sufficiently small values of $x-a$, since $D^+f(a) \geqslant 0$. Hence $\phi(x) = 0$ for some values of x between a and b. Let ξ be the greatest such value. Then $D^+\phi(\xi) \leqslant 0$, $D^+f(\xi)+\epsilon \leqslant 0$, contrary to hypothesis. Hence $f(b)-f(a) \geqslant -\epsilon(b-a)$ for every positive ϵ, and the result follows.]

(v) The derivates and incrementary ratios of a continuous function have the same bounds in any interval; i.e. if any one of the derivates satisfies $\alpha \leqslant Df \leqslant \beta$, then $\alpha \leqslant \{f(x_2)-f(x_1)\}/(x_2-x_1) \leqslant \beta$, and conversely.

[Consider $\phi(x) = f(x)-\alpha x$, and use the previous example.]

(vi) If one of the derivates of $f(x)$ is continuous at a certain point, then $f(x)$ has a differential coefficient at the point.

11.4. Functions of bounded variation. We say that $f(x)$ is of bounded variation in (a, b) if, in this interval, it can be expressed in the form $\phi(x)-\psi(x)$, where ϕ and ψ are non-decreasing bounded functions.

It is easily seen that the sum, difference, or product of two functions of bounded variation is also of bounded variation.

An alternative definition is obtained by assuming that, if the interval (a, b) is divided up by points $a = x_0 < x_1 < \ldots < x_n = b$, then

$$\sum_{\nu=0}^{n-1} |f(x_{\nu+1})-f(x_\nu)|$$

is less than a constant independent of the mode of division. The upper bound of these sums is called the total variation.

It is easily seen that, if the first condition holds, then so does the second. For

$$|f(x_{\nu+1})-f(x_\nu)| \leqslant \phi(x_{\nu+1})-\phi(x_\nu)+\psi(x_{\nu+1})-\psi(x_\nu),$$

so that

$$\sum_{\nu=0}^{n-1} |f(x_{\nu+1})-f(x_\nu)| \leqslant \phi(b)-\phi(a)+\psi(b)-\psi(a).$$

To prove the converse, let p be the sum of those differences $f(x_{\nu+1})-f(x_\nu)$ which are positive, $-n$ the sum of those which are negative. Then, if v is the sum $\sum |f(x_{\nu+1})-f(x_\nu)|$, we have

$$v = p+n, \qquad f(b)-f(a) = p-n,$$

and so

$$v = 2p+f(a)-f(b), \qquad v = 2n+f(b)-f(a).$$

Hence, if v is bounded for all modes of division, so are p and n. Let V, P, and N be the upper bounds of v, p, and n. Then

$$V = 2P+f(a)-f(b), \qquad V = 2N+f(b)-f(a).$$

Let $V(x)$, $P(x)$, and $N(x)$ be the corresponding numbers for the interval (a, x). They are obviously bounded non-decreasing functions of x; and

$$V(x) = 2P(x)+f(a)-f(x), \qquad V(x) = 2N(x)+f(x)-f(a),$$

so that
$$f(x) = f(a)+P(x)-N(x).$$

This is the required expression for $f(x)$.

The functions $V(x)$, $P(x)$, and $N(x)$ are called the total variation and the positive and negative variations of $f(x)$ in (a, x).

If $f(x)$ is continuous and of bounded variation, its variation $V(x)$ is continuous. We can find a mode of division of the interval (a, x), with a point of division x' as near x as we please, such that

$$v > V(x)-\epsilon$$

and also
$$|f(x)-f(x')| < \epsilon.$$

Let
$$v' = v-|f(x)-f(x')|.$$

Then v' is a sum corresponding to the interval (a, x'), and so

$$V(x') \geqslant v' > V(x)-2\epsilon.$$

Since $V(x')$ is non-decreasing, it follows that $V(x') \to V(x)$ as $x' \to x$ from below. Similarly $V(x') \to V(x)$ as $x' \to x$ from above. Hence $V(x)$ is continuous.

