(l) \geq \frac{1}{2RL} \sum_{n \in [RL]} F(n) = \frac{1}{2L} \sum_{n \in [L]} f(n) \geq \frac{\|T_*^{\mathrm{os}} f\|_\infty}{2L} \geq \frac{\|T_*^{\mathrm{os}} f\|_p}{2L^{1+1/p}}$$

$$\geq \frac{(1 - \epsilon) \mathcal{C}^{\mathrm{os}}(\mathbb{X}_{[L]}, p)}{2L^{1+1/p} R^{1/p}} \|F\|_p.$$

Consequently,

$$\|T_*^{\mathrm{os}} F \cdot \mathbb{1}_{\{-RL+1,\ldots,RL\}}\|_p^p \geq \left((1 - \epsilon)^{2p} + RL \cdot \frac{(1 - \epsilon)^p}{2^p L^{p+1} R} \right) \mathcal{C}^{\mathrm{os}}(\mathbb{X}_{[L]}, p)^p \|F\|_p^p$$

$$> \mathcal{C}^{\mathrm{os}}(\mathbb{X}_{[L]}, p)^p \|F\|_p^p,$$

provided that ϵ is sufficiently small (with respect to L and p). □

Next, we consider the case $p \in [1, \infty)$ for other systems.

LEMMA 3.3. (Case (B), $p \in [1, \infty)$) *Assume that case (B) from Fact 1.8 holds. Then for each $p \in [1, \infty)$, we have $\mathcal{C}^{\mathrm{os}}(\mathbb{X}, p) = \mathcal{C}^{\mathrm{os}}(\mathbb{X}_1, p)$, $\mathcal{C}^{\mathrm{c}}(\mathbb{X}, p) = \mathcal{C}^{\mathrm{c}}(\mathbb{X}_1, p)$, and $\mathcal{C}^{\mathrm{u}}(\mathbb{X}, p) = \mathcal{C}^{\mathrm{u}}(\mathbb{X}_1, p)$.*

Proof. We only prove $\mathcal{C}^{\mathrm{os}}(\mathbb{X}, p) = \mathcal{C}^{\mathrm{os}}(\mathbb{X}_1, p)$ for $p \in (1, \infty)$ and the remaining equalities or the case $p = 1$ can be verified by using very similar arguments.

Given $\epsilon \in (0, 1)$, one can find $L \in \mathbb{N}$ and $F \colon \mathbb{Z} \to [0, \infty)$ such that $F(l) = 0$ for each $l \in \mathbb{Z} \setminus [L]$ and

$$\|T_*^{\mathrm{os}} F \cdot \mathbb{1}_{[L]}\|_p \geq (1 - \epsilon) \mathcal{C}^{\mathrm{os}}(\mathbb{X}_1, p) \|F\|_p.$$

Choose $E_L \subseteq X$ as in case (B) from Fact 1.8, and define $f \in L^p(\mathbb{X})$ by

$$f(x) := \begin{cases} F(L + 1 - l) & \text{for } x \in T^{-l}(E_L), \ l \in [L], \\ 0 & \text{otherwise.} \end{cases}$$

Then it is easy to see that $\|T_*^{\mathrm{os}} f\|_p \geq (1 - \epsilon) \mathcal{C}^{\mathrm{os}}(\mathbb{X}_1, p) \|f\|_p.$ □

It remains to consider the case $p = 1$ for atomic systems. We only focus on T_*^{c} and T_*^{u}, since the claim for T_*^{os} has already been proven in Lemma 3.1.

LEMMA 3.4. (Case (A), $p = 1$, centered operator) *Assume that case (A) from Fact 1.8 holds. Then we have $\mathcal{C}^{\mathrm{c}}(\mathbb{X}, 1) < \mathcal{C}^{\mathrm{c}}(\mathbb{X}_1, 1)$.*

Proof. It suffices to show that for each fixed $L \in \mathbb{N}$, we have $\mathcal{C}^{\mathrm{c}}(\mathbb{X}_{[L]}, 1) < \mathcal{C}^{\mathrm{c}}(\mathbb{X}_1, 1)$, where $\mathbb{X}_{[L]}$ is as in Lemma 3.2. We follow the proof of [16, Theorem 3].

First, observe that there exists $f_0 \colon [L] \to [0, \infty)$ such that $\|T_*^{\mathrm{c}} f_0\|_{1,\infty} = \mathcal{C}^{\mathrm{c}}(\mathbb{X}_{[L]}, 1) \|f_0\|_1$. Indeed, let $(f_j)_{j \in \mathbb{N}}$ be a sequence of non-negative functions such that

$$\lim_{j \to \infty} \frac{\|T_*^{\mathrm{c}} f_j\|_{1,\infty}}{\|f_j\|_1} = \mathcal{C}^{\mathrm{c}}(\mathbb{X}_{[L]}, 1).$$

By applying the appropriate translation and scaling, we can assume that $f_j(1) = \|f_j\|_\infty = 1$ holds. Choose a subsequence $(j_i)_{i \in \mathbb{N}}$ such that $\lim_{i \to \infty} f_{j_i}(l)$ exists for

each $l \in [L]$. Then f_0 defined by $f_0(l) := \lim_{i \to \infty} f_{j_i}(l)$ is the function for which we are looking. Similarly, observe that there exists $\lambda \in (0, 1]$ such that $\lambda |E_\lambda^c(f_0)| = \mathcal{C}^c(\mathbb{X}_{[L]}, 1) \|f_0\|_1$. We define $f := f_0/\lambda$ and note that $f(l) \in [0, L]$ holds for each $l \in [L]$ (if $f(l) > L$, then $\|T_*^c f\|_{1,\infty} < \|f\|_1$, which leads to a contradiction). Below, we identify f with its L-periodic extension.

