JOURNAL OF APPROXIMATION THEORY 2, 249–257 (1969)

On Certain Lemmas of Marcinkiewicz and Carleson*

A. ZYGMUND

University of Chicago, Chicago, Illinois 60637

1. In this lecture I am going to discuss things that are known, but probably not *too* well known. Correspondingly, the character of the presentation be more expository than exploratory, though there may be some elements of novelty here. I shall discuss certain metric properties of sets and functions, properties that are important for applications of the Lebesgue integral to classical analysis.

Two such properties are fundamental, and I shall begin with the one which is familiar to all: it is the fact that the derivative of the indefinite integral of an integrable function exists almost everywhere, and almost everywhere is equal to the integrand, More specifically, the result is as follows: if $f(x) = f(x_1, x_2, \ldots, x_n)$ is an integrable function defined over the n-dimensional Euclidean space E_n, and if

$$F(E) = \int_E f$$

is the indefinite integral of f, then at almost all points x we have

$$\frac{F(Q)}{|Q|} \to f(x), \tag{1.1}$$

where Q designates a cube containing the point x and shrinking to x, and $|Q|$ is the measure of Q. The result holds if Q is an n-dimensional interval (parallelepiped) containing x, provided the ratio of the largest and the smallest edge of Q remains bounded (we consider only intervals with edges parallel to the co-ordinate axes).

Also, and this easily follows from the preceding, we have at almost all points x the somewhat stronger result, namely

$$\frac{\int_Q |f(y) - f(x)|\, dy}{|Q|} \to 0. \tag{1.2}$$

The points at which (1.2)—or a suitable generalization of it—holds, are usually called *Lebesgue points* of the function f.

---

* A lecture delivered at the Second Symposium on Inequalities at USAF Academy, Colorado, August 1967.

One need not stress the importance of the fact; we all know that without it the present-day analysis would be impossible. What is, however, less known is that, in a number of problems, the result does not suffice and must be supplemented by another result whose general importance seems to have been recognized for the first time by Marcinkiewicz.

Some of us who worked in Fourier series in the period between the two world wars remember a number of problems which remained unsolved for a long time and which were rather tantalizing; tantalizing, because, without appearing to be out of reach (unlike, e.g., the problem of convergence almost everywhere of Fourier series of continuous functions), they were still quite elusive. I shall mention three examples as illustrations.

(a) In the early nineteen twenties Carleman proved that if $f(x)$ is a periodic function of the class L^2 and $s_k(x)$ is the kth partial sum of the Fourier series of f, then for almost all values of x we have the relation

$$\frac{1}{n+1}\sum_{k=0}^{n}|s_k(x)-f(x)|^2 \to 0 \tag{1.3}$$

("strong summability" of Fourier series). The question naturally arose whether this result, or a suitable modification of it, holds for functions that are merely integrable. During a number of years much effort was being spent on it and a number of generalizations were obtained. For example, it was shown that the result holds for f in any class L^p, provided p is strictly greater than 1, and that in this case we can even replace the exponent 2 in (1.3) by any positive q, arbitrarily large. (See, e.g., [*10*, II, p. 180] and references there.) The relation (1.3) and its generalizations were usually shown to hold at the Lebesgue points of f, so that when Hardy and Littlewood [*3*] showed that, for f merely integrable, (1.3) need not hold at such points the possibility arose (cf. [*3*]) that perhaps, after all, strong summability need not hold almost everywhere for functions that are merely integrable. The question remained unsolved until 1939 when Marcinkiewicz showed that (1.3) is indeed true almost everywhere for f integrable (see [*4*] or [*10*, II, p. 184]).