A continuous function of bounded variation is the difference between two continuous non-decreasing functions. For if $f(x)$ is continuous, so are $P(x)$ and $N(x)$.

11.41. The differential coefficient of a function of bounded variation. The object of the next three sections is to prove that *a function of bounded variation has a finite differential coefficient almost everywhere.*

Our proof depends on the following lemmas, due to Sierpinski.* They are of the same type as the Heine-Borel theorem, but apply to sets which need not even be measurable.

LEMMA 1. *Suppose that each point x of a set E in (a,b) is the left-hand end-point of one or more intervals $(x, x+h_x)$ of a family H. Then there is a finite non-overlapping set S of intervals of H which includes a sub-set E' of E such that $m_e(E') > m_e(E)-\epsilon$.*

Let E_n be the set of points of E which are associated with some $h_x > 1/n$. Then E is the outer limiting set of the sets E_n, we have $\lim m_e(E_n) = m_e(E)$ (§ 10.29), and we can take n so large that $m_e(E_n) > m_e(E)-\frac{1}{2}\epsilon$.

* Sierpinski (1). A similar lemma is given by W. H. and G. C. Young (1).

Let a_1 be the lower bound of E_n, b_1 its upper bound, and let $l = b_1 - a_1$. Let $\eta = \frac{1}{2}\epsilon/(nl+1)$. Then there is a point x_1 of E_n such that $a_1 \leqslant x_1 < a_1 + \eta$. Let $(x_1, x_1 + h_1)$ be an associated interval for which $h_1 > 1/n$.

If there are points of E_n to the right of $x_1 + h_1$, let a_2 be their lower bound. Then there is a point x_2 of E_n in $(a_2, a_2 + \eta)$. Let $(x_2, x_2 + h_2)$ be an associated interval with $h_2 > 1/n$.

Continuing the process, we reach b_1 in a finite number of steps, since each step takes us at least $1/n$ nearer to it. In fact, if there are N steps, then $(N-1)/n < l$, i.e. $N < nl + 1$.

Let S denote the set of intervals $(x_\nu, x_\nu + h_\nu)$ so constructed, and T the set of intervals $(x_\nu - \eta, x_\nu)$. Then $E_n \prec S + T$, and $m(T) < N\eta < \frac{1}{2}\epsilon$. Hence

$$m_e(E) - \tfrac{1}{2}\epsilon < m_e(E_n) \leqslant m_e(E_n S) + m_e(E_n T) < m_e(E_n S) + \tfrac{1}{2}\epsilon,$$

and the set $E' = E_n S$ has the required property.

LEMMA 2. *Suppose in addition that for every x there are arbitrarily small intervals $(x, x + h_x)$. Then we may conclude in addition that*

$$m(S) < m_e(E) + \epsilon.$$

The additional condition is necessary; we might, for example, take E to be a single point x, and associate with it the interval $(x, x+1)$. Then Lemma 1 would hold, but not Lemma 2.

Let O be an open set containing E, such that

$$m(O) < m_e(E) + \epsilon.$$

Let H_1 be the sub-class of the family of intervals H consisting of those intervals that lie in O. In view of the additional condition imposed in Lemma 2, every point of E is the left-hand end-point of one or more intervals of H_1. We can now apply Lemma 1 with H replaced by H_1. We obtain a new set of intervals S which has the same property as that constructed in the proof of Lemma 1. But now S is a set of non-overlapping intervals included in O. Hence

$$m(S) \leqslant m(O) < m_e(E) + \epsilon.$$

This proves the lemma.

In these lemmas the intervals of which S consists may be regarded as either open or closed, whichever is most convenient in any particular case. For if the result has been obtained with S consisting of closed intervals, we can replace them by open intervals by removing a finite number of points, i.e. a set of measure zero. This clearly does not affect the result.

LEMMA 3. *We may suppose S in the above construction to be included in any given set of intervals G which contains E.*

For we may replace O by OG in the construction.