We have $E_1^c(f) = \{l_1, \ldots, l_J\}$ for some $J \in [L]$ and $l_1 < \cdots < l_J$. For each $j \in [J]$, we choose $N_j \in \mathbb{N} \cup \{0\}$ such that $f(l_j - N_j) + \cdots + f(l_j + N_j) \geq 2N_j + 1$. Consider the set

$$\mathcal{S} := \{(x_1, \ldots, x_L) \in [0, L]^L : x_{l_j - N_j} + \cdots + x_{l_j + N_j} \geq 2N_j + 1 \text{ for each } j \in [J]\}.$$

Notice that $(f(1), \ldots, f(L)) \in \mathcal{S}$. Consequently, $\mathcal{S}$ is a non-empty, compact, convex polyhedron contained in $\mathbb{R}^L$. We introduce a linear operator $\Lambda \colon \mathbb{R}^L \to \mathbb{R}$ defined by

$$\Lambda(x_1, \ldots, x_L) := x_1 + \cdots + x_L.$$

Then there exists a vertex of $\mathcal{S}$, say $(x_1^*, \ldots, x_L^*)$, such that

$$\Lambda(x_1^*, \ldots, x_L^*) = \min\{\Lambda(x_1, \ldots, x_L) : (x_1, \ldots, x_L) \in \mathcal{S}\} \leq \Lambda(f(1), \ldots, f(L)).$$

We observe that $(x_1^*, \ldots, x_L^*)$, as a vertex of $\mathcal{S}$, is the only solution of the linear system consisting of all those of the equations $x_1 = 0, \ldots, x_L = 0, x_1 = L, \ldots, x_L = L$, and $x_{l_1 - N_1} + \cdots + x_{l_1 + N_1} = 2N_1 + 1, \ldots, x_{l_J - N_J} + \cdots + x_{l_J + N_J} = 2N_J + 1$, which are satisfied with $(x_1^*, \ldots, x_L^*)$ in place of $(x_1, \ldots, x_L)$. Applying a standard argument from the theory of linear systems, we obtain that $x_l^* \in \mathbb{Q}$ for each $l \in [L]$.

Consider $f^* \colon [L] \to [0, \infty)$ defined by $f^*(l) := x_l^*$. Then we have $\|f^*\|_1 \leq \|f\|_1$ and $E_1^c(f^*) \supseteq E_1^c(f)$. Consequently, $|E_{f^*}| = \mathcal{C}^c(\mathbb{X}_{[L]}, 1) \|f^*\|_1$, which implies $\mathcal{C}^c(\mathbb{X}_{[L]}, 1) \in \mathbb{Q}$. By using Proposition 2.1 and the fact that $\mathcal{C}^c(\mathbb{X}_\mathbb{Z}, 1) = (11 + \sqrt{61})/12 \notin \mathbb{Q}$ (see [16] and [13, Theorem 1]), we conclude that $\mathcal{C}^c(\mathbb{X}_{[L]}, 1) < \mathcal{C}^c(\mathbb{X}_\mathbb{Z}, 1)$. □

LEMMA 3.5. (Case (A), $p = 1$, uncentered operator) *Assume that case (A) from Fact 1.8 holds. Then we have $\mathcal{C}^u(\mathbb{X}, 1) < \mathcal{C}^u(\mathbb{X}_1, 1)$.*

Proof. It suffices to show that for each fixed $L \in \mathbb{N}$, we have $\mathcal{C}^u(\mathbb{X}_{[L]}, 1) < \mathcal{C}^u(\mathbb{X}_1, 1)$, where $\mathbb{X}_{[L]}$ is as in Lemma 3.2. We also recall the well-known fact that $\mathcal{C}^u(\mathbb{X}_1, 1) = 2$.