(b) A very well-known result asserts that the Fourier series of any integrable function $f(x)$ is summable (C,δ), $\delta > 0$, almost everywhere; more precisely, at each Lebesgue point of f. It is also very well known that the result fails for $\delta = 0$: there are integrable functions whose Fourier series diverge at each point. Thus the result of Hardy and Littlewood that the termwise differentiated Fourier series of a function f is summable $(C, 1+\delta)$, $\delta > 0$, at each point where f' exists and is finite, appeared final, the more so as they showed by examples that the conclusion fails for $\delta = 0$. But Marcinkiewicz showed that though the conclusion may fail at individual points it is nevertheless valid almost everywhere; more precisely, if $f'(x)$ exists at each point of a set E, then the termwise differentiated Fourier series of f is summable $(C, 1)$ almost everywhere in E (see [5] or [10, II, p. 81]).

(c) One of the classical results of the theory of Fourier series asserts that if a periodic and continuous function $f(x)$ satisfies the condition

$$f(x+h)-f(x)=o\left\{\frac{1}{\log 1/|h|}\right\} \qquad (h \to 0) \tag{1.4}$$

uniformly in x, then the Fourier series of f converges uniformly (the Dini–Lipschitz test). It is easy to show by examples that a continuous function f may satisfy the condition (1.4) at *some* point x without its Fourier series converging at that point. The question remained: if a periodic and merely integrable f satisfies (1.4) at each point x of a set E, does the Fourier series of f necessarily converge almost everywhere in E? It was again Marcinkiewicz (see [6] or [10, II, p. 170]) who showed that it is actually so. He even proved a stronger result: the conclusion holds if at each point $x \in E$ we have instead of (1.4) the obviously weaker relation:

$$\frac{1}{h}\int_0^h |f(x+t)-f(x)|\,dt=O\left\{\frac{1}{\log 1/|h|}\right\} \qquad (h \to 0). \tag{1.5}$$

(Observe that we have "O" here). Incidentally, he also showed that the result is best possible: the conclusion fails if the expression $1/(\log 1/h)$ on the right of (1.5) is replaced by any function of h tending to 0 more slowly [5].

In all three cases the solution was made possible by an application of the same theorem which expresses a certain metric property of sets and functions and which succeeds where the theorem about the differentiability of integrals seems to be insufficient. And it is a curious fact that this property, in a somewhat modified form, plays an important role in Carleson's proof of his fundamental theorem on the convergence almost everywhere of Fourier series of functions of the class L^2 (see [2], Lemma 5). I shall now describe that property.

2. Given any closed set P situated in the Euclidean space E_n we shall call the distance of any point x from P the *distance function*; it will be denoted by $\delta(x;P)$, or simply by $\delta(x)$. Thus $\delta(x)=0$ if and only if x is in P. If $n=1$ and (a,b) is any interval contiguous to P and situated between the terminal points of P, then the graph of $\delta(x)$ over (a,b) is an isosceles triangle with base (a,b) and altitude $\frac{1}{2}(b-a)$; outside the terminal points of P the graph of $\delta(x)$ is a

linear function. If $n > 1$, the graph of $\delta(x)$ is in general much less simple, but since if we move from a point x to another point y the distance from P does not increase by more than $|x-y|$, it is clear that

$$|\delta(x) - \delta(y)| \leqslant |x - y|,$$

that is $\delta(x)$ satisfies a Lipschitz condition of order 1.

Marcinkiewicz's lemma (or theorem) may be stated as follows (see [*10*, I, pp. 129–131 and p. 377].)

(A) *Let P be a closed subset of E_n and $\delta(x) = \delta(x;P)$ the corresponding distance function. Let λ be a positive number and $f(x)$ a non-negative function integrable over the complement Q of P. Then for almost all points $x \in P$ the integral*

$$J_\lambda(x) = J_\lambda(x;f,P) = \int_{E_n} \frac{\delta^\lambda(y)\, f(y)}{|x-y|^{n+\lambda}}\, dy \tag{2.1}$$

is finite.

In particular, if P is bounded and K is any finite sphere containing P, the integral

$$\int_K \frac{\delta^\lambda(y)}{|x-y|^{n+\lambda}}\, dy \tag{2.2}$$

is finite almost everywhere in P.

