

## A THEOREM OF BESICOVITCH AND A GENERALIZATION OF THE BIRKHOFF ERGODIC THEOREM

PAUL HAGELSTEIN, DANIEL HERDEN, AND ALEXANDER STOKOLOS

(Communicated by Alexander Iosevich)

ABSTRACT. A remarkable theorem of Besicovitch is that an integrable function  $f$  on  $\mathbb{R}^2$  is strongly differentiable if its associated strong maximal function  $M_S f$  is finite a.e. We provide an analogue of Besicovitch's result in the context of ergodic theory that provides a generalization of Birkhoff's Ergodic Theorem. In particular, we show that if  $f$  is a measurable function on a standard probability space and  $T$  is an invertible measure-preserving transformation on that space, then the ergodic averages of  $f$  with respect to  $T$  converge a.e. if and only if the associated ergodic maximal function  $T^* f$  is finite a.e.

### 1. INTRODUCTION

Let  $f$  be an integrable function on  $\mathbb{R}^2$ . The *Lebesgue Differentiation Theorem* tells us that for a.e.  $x \in \mathbb{R}^2$  the averages of  $f$  over disks shrinking to  $x$  tend to  $f(x)$  itself. More precisely, we have that

$$\lim_{r \rightarrow 0} \frac{1}{|B(x, r)|} \int_{B(x, r)} f = f(x) \quad \text{a.e.},$$

where  $B(x, r)$  denotes the open disk centered at  $x$  of radius  $r$  and  $|B(x, r)|$  denotes the area of that disk. For a proof of this result, the reader is encouraged to consult [13].

The issue of the averages of  $f$  over *rectangles* shrinking to  $x$  is more subtle. There do exist integrable functions  $f$  on  $\mathbb{R}^2$  such that, for a.e.  $x \in \mathbb{R}^2$ , there exists a sequence of rectangles  $\{R_{x,j}\}$  shrinking toward  $x$  for which

$$\lim_{j \rightarrow \infty} \frac{1}{|R_{x,j}|} \int_{R_{x,j}} f$$

fails to converge. Such functions  $f$  can even be characteristic functions of sets! (See [5] for a nice exposition of this result.) This result is closely related to the well-known *Kakeya Needle Problem*, and the interested reader may consult [4] for more information on this topic.

If we restrict the class of rectangles over which we allow ourselves to average, we obtain better results. In [9], Jessen, Marcinkiewicz, and Zygmund proved that if  $\mathcal{B}_2$  consists of all the open rectangles in  $\mathbb{R}^2$  whose sides are parallel to the coordinate

---

Received by the editors October 18, 2019, and, in revised form, December 9, 2020.

2020 *Mathematics Subject Classification*. Primary 37A30, 42B25.

*Key words and phrases*. Differentiation of integrals, maximal operators.

The first author was partially supported by a grant from the Simons Foundation (#521719).

axes, then for any function  $f \in L^p(\mathbb{R}^2)$  with  $1 < p \leq \infty$  one has

$$\lim_{j \rightarrow \infty} \frac{1}{|R_j|} \int_{R_j} f = f(x)$$

for a.e.  $x \in \mathbb{R}^2$ , where here  $\{R_j\}$  is any sequence of rectangles in  $\mathcal{B}_2$  shrinking toward  $x$ . Such a function  $f$  is said to be *strongly differentiable*. Jessen, Marcinkiewicz, and Zygmund proved this by showing that the *strong maximal operator*  $M_S$ , defined by

$$M_S f(x) = \sup_{x \in R \in \mathcal{B}_2} \frac{1}{|R|} \int_R |f|,$$

satisfies for every  $1 < p < \infty$  the *weak type*  $(p, p)$  estimate

$$|\{x \in \mathbb{R}^2 : M_S f(x) > \alpha\}| \leq C_p \left( \frac{\|f\|_{L^p}}{\alpha} \right)^p.$$

Most mathematicians interested in multiparameter harmonic analysis are well aware of the above result. Less well-known is a remarkable theorem that happens to be the paper in *Fundamenta Mathematicae* immediately preceding the famous paper of Jessen, Marcinkiewicz, and Zygmund. In this paper [3], *On differentiation of Lebesgue double integrals*, Besicovitch proved the following.