Take $f \colon [L] \to [0, \infty)$ and $\lambda \in (0, \infty)$ such that $E_\lambda^u(f)$ is non-empty. Again, we identify f with its L-periodic extension. Then, $E_\lambda^u(f) = E_{\text{left}} \cup E_{\text{right}}$, where

$$E_{\text{left}} := \{l \in [L] : f(l - \underline{r}) + \cdots + f(l) \geq (n + 1)\lambda \text{ for some } \underline{r} \in \mathbb{N} \cup \{0\}\},$$
$$E_{\text{right}} := \{l \in [L] : f(l) + \cdots + f(l + \overline{r}) \geq (n + 1)\lambda \text{ for some } \overline{r} \in \mathbb{N} \cup \{0\}\}.$$

Applying the result for the one-sided maximal operator, we see that

$$|E_{\text{left}}| \leq \|f\|_1/\lambda \quad \text{and} \quad |E_{\text{right}}| \leq \|f\|_1/\lambda.$$

Since $E_{\text{left}} \cap E_{\text{right}} \neq \emptyset$, we deduce that, under the conditions specified above, the following estimate using the ceiling function must be satisfied:

$$\lambda |E_\lambda^u(f)| \leq \max_{E, E' \subseteq [L] : E \cap E' \neq \emptyset} \frac{|E \cup E'|}{\max\{|E|, |E'|\}} \|f\|_1 = \frac{2\lceil L/2 \rceil - 1}{\lceil L/2 \rceil} \|f\|_1.$$

Consequently, $\mathcal{C}^{\mathrm{u}}(\mathbb{X}_{[L]}, 1) \leq (2\lceil L/2 \rceil - 1)/\lceil L/2 \rceil < 2$, as desired. In fact, $\mathcal{C}^{\mathrm{u}}(\mathbb{X}_{[L]}, 1) = (2\lceil L/2 \rceil - 1)/\lceil L/2 \rceil$, which can be proven by taking $\tilde{f} := \mathbb{1}_{\lceil L/2 \rceil}$ and $\tilde{\lambda} := \lceil L/2 \rceil^{-1}$, and observing that $E^{\mathrm{u}}_{\tilde{\lambda}}(\tilde{f}) = [2\lceil L/2 \rceil - 1]$. $\square$

REFERENCES

[1] V. Bergelson and A. Leibman. A nilpotent Roth theorem. *Invent. Math.* **147** (2002), 427–470.
[2] G. Birkhoff. Proof of the ergodic theorem. *Proc. Natl. Acad. Sci. USA* **17** (1931), 656–660.
[3] J. Bourgain. On the maximal ergodic theorem for certain subsets of the integers. *Israel J. Math.* **61** (1988), 39–72.
[4] J. Bourgain. On the pointwise ergodic theorem on L^p for arithmetic sets. *Israel J. Math.* **61** (1988), 73–84.
[5] J. Bourgain. Pointwise ergodic theorems for arithmetic sets. *Publ. Math. Inst. Hautes Études Sci.* **69** (1989), 5–45, with an appendix by the author, H. Furstenberg, Y. Katznelson and D. S. Ornstein.
[6] Z. Buczolich and R. Mauldin. Divergent square averages. *Ann. of Math. (2)* **171** (2010), 1479–1530.
[7] A. Calderón. Ergodic theory and translation invariant operators. *Proc. Natl. Acad. Sci. USA* **59** (1968), 349–353.
[8] J.-P. Conze. Convergence des moyennes ergodiques pour des sous-suites. *Mém. Soc. Math. Fr. (N.S.)* **35** (1973), 7–15.
[9] L. Grafakos and S. Montgomery-Smith. Best constants for uncentered maximal functions. *Bull. Lond. Math. Soc.* **29** (1997), 60–64.
[10] P. R. Halmos. *Lectures on Ergodic Theory*. Mathematical Society of Japan, Tokyo, 1956.
[11] A. Ionescu, Á. Magyar, M. Mirek and T. Z. Szarek. Polynomial averages and pointwise ergodic theorems on nilpotent groups. *Invent. Math.* **231** (2023), 1023–1140.
[12] R. L. Jones, A. Seeger and J. Wright. Strong variational and jump inequalities in harmonic analysis. *Trans. Amer. Math. Soc.* **360** (2008), pp. 6711–6742.
[13] D. Kosz, M. Mirek, P. Plewa and B. Wróbel. Some remarks on dimension-free estimates for the discrete Hardy–Littlewood maximal functions. *Israel J. Math.* **254** (2023), 1–38.
[14] B. Krause, M. Mirek and T. Tao. Pointwise ergodic theorems for non-conventional bilinear polynomial averages. *Ann. of Math. (2)* **195** (2022), 997–1109.
[15] D. Lépingle. La variation d'ordre p des semi-martingales. *Z. Wahrscheinlichkeitstheorie und verw. Gebiete* **36** (1976), 295–316.
[16] A. D. Melas. The best constant for the centered Hardy–Littlewood maximal inequality. *Ann. of Math. (2)* **157** (2003), 647–688.
[17] M. Mirek, W. Słomian and T. Z. Szarek. Some remarks on oscillation inequalities. *Ergod. Th. & Dynam. Sys.* doi:10.1017/etds.2022.77. Published online 29 November 2022.
[18] G. Pisier and Q. H. Xu. The strong p-variation of martingales and orthogonal series. *Probab. Theory Related Fields* **77** (1988), 497–514.
[19] W. Słomian. Bootstrap methods in bounding discrete Radon operators. *J. Funct. Anal.* **283** (2022), 109650.
[20] E. M. Stein. On limits of sequences of operators. *Ann. of Math. (2)* **74** (1961), 140–170.
[21] J. von Neumann. Proof of the quasi-ergodic hypothesis. *Proc. Natl. Acad. Sci. USA* **18** (1932), 70–82.