The usual proof of (A) actually gives a little more (see [*10*, I], pp. 129–131), namely *the function $J_\lambda(x)$ is integrable over P.* A few years ago Professor R. O'Neil pointed out to the author that if $f \in L^p(E_n - P)$, $1 \leqslant p < \infty$, then $J_\lambda \in L^p(P)$ and we have the obvious inequalities for the norms. His proof was based on Hardy–Littlewood maximal theorems. In what follows we give a slightly different proof of (A) and its generalization by using a modification of the integral J_λ.

Whatever the behavior of the integral $J_\lambda(x)$ in P, it generally diverges outside P; this is certainly true of the integral (2.2) which is infinite at the points x interior to $K - P$. Let us however consider the following modification of J_λ:

$$H_\lambda(x) = \int_{E_n} \frac{\delta^\lambda(y) f(y)}{|x-y|^{n+\lambda} + \delta^{n+\lambda}(x)}\, dy. \tag{2.3}$$

It has two obvious properties: (a) it coincides with $J_\lambda(x)$ for $x \in P$; (b) it is finite at each point x not in P, provided $f \in L^p(E_n - P)$, $1 \leqslant p \leqslant \infty$. To prove the latter we consider separately the y's close to x, in which case the denominator stays away from 0, and the more distant y's to which we can apply Hölder's inequality.

We shall also consider another modification of J_λ, namely

$$H_\lambda'(x) = \int_{E_n} \frac{\delta^\lambda(y) f(y)}{|x-y|^{n+\lambda} + \delta^{n+\lambda}(y)} dy, \tag{2.4}$$

which, like H_λ, is finite at each point not in P. In view of the inequality $\delta(y) \leqslant |x-y| + \delta(x)$ we have, by Jensen's inequality,

$$\delta^{n+\lambda}(y) \leqslant 2^{n+\lambda-1}\{|x-y|^{n+\lambda} + \delta^{n+\lambda}(x)\}$$

and a similar inequality with x and y interchanged. We immediately deduce from this that

$$A^{-1} H_\lambda'(x) \leqslant H_\lambda(x) \leqslant A H_\lambda'(x) \, (A = 2^{n+\lambda-1} + 1), \tag{2.5}$$

so that inequalities for H_λ' immediately lead to inequalities for H_λ, but H_λ' is sometimes easier to deal with than H_λ.

Also, since the values of H_λ and H_λ' are independent of the values of f on P, we shall assume for the sake of simplicity of enunciation that f is defined over the whole of E_n and, say, is 0 in P.

(B) *If* $f \in L^p(E_n)$, $1 \leqslant p < \infty$, *then* $H_\lambda \in L^p(E_n)$ *and*

$$\left\{\int_{E_n} H^p(x)\,dx\right\}^{1/p} \leqslant Ap\left(\int_{E_n} f^p\,dx\right)^{1/p} \qquad (A = A_{n,\lambda}). \tag{2.6}$$

If f *is bounded, say* $0 \leqslant f \leqslant 1$, *and has support in a sphere* $K \supset P$, *then*

$$\int_K \exp\{\gamma H_\lambda(x)\}\,dx \leqslant A|K| \qquad (A = A_{n,\lambda}) \tag{2.7}$$

provided γ *is small enough,* $\gamma \leqslant A'_{n,\lambda}$.

3. We first prove (2.6), with H_λ' instead of H_λ.

Let $g(x)$ be any non-negative locally integrable function and let $\bar{g}(x)$ denote the corresponding Hardy–Littlewood maximal function

$$\bar{g}(x) = \operatorname*{Sup}_{\rho} \left\{\rho^{-n} \int_{|z| \leqslant \rho} g(x+z)\,dz\right\}.$$

It is a familiar fact that if $g \in L^r(E_n)$, $1 < r < \infty$, then $\bar{g}$ is likewise in $L^r(E_n)$, and

$$\|\bar{g}\|_r \leqslant A_n \frac{r}{r-1} \|g\|_r.^{1} \tag{3.1}$$

[1] The fact that $g \in L^r(E_n)$ implies $\bar{g} \in L^r(E_n)$, $1 < r < \infty$, and the inequality $\|\bar{g}\|_r \leqslant \{2r/(r-1)\}^n \|g\|_r$ is, using repeated integration, a simple corollary of the Hardy–Littlewood classical result for $n = 1$. The estimate (3.1), where we have $r/(r-1)$ in the first power, is slightly deeper and is due to Wiener [9]. See also [1], where it is deduced from the case $n = 1$ by the "method of rotation."