**Theorem 1** (Besicovitch). *Let  $f$  be an integrable function on  $\mathbb{R}^2$ . If  $M_S f$  is finite a.e., then for a.e.  $x$  we have*

$$\lim_{j \rightarrow \infty} \frac{1}{|R_j|} \int_{R_j} f = f(x)$$

whenever  $\{R_j\}$  is a sequence of sets in  $\mathcal{B}_2$  shrinking to  $x$ .

Of course, if  $f \in L^p(\mathbb{R}^2)$  for  $1 < p < \infty$ , the quantitative weak type  $(p, p)$  bound satisfied by  $M_S$  implies that  $M_S f$  will be finite a.e. It is for this reason that this paper of Besicovitch has received comparatively little attention. However, it is worth observing that the above theorem of Besicovitch provides a means for obtaining a.e. differentiability results that bypasses the need for finding quantitative weak type bounds on the associated maximal operator.

Many results in the study of differentiation of integrals have analogous results in ergodic theory; for instance the Lebesgue Differentiation Theorem is structurally very similar to the *Birkhoff Ergodic Theorem* on integrable functions. This observation may be found at least as far back as the work of Wiener [14]. In that regard, we consider what the companion result of Besicovitch's Theorem might be when replacing the strong maximal operator  $M_S$  by an ergodic maximal operator. We are led immediately to the following theorem.

**Theorem 2.** *Let  $T$  be an invertible measure-preserving transformation on the standard probability space  $(X, \Sigma, \mu)$  and let  $f$  be a  $\mu$ -measurable function on that space. If  $T^* f(x)$  is finite  $\mu$ -a.e., where  $T^* f$  is the ergodic maximal function defined by*

$$T^* f(x) = \sup_{n \geq 1} \frac{1}{n} \left| \sum_{k=0}^{n-1} f(T^k x) \right|,$$

then the limit

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

exists  $\mu$ -a.e.

We remark that if  $f$  is integrable, then by the Birkhoff Ergodic Theorem the above limit automatically holds a.e.

The purpose of this paper is to provide a proof of the above theorem. Section 2 provides a proof of this theorem in the special case that  $T$  is an ergodic transformation. Section 3 provides a proof of the general case by means of the ergodic decomposition theorem. In the last section we indicate further directions of research, both in ergodic theory as well as the theory of differentiation of integrals.

We remark that our techniques in Section 2 are strongly influenced by the work of Aaronson. In fact, the key lemma of this section is stated without proof in [1] and as Exercise 2.3.1 in [2]. For completeness, we provide a proof, especially as it may be beneficial for harmonic analysts reading the paper without an extensive background in ergodic theory.

## 2. MAXIMAL FUNCTIONS ASSOCIATED TO ERGODIC TRANSFORMATIONS

The purpose of this section is to state and prove the following lemma.

**Lemma 1.** *Let  $T$  be an ergodic measure-preserving transformation on the probability space  $(X, \Sigma, \mu)$  and let  $f$  be a  $\mu$ -measurable function on that space. If  $T^*f(x)$  is finite  $\mu$ -a.e., where  $T^*f$  is the ergodic maximal function defined by*

$$T^*f(x) = \sup_{n \geq 1} \frac{1}{n} \left| \sum_{k=0}^{n-1} f(T^k x) \right| ,$$

then the limit

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

exists  $\mu$ -a.e.