11.42. *If* $f(x)$ is non-decreasing in (a,b), it has almost everywhere in (a,b) a differential coefficient $f'(x)$.*

Let E be the set where $D_+f < D^+f$. We shall first prove that $m_e(E) = 0$.

Now E is the sum of the sets $E(u,v)$ where

$$D_+f < u < v < D^+f,$$

u and v running through all rational numbers ($u < v$). Hence it is sufficient to prove that $m_e\{E(u,v)\} = 0$ for every pair of such numbers.

Suppose on the contrary that one of these sets $E(u,v)$ has a positive exterior measure, say μ. Every point x of it is the left-hand end-point of arbitrarily small intervals $(x, x+h)$ for which

$$f(x+h)-f(x) < hu.$$

Hence by Lemma 2 there is a finite set S of such intervals, containing a part E' of $E(u,v)$ such that $m_e(E') > \mu-\epsilon$, and such that $\sum_1 h < \mu+\epsilon$, where $\sum_1$ denotes a summation over S. Hence

$$\sum_1 \{f(x+h)-f(x)\} < u \sum_1 h < u(\mu+\epsilon).$$

Again, every point of E' is the left-hand end-point of intervals $(x, x+k)$ such that

$$f(x+k)-f(x) > kv,$$

and by Lemma 3 there is a finite set of these intervals, included in S and of measure greater than $m_e(E')-\epsilon > \mu-2\epsilon$. If $\sum_2$ denotes a summation over these intervals,

$$\sum_2 \{f(x+k)-f(x)\} > v \sum_2 k > v(\mu-2\epsilon).$$

But since $f(x)$ is non-decreasing, and the k-intervals are included in the h-intervals,

$$\sum_2 \{f(x+k)-f(x)\} \leqslant \sum_1 \{f(x+h)-f(x)\}.$$

Hence $v(\mu-2\epsilon) < u(\mu+\epsilon)$, which is false if ϵ is small enough. Hence $f'_+(x)$ (and similarly $f'_-(x)$) exists almost everywhere.

* This proof is due to Rajchman and Saks (1).

Further, we can argue in the above way with D^+ replaced by D^-; every point of E' is then the right-hand end-point of arbitrarily small intervals $(x-k, x)$ such that $f(x)-f(x-k) > kv$, and the conclusion follows as before. Hence almost everywhere $D_+ f \geqslant D^- f$, i.e. almost everywhere $f'_+(x) \geqslant f'_-(x)$. Similarly we can prove the reversed inequality, and the result follows.

11.43. There is a more general theorem on the possible sets where $f'_-(x) \neq f'_+(x)$, and the result has nothing to do with monotony.

The set of points where the right-hand and left-hand derivatives of any function exist and are different is enumerable.

Let E be the set where $f'_-(x) < f'_+(x)$, and let all rational numbers be arranged in a sequence $r_1, r_2, \ldots$. Then if x is a point of E, there is a smallest integer k such that

$$f'_-(x) < r_k < f'_+(x).$$

There is then a smallest integer m such that $r_m < x$, and such that

$$\{f(\xi)-f(x)\}/(\xi-x) < r_k$$

for $r_m < \xi < x$; and a smallest integer n such that $r_n > x$, and

$$\{f(\xi)-f(x)\}/(\xi-x) > r_k$$

for $x < \xi < r_n$. The two inequalities together give

$$f(\xi)-f(x) > r_k(\xi-x) \qquad (r_m < \xi < r_n,\ \xi \neq x). \tag{1}$$

Thus to every x corresponds a unique triad of numbers (k, m, n); and no two values of x correspond to the same triad; for if x_1 and x_2 correspond to the same triad, we have, on putting $x = x_1$, $\xi = x_2$ in (1), $f(x_2)-f(x_1) > r_k(x_2-x_1)$, and, on putting $x = x_2$, $\xi = x_1$, the same inequality reversed.

Since the set of triads (k, m, n) is enumerable, it follows that E is enumerable or finite. This is the required result. Since the measure of an enumerable set is zero, this theorem can be used to give an alternative ending to the proof of the theorem of the previous section.