Also the following observation is very well known (and immediate): if λ and δ are positive numbers, $g(x)$ is non-negative and in $L^r(E_n)$, $1 \leqslant r \leqslant \infty$, then

$$\int_{E_n} \frac{\delta^\lambda g(x+z)}{|z|^{n+\lambda}+\delta^{n+\lambda}}\,dz \leqslant A\bar{g}(x) \qquad (A = A_{n,\lambda}). \tag{3.2}$$

For, decomposing the integral into two, extended, respectively, over $|z| \leqslant \delta$ and $|z| \geqslant \delta$, we see that the first is majorized by

$$\delta^{-n}\int_{|z|\leqslant\delta} g(x+z)\,dz \leqslant \bar{g}(x),$$

and the second by

$$\int_{|z|\geqslant\delta} \frac{g(x+z)}{|z|^{n+\lambda}}\,dz \leqslant A_{n,\lambda}\bar{g}(x),$$

as a simple integration by parts shows. This proves (3.2).

Let now $g(x)$ be any non-negative function such that $\|g\|_{p'} = 1$, where $p' = p/(p-1)$. Then

$$\int_{E_n} H_\lambda'(x)g(x)\,dx = \int_{E_n} f(y)\delta^\lambda(y)\,dy\left\{\int_{E_n}\frac{g(x)\,dx}{|x-y|^{n+\lambda}+\delta^{n+\lambda}_{(y)}}\right\}$$

$$\leqslant A_{n,\lambda}\int_{E_n} f(y)\bar{g}(y)\,dy$$

$$\leqslant A_{n,\lambda}\|f\|_p\|\bar{g}\|_{p'} \leqslant A_{n,\lambda}\|f\|_p\cdot A_n\frac{p'}{p'-1}\|g\|_{p'}$$

$$= p\,.\,A_{n,\lambda}\|f\|_p,$$

and since the least upper bound of the left-hand side here for all such g is the left-hand side of (2.6) with H_λ' for H_λ, this proves the first part of (B).

Passing to (2.7) we observe that the left-hand side there is

$$|K| + \sum_{p=1}^{\infty}\frac{\gamma^p}{p!}\int_K H_\lambda{}^p\,dx \leqslant |K| + \sum_{p=1}^{\infty}\frac{\gamma^p}{p!}A^p p^p\int_K f^p\,dx$$

$$\leqslant |K|\left(1+\sum_{p=1}^{\infty}\frac{(\gamma Ap)^p}{p!}\right).$$

Since the last series converges for $\gamma Ae < 1$, (2.7) follows.

4. Let us consider (2.7) in the special case $f \equiv 1$ in K, and let $\omega(\eta) = \omega(\eta; K)$ be the distribution function of H_λ in K, that is, the measure of the set of the points $x \in K$ such that $H_\lambda(x) > \eta > 0$. An immediate corollary of (2.7) is

(C) *If γ is sufficiently small, $0<\gamma\leqslant A'_{n,\lambda}$, then the distribution function of $H_\lambda(x;1,P)$ in K satisfies an inequality*

$$\omega(\eta)<A_{n,\lambda}|K|\,e^{-\gamma\eta}. \tag{4.1}$$

It is clear that, conversely, (4.1) gives (2.7) with any smaller value of γ.