*Proof.* Since  $T^*f(x) < \infty$   $\mu$ -a.e., we can choose some  $M \geq 0$  and a set  $B \in \Sigma$  such that  $\mu(B) > 0$  with  $T^*f(x) \leq M$  for all  $x \in B$ . Letting

$$f_n(x) = \sum_{k=0}^{n-1} f(T^k x) ,$$

we have

$$(2.1) \quad \frac{|f_n(x)|}{n} \leq M \text{ for all } x \in B \text{ and } n \geq 1.$$

The goal now is to show that  $\frac{f_n(x)}{n}$  converges to a finite constant  $\mu$ -a.e. on  $B$ , in which case the  $T$ -invariant  $\mu$ -measurable set

$$\left\{ x \in X : \frac{f_n(x)}{n} \text{ converges for } n \rightarrow \infty \right\}$$

contains a subset of positive measure. As  $T$  is ergodic, this would complete the proof.

The integer

$$\phi(x) = \inf\{n \geq 1 : T^n x \in B\}$$

is finite for  $\mu$ -a.e.  $x \in B$ . Without loss of generality we assume  $\phi(x) < \infty$  for all  $x \in B$ . Let  $T_B$  be the induced transformation of  $T$  on  $B$  given by

$$T_B(x) = T^{\phi(x)} x .$$

Note  $T_B$  is an ergodic measure-preserving transformation on  $(B, \Sigma \cap B, \mu)$  and  $\int_B \phi \, d\mu = 1$ . (See [10–12] in this regard.)

Define now the function  $g$  on  $B$  by

$$g(x) = f_{\phi(x)}(x) .$$

Moreover define the functions  $\phi_n, g_n$  on  $B$  by

$$\phi_n(x) = \sum_{k=0}^{n-1} \phi(T_B^k x) \quad \text{and} \quad g_n(x) = \sum_{k=0}^{n-1} g(T_B^k x) .$$

Introducing the measure  $\mu_B = \mu/\mu(B)$ , we have the probability space  $(B, \Sigma \cap B, \mu_B)$ , and applying the Birkhoff Ergodic Theorem to the integrable function  $\phi$ , we have

$$(2.2) \quad \lim_{n \rightarrow \infty} \frac{\phi_n(x)}{n} = \int_B \phi \, d\mu_B = \frac{1}{\mu(B)} \int_B \phi \, d\mu = \frac{1}{\mu(B)} \quad \mu\text{-a.e. on } B .$$

In particular, for  $\mu$ -a.e.  $x$  on  $B$  we have that

$$(2.3) \quad \frac{\phi(T_B^n x)}{n} = \frac{\phi_{n+1}(x) - \phi_n(x)}{n} = \frac{n+1}{n} \cdot \frac{\phi_{n+1}(x)}{n+1} - \frac{\phi_n(x)}{n} \rightarrow 0$$

as  $n \rightarrow \infty$ . Note that  $g$  is a measurable function with  $|g(x)| \leq M\phi(x)$  on  $B$ . Thus  $\int_B |g| \, d\mu \leq M < \infty$ , and  $g$  is integrable. Hence by the Birkhoff Ergodic Theorem there exists a finite constant  $c$  such that

$$(2.4) \quad \lim_{n \rightarrow \infty} \frac{g_n(x)}{n} = c \quad \mu\text{-a.e. on } B .$$

For the remainder of this proof, fix any  $x \in B$  such that (2.3) and (2.4) hold. For  $n \geq 0$ , we define the integers  $k_n = k_n(x)$  and  $l_n = l_n(x)$  such that

$$n = \phi_{k_n}(x) + l_n \quad \text{with} \quad 0 \leq l_n < \phi(T_B^{k_n} x) .$$

Note that  $k_n$  and  $l_n$  are uniquely determined with  $k_n \rightarrow \infty$  for  $n \rightarrow \infty$ . With (2.3), we have the estimate

$$0 \leq \frac{l_n}{k_n} \leq \frac{\phi(T_B^{k_n} x)}{k_n} \rightarrow 0 \quad \text{as } n \rightarrow \infty .$$