11.5. Integrals. A function which is the Lebesgue indefinite integral of another function is called *an integral*.

An integral is continuous. For if $F(x)$ is the integral of $f(x)$, then

$$F(x+h)-F(x) = \int_x^{x+h} f(t)\,dt,$$

which tends to 0 with h, by § 10.73 (v).

The integral of a positive function is a non-decreasing function. For if $f(x) \geqslant 0$, $h > 0$,

$$F(x+h)-F(x) = \int\limits_{x}^{x+h} f(t)\, dt \geqslant 0.$$

An integral is a function of bounded variation. For let

$$F(x) = F(a) + \int\limits_{a}^{x} f(t)\, dt,$$

and let $f_1(x) = f(x)$ where $f(x) \geqslant 0$, and $f_1(x) = 0$ elsewhere, and $-f_2(x) = f(x) - f_1(x)$. Then $f_1(x) \geqslant 0$, $f_2(x) \geqslant 0$, and

$$F(x) = F(a) + \int\limits_{a}^{x} f_1(t)\, dt - \int\limits_{a}^{x} f_2(t)\, dt$$
$$= F(a) + F_1(x) - F_2(x),$$

where $F_1(x)$ and $F_2(x)$ are bounded non-decreasing functions.

11.51. Differentiation of the indefinite integral. Let $f(x)$ be integrable over (a, b), and let

$$F(x) = \int\limits_{a}^{x} f(t)\, dt.$$

Since $F(x)$ is a function of bounded variation, it has a finite differential coefficient $F'(x)$ almost everywhere. Our next object is to prove that $F'(x) = f(x)$ almost everywhere.

11.52. The proof depends on the following lemma.

If $\phi(x)$ is integrable, and $\int\limits_{a}^{x} \phi(t)\, dt = 0$ for all values of x in (a, b), then $\phi(x) = 0$ for almost all values of x in (a, b).

If this is not so, then either $\phi(x) > 0$ in a set of positive measure, or $\phi(x) < 0$ in a set of positive measure—suppose, for example, the former. Any set of positive measure contains a closed set of positive measure, since its complement can be included in an open set less than the whole interval. Hence $\phi(x) > 0$ in a closed set of positive measure—say E.

Now the integral of ϕ over any interval is zero; hence, by § 10.71, the integral over any open set is zero. Hence the integral over any closed set is zero, and in particular

$$\int_E \phi(x)\, dx = 0.$$

Hence, by § 10.73, $\phi(x) = 0$ almost everywhere in E, contrary to hypothesis. This proves the lemma.

11.53. *If $f(x)$ is bounded, and $F(x)$ is its integral, then $F'(x) = f(x)$ almost everywhere.*

Let $|f(x)| \leqslant M$. Then

$$\left|\frac{F(x+h)-F(x)}{h}\right| = \left|\frac{1}{h}\int_x^{x+h} f(t)\,dt\right| \leqslant M,$$

and

$$\lim_{h\to 0}\frac{F(x+h)-F(x)}{h} = F'(x)$$

almost everywhere. Hence, by the theorem of bounded convergence,* as $h \to 0$,

$$\int_a^x \frac{F(t+h)-F(t)}{h}\,dt \to \int_a^x F'(t)\,dt.$$

But the left-hand side is equal to

$$\frac{1}{h}\int_{a+h}^{x+h} F(t)\,dt - \frac{1}{h}\int_a^x F(t)\,dt = \frac{1}{h}\int_x^{x+h} F(t)\,dt - \frac{1}{h}\int_a^{a+h} F(t)\,dt,$$

which tends to $F(x)-F(a)$, since F is continuous. Hence

$$\int_a^x F'(t)\,dt = F(x)-F(a), \tag{1}$$

i.e.

$$\int_a^x \{F'(t)-f(t)\}\,dt = 0, \tag{2}$$

for all values of x. The result now follows from the lemma.