5. Let K be any finite closed sphere in E_n, and let $K_1, K_2, \ldots$ be a sequence, finite or not, of non-overlapping spheres contained in K. The center of K_j we denote by ξ_j, the radius by r_j. Let $K_j{}^*$ be the sphere concentric with K_j of radius $\frac{1}{2}r_j$. Let $K_j{}^0$ be the interior of K_j and $P=K-\bigcup K_j{}^0$. Let $\delta(x)$ be the distance of x from P. If $x\in K_j{}^*$, then $\frac{1}{2}r_j\leqslant\delta(x)\leqslant r_j$. Consider the function $H_\lambda{}'(x)$ for f equal to the characteristic function of the set $\bigcup K_j{}^*$. Thus

$$H_\lambda{}'(x)=\sum_j\int_{K_j{}^*}\frac{\delta^\lambda(y)}{|x-y|^{n+\lambda}+\delta^{n+\lambda}(y)}\,dy$$

and an elementary argument (we consider separately the cases when x is or is not in $K_j{}^*$) shows that $H_\lambda{}'(x)$ is contained between two positive multiples, depending on n and λ only, of the sum

$$S_\lambda(x)=\sum_j\frac{r_j^{n+\lambda}}{|x-\xi_j|^{n+\lambda}+r_j^{n+\lambda}}. \tag{5.1}$$

Hence, using (C), or rather its analog for H'_λ, we obtain the following result:

(D) *With the notation just introduced, the distribution function on K of the sum $S_\lambda(x)$ satisfies the inequality* (4.1)

For $n=\lambda=1$ this is Lemma 5 of Carleson's paper [*2*]. In his proof, which is very short, he uses properties of harmonic functions. We now see that his lemma has close connection with the results of Marcinkiewicz, and indicating this was one of the purposes of this lecture. Obviously the result holds if the spheres K, K_j are replaced by cubes.

6. In all the foregoing the parameter λ was a strictly positive number. If $\lambda=0$ the arguments break down, and one can also show by examples that the theorems are false. However, already Marcinkiewicz considered in this case the substitute function

$$J_0(x;f,P)=\int_K\frac{\log\{1/\delta(y)\}^{-1}f(y)}{|x-y|^n}\,dy \tag{6.1}$$

which has a number of properties in common with J_λ, $\lambda>0$. Since the function $|x|^{-n}$ is not integrable at infinity it is convenient to integrate in (6.1) over a finite sphere K, which is supposed to contain our closed set P. Morevoer, it will be convenient to assume that the diameter of the sphere is $\leqslant \frac{1}{2}$, so that the integrand in (6.1) is non-negative. Correspondingly, we shall also consider the function

$$H_0(x;f,P) = \int_K \frac{f(y)\log\{1/\delta(y)\}^{-1}}{|x-y|^n + \delta^n(x)}\,dy \tag{6.2}$$

which for $x \in P$ reduces to J_0, and the function $H_0'(x)$ which is obtained from $H_0(x)$ by replacing the term $\delta^n(x)$ in the denominator by $\delta^n(y)$. The inequality (2.5) holds also in this case. We shall only consider the behaviour of H_0 and H_0' on K and we have then the following theorem, in which the diameter of K is $\leqslant \frac{1}{2}$.

(E) *If $f \in L^p(K)$, $1 \leqslant p < \infty$, then $H_0 \in L^p(K)$ and*

$$\left\{\int_K H_0{}^p(x)\,dx\right\}^{1/p} \leqslant Ap\left\{\int_K f^p\,dx\right\}^{1/p} \qquad (A = A_n) \tag{6.3}$$

and if $f \equiv 1$ in K, then

$$\int_K \exp\gamma H_0(x)\,dx \leqslant A|K| \qquad (A = A_n, \gamma \leqslant A_n') \tag{6.4}$$