Hence

$$(2.5) \quad \lim_{n \rightarrow \infty} \frac{l_n}{k_n} = 0 .$$

With (2.2), we have

$$(2.6) \quad \lim_{n \rightarrow \infty} \frac{n}{k_n} = \lim_{n \rightarrow \infty} \frac{\phi_{k_n}(x)}{k_n} + \lim_{n \rightarrow \infty} \frac{l_n}{k_n} = \lim_{n \rightarrow \infty} \frac{\phi_n(x)}{n} = \frac{1}{\mu(B)} .$$

Observe that

$$f_n(x) = f_{\phi_{k_n}(x)}(x) + f_{l_n}(T_B^{k_n} x) = g_{k_n}(x) + f_{l_n}(T_B^{k_n} x) .$$

With (2.1), (2.5), and (2.6), we have the estimate

$$0 \leq \frac{|f_{l_n}(T_B^{k_n} x)|}{n} \leq \frac{Ml_n}{n} = M \cdot \frac{k_n}{n} \cdot \frac{l_n}{k_n} \rightarrow 0 \quad \text{for } n \rightarrow \infty .$$

Together with (2.4) and (2.6), we conclude

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{f_n(x)}{n} &= \lim_{n \rightarrow \infty} \frac{g_{k_n}(x)}{n} + \lim_{n \rightarrow \infty} \frac{f_{l_n}(T_B^{k_n} x)}{n} \\ &= \lim_{n \rightarrow \infty} \frac{g_{k_n}(x)}{n} = \lim_{n \rightarrow \infty} \frac{g_{k_n}(x) k_n}{k_n n} = c\mu(B). \end{aligned}$$

As this holds for all  $x$  satisfying (2.3) and (2.4), and both of these hold for  $\mu$ -a.e.  $x$  on  $B$ , the lemma follows.  $\square$

### 3. MAXIMAL FUNCTIONS ASSOCIATED TO MEASURE-PRESERVING TRANSFORMATIONS

The purpose of this section is to provide a proof of Theorem 2 by using Lemma 1 combined with the Ergodic Decomposition Theorem.

The version of the Ergodic Decomposition Theorem we use follows from Theorem 2.2.9 in [2] and is stated as follows:

**Theorem 3** (Ergodic Decomposition Theorem). *Let  $T$  be an invertible measure-preserving transformation on a standard probability space  $(X, \Sigma, \mu)$ . Then there is a probability space  $(Y, \Lambda, \eta)$  and a collection of probabilities*

$$\{\mu_y : y \in Y\}$$

on  $(X, \Sigma)$  such that

- (i) for  $y \in Y$ ,  $T$  is an invertible measure-preserving ergodic transformation of  $(X, \Sigma, \mu_y)$ , and
- (ii) for  $A \in \Sigma$ , the map  $y \mapsto \mu_y(A)$  is measurable, with

$$\mu(A) = \int_Y \mu_y(A) d\eta(y).$$

*Proof of Theorem 2.* Let  $f$  be a  $\mu$ -measurable function on  $X$  and suppose that  $T^* f(x)$  is finite for  $\mu$ -a.e.  $x$ . Let  $A$  be the set of points in  $X$  such that

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

exists. Note that  $A$  is indeed  $\mu$ -measurable, being the complement of the set

$$\bigcup_{\alpha, \beta \in \mathbb{Q}} \left\{ x \in X : \liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k x) < \alpha < \beta < \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k x) \right\}.$$

It suffices to show that  $\mu(A) = 1$ . If  $\mu(A) < 1$ , by the ergodic decomposition above there would exist  $y \in Y$  such that  $\mu_y(A) < 1$ . However, as  $T$  is an ergodic transformation on the space  $(X, \Sigma, \mu_y)$ , by Lemma 1 we would have  $\mu_y(A) = 1$ , a contradiction.  $\square$

*Remark.* The condition of invertibility of  $T$  in Theorem 2 enables the use of the version of the Ergodic Decomposition Theorem we provide here. The conclusion of Theorem 2 holds whenever  $T$  and  $(X, \Sigma, \mu)$  permit a decomposition as in the conclusion of the Ergodic Decomposition Theorem.