The proof is parallel to that of B. In the case of (6.3) it is enough to observe that if $g(x)$ is non-negative in K and the integral of $g^{p'}$ over K is 1, then the integral of $H_0'(x)g(x)$ over K can be written

$$\int_K f(y)\left\{(\log 1/\delta(y))^{-1}\int_K \frac{g(x)\,dx}{|x-y|^n + \delta^n(y)}\right\}dy,$$

and that the expression in curly brackets is majorized by the sum

$$\{\log 1/\delta(y)\}^{-1}\int_{|z|\leqslant\delta(y)} \frac{g(y+z)}{\delta^n(y)}\,dz + \{\log 1/\delta(y)\}^{-1}\int_{\delta(y)\leqslant|z|\leqslant 1} \frac{g(y+z)}{|z|^n}\,dz,$$

of which the first term does not exceed $\bar{g}(y)(\log 2)^{-1}$, and the second does not exceed $A_n\bar{g}(y)$. The proof of (6.4) is identical with that of (2.7). We also have analogous results for the function

$$S_0(x) = \sum_j \frac{r_j{}^n\{\log(1/r_j)\}^{-1}}{|x-\xi_j|^n + r_j{}^n}$$

[cf. (5.1)].

We add that, in the case $p = 1$, the integrability of H_0 over K implies the finiteness of J_0 almost everywhere in P. The latter result is, essentially, one of the original results of Marcinkiewicz.[2]

[2] Marcinkiewicz [7] himself considered only the case $n = 1$ and instead of $\delta(x)$ the function $\delta^*(x)$ equal to 0 in P and to $b_i - a_i$ in each interval (a_i, b_i) contiguous to P. But the function δ seems to be more natural than δ^* and extensions to higher dimensions more routine.

7. We conclude with an incomplete result. Following ideas of Ostrow and Stein [*8*], we may consider instead of H_λ the somewhat more general integral

$$T_\lambda(x) = \int_{E_n} \frac{f(y)\,\delta^\lambda(y)\,\phi(x-y)}{|x-y|^{n+\lambda} + \delta^{n+\lambda}(x)}\,dy,$$

where ϕ is a non-negative locally integrable function satisfying the condition

$$\int_{|x|\leqslant\rho} \phi(x)\,dx \leqslant M\rho^n \qquad (0<\rho<\infty). \tag{7.1}$$

It is very easy to show that if $f \in L$ then T_λ is likewise in L and $\|T_\lambda\|_1 \leqslant MA_{n,\lambda}\|f\|_1$; in particular the integral

$$\int_{E_n} \frac{f(y)\,\delta^\lambda(y)\,\phi(x-y)}{|x-y|^{n+\lambda}}\,dy$$

generalizing J_λ, is finite almost everywhere in P. Whether, however, $f \in L^p$, $1<p<\infty$, implies $T_\lambda \in L^p$ seems to be an open problem.

References

1. A. P. Calderón and A. Zygmund, On singular integrals. *Am. J. Math.* **78** (1956), 289–309.

2. L. Carleson, On convergence and growth of partial sums of Fourier series. *Acta Math.* **116** (1966), 135–157.

3. G. H. Hardy and J. E. Littlewood, On the strong summability of Fourier series. *Fund. Math.* **25** (1935), 162–189.

4. J. Marcinkiewicz, Sur la sommabilité forte des séries de Fourier. *J. London Math. Soc.* **14** (1939), 162–168.

5. J. Marcinkiewicz, Sur les séries de Fourier. *Fund. Math.* **27** (1936), 38–69.

6. J. Marcinkiewicz, On the convergence of Fourier series. *J. London Math. Soc.* **10** (1935), 264–268.

7. J. Marcinkiewicz, Sur quelques intégrales du type de Dini. *Ann. Soc. Polon. Math.* **17** (1938) 42–50.

8. E. M. Ostrow and E. M. Stein, A generalization of lemmas of Marcinkiewicz and Fine with applications to singular integrals. *Ann. Scuola Normale Sup. Pisa* **11** (1957), 117–135.

9. N. Wiener, The ergodic theorem. *Duke Math. J.* **5** (1935), 1–18.

10. A. Zygmund, "Trigonometric Series," Vols. I and II. Cambridge University Press, 1959.