## 4. FUTURE DIRECTIONS

The theorem of Besicovitch and its analogue in the context of ergodic theory do suggest the following future directions of research, some related to very recent work of Hagelstein and Parissis [7].

**Problem.** *Differentiation* of a function  $f$  relates to averages of  $f$  over sets of arbitrarily small diameter, whereas the strong maximal operator  $M_{\mathcal{B}}$  involves rectangles of any size. This suggests that we might be able to strengthen Theorem 1 by the following:

Given a collection  $\mathcal{B} = \{R_j\}$  of open sets in  $\mathbb{R}^n$  we define the maximal operator  $M_{\mathcal{B}}f$  by

$$M_{\mathcal{B}}f(x) = \sup_{x \in R \in \mathcal{B}} \frac{1}{|R|} \int_R |f|.$$

For  $r > 0$  we set

$$\mathcal{B}_r = \{R \in \mathcal{B} : \text{diam } R < r\}.$$

Define the maximal operator  $\tilde{M}_{\mathcal{B}}$  by

$$\tilde{M}_{\mathcal{B}}f(x) = \lim_{r \rightarrow 0} M_{\mathcal{B}_r}f(x).$$

**Conjecture 1.** *If  $\mathcal{B}_2$  is the collection of rectangles in  $\mathbb{R}^2$  whose sides are parallel to the coordinate axes and  $\tilde{M}_{\mathcal{B}}f(x)$  is finite a.e., then  $f$  is strongly differentiable.*

Of course, this conjecture may be generalized in many ways, e.g., if  $\mathcal{B}$  is a translation invariant density basis of open sets in  $\mathbb{R}^n$  and  $\tilde{M}_{\mathcal{B}}f(x)$  is finite a.e., then  $\mathcal{B}$  differentiates  $f$ .

**Problem.** Theorem 1 was generalized by de Guzmán and Menárguez to encompass homothety invariant Busemann-Feller bases associated to convex sets in  $\mathbb{R}^n$  with a center of symmetry. (For a proof one may consult Chapter IV of [5].) It is natural to consider to what extent Theorem 1 may be further generalized. One possible generalization is provided by the following:

**Conjecture 2.** *Let  $\mathcal{B}$  be a translation invariant density basis consisting of open sets in  $\mathbb{R}^n$ . If  $M_{\mathcal{B}}f(x) < \infty$  a.e., then*

$$\lim_{j \rightarrow \infty} \frac{1}{|R_j|} \int_{R_j} f = f(x) \text{ a.e.}$$

where the limit is over an arbitrary sequence of sets  $\{R_j\}$  in  $\mathcal{B}$  shrinking to  $x$ .

**Problem.** It would be natural to desire to obtain a *multiparameter* analogue of Theorem 2 in the spirit of previous work of Hagelstein and Stokolos in [8] and Hagelstein and Parissis in [6]. In particular we make the following conjecture.

**Conjecture 3.** *Let  $U$  and  $V$  be a nonperiodic collection of invertible measure-preserving transformations on a standard probability space  $(X, \Sigma, \mu)$  and define the associated strong ergodic maximal operator by*

$$M_{U,V}f(x) = \sup_{m,n \geq 1} \frac{1}{mn} \left| \sum_{j=0}^{m-1} \sum_{k=0}^{n-1} f(U^j V^k x) \right|.$$

If  $M_{U,V}f(x) < \infty$   $\mu$ -a.e., then

$$\lim_{m,n \rightarrow \infty} \frac{1}{mn} \sum_{j=0}^{m-1} \sum_{k=0}^{n-1} f(U^j V^k x) \text{ converges } \mu\text{-a.e.}$$

Of course analogues of this conjecture exist where  $(m, n)$  are allowed to only lie in a specified set  $\Gamma$  in  $\mathbb{Z}_+^2$ .

All of these topics are subjects of ongoing research.

#### ACKNOWLEDGMENT

It is our pleasure to thank Jon Aaronson for his helpful comments and advice regarding this paper.

#### REFERENCES

- [1] J. Aaronson, *An ergodic theorem with large normalising constants*, Israel J. Math. **38** (1981), no. 3, 182–188, DOI 10.1007/BF02760803. MR605376
- [2] J. Aaronson, *An introduction to infinite ergodic theory*, Mathematical Surveys and Monographs, vol. 50, American Mathematical Society, Providence, RI, 1997. MR1450400
- [3] A. S. Besicovitch, *On differentiation of Lebesgue double integrals*, Fund. Math. **25** (1935), 209–216.
- [4] K. J. Falconer, *The geometry of fractal sets*, Cambridge Tracts in Mathematics, vol. 85, Cambridge University Press, Cambridge, 1986. MR867284
- [5] M. de Guzmán, *Differentiation of integrals in  $R^n$* , Lecture Notes in Mathematics, Vol. 481, Springer-Verlag, Berlin-New York, 1975. With appendices by Antonio Córdoba, and Robert Fefferman, and two by Roberto Moriyón. MR0457661
- [6] P. Hagelstein and I. Parissis, *Hölder continuity of Tauberian constants associated with discrete and ergodic strong maximal operators*, New York J. Math. **23** (2017), 1219–1236. MR3711277
- [7] P. Hagelstein and I. Parissis, *Tauberian constants associated to centered translation invariant density bases*, Fund. Math. **243** (2018), no. 2, 169–177, DOI 10.4064/fm409-2-2018. MR3846848
- [8] P. Hagelstein and A. Stokolos, *Transference of weak type bounds of multiparameter ergodic and geometric maximal operators*, Fund. Math. **218** (2012), no. 3, 269–284, DOI 10.4064/fm218-3-4. MR2982778
- [9] B. Jessen, J. Marcinkiewicz, and A. Zygmund, *Note on the differentiability of multiple integrals*, Fund. Math. **25** (1935), 217–234.
- [10] M. Kac, *On the notion of recurrence in discrete stochastic processes*, Bull. Amer. Math. Soc. **53** (1947), 1002–1010, DOI 10.1090/S0002-9904-1947-08927-8. MR22323
- [11] S. Kakutani, *Induced measure preserving transformations*, Proc. Imp. Acad. Tokyo **19** (1943), 635–641. MR14222
- [12] K. Petersen, *Ergodic theory*, Cambridge Studies in Advanced Mathematics, vol. 2, Cambridge University Press, Cambridge, 1983. MR833286
- [13] E. M. Stein, *Singular integrals and differentiability properties of functions*, Princeton Mathematical Series, No. 30, Princeton University Press, Princeton, N.J., 1970. MR0290095
- [14] N. Wiener, *The ergodic theorem*, Duke Math. J. **5** (1939), no. 1, 1–18, DOI 10.1215/S0012-7094-39-00501-6. MR1546100

DEPARTMENT OF MATHEMATICS, BAYLOR UNIVERSITY, SID RICHARDSON BUILDING, 1410 S.  
4TH STREET, WACO, TEXAS 76706

*Email address:* [paul.hagelstein@baylor.edu](mailto:paul.hagelstein@baylor.edu)

DEPARTMENT OF MATHEMATICS, BAYLOR UNIVERSITY, SID RICHARDSON BUILDING, 1410 S.  
4TH STREET, WACO, TEXAS 76706

*Email address:* [daniel.herden@baylor.edu](mailto:daniel.herden@baylor.edu)

DEPARTMENT OF MATHEMATICAL SCIENCES, GEORGIA SOUTHERN UNIVERSITY, 203 GEORGIA  
AVENUE, STATESBORO, GEORGIA 30460-8093

*Email address:* [astokolos@georgiasouthern.edu](mailto:astokolos@georgiasouthern.edu)