11.54. To extend the theorem to unbounded functions, we require another lemma.

If $\phi(x)$ is continuous and non-decreasing in (a, b), then $\phi'(x)$ is integrable, and

$$\int_a^b \phi'(x)\,dx \leqslant \phi(b)-\phi(a).$$

For $\{\phi(x+h)-\phi(x)\}/h \geqslant 0$, and $\{\phi(x+h)-\phi(x)\}/h$ tends to

* To apply the theorem as given in § 10.5, we make $h \to 0$ through an enumerable sequence; so also in the next section.

$\phi'(x)$ almost everywhere as $h \to 0$. Hence, by Fatou's theorem (§ 10.81),

$$\varliminf_{h\to 0} \int_a^b \frac{\phi(x+h)-\phi(x)}{h}\,dx \geqslant \int_a^b \phi'(x)\,dx.$$

Also, since ϕ is continuous, the left-hand side is equal to $\phi(b)-\phi(a)$, as in the above proof. Hence the result.

11.55. *If $f(x)$ is any integrable function, $F'(x) = f(x)$ almost everywhere.*

We may as usual suppose that $f(x) \geqslant 0$. We define $\{f(x)\}_n$ as in § 10.7. Since $f(t)-\{f(t)\}_n \geqslant 0$, the function

$$\int_a^x [f(t)-\{f(t)\}_n]\,dt$$

is non-decreasing, so that its differential coefficient is never negative. Hence

$$\frac{d}{dx}\left\{\int_a^x f(t)\,dt\right\} \geqslant \frac{d}{dx}\left\{\int_a^x \{f(t)\}_n\,dt\right\}$$

wherever these differential coefficients exist. Hence, by the theorem for bounded functions, $F'(x) \geqslant \{f(x)\}_n$ almost everywhere. Making $n \to \infty$ we see that $F'(x) \geqslant f(x)$ almost everywhere. Hence

$$\int_a^b F'(x)\,dx \geqslant \int_a^b f(x)\,dx.$$

The above lemma, however, gives this inequality reversed. Hence in fact the two sides are equal, i.e.

$$\int_a^b \{F'(x)-f(x)\}\,dx = 0.$$

Since the integrand is never negative, it must be zero almost everywhere. This is the required result.

11.6. The Lebesgue set. The theorem that $F'(x) = f(x)$ almost everywhere was extended by Lebesgue as follows.

If $f(x)$ is integrable,

$$\lim_{h\to 0} \frac{1}{h} \int_x^{x+h} |f(t)-\alpha|\,dt = |f(x)-\alpha|$$

for all values of α, except when x belongs to a set of measure zero; that is, $|f(x)-\alpha|$ is the derivative of its indefinite integral for all values of α and almost all values of x.

If α were fixed there would be nothing to prove, since $|f(x)-\alpha|$ is integrable, and the result follows from the above fundamental theorem.

Consider next all rational values of α, say $\alpha_1, \alpha_2,\ldots$. The sets in which the theorem is false for $\alpha_1, \alpha_2,\ldots$ are all of measure zero, and so their aggregate is of measure zero. Hence $|f(x)-\alpha|$ is the derivative of its integral for all rational values of α, except when x belongs to a set E of measure zero.

Now let x be a point not in E, α an irrational number, and β a rational number near to α. Since

$$\big||f(t)-\alpha|-|f(t)-\beta|\big| \leqslant |\beta-\alpha|.$$

we have

$$\left|\frac{1}{h}\int_x^{x+h} |f(t)-\alpha|\,dt - \frac{1}{h}\int_x^{x+h} |f(t)-\beta|\,dt\right| \leqslant |\beta-\alpha|.$$

But
$$\left|\frac{1}{h}\int_x^{x+h} |f(t)-\beta|\,dt - |f(x)-\beta|\right| \leqslant \epsilon$$

if $|h| < h_0(\beta,\epsilon)$. Hence

$$\begin{aligned}
&\left|\frac{1}{h}\int_x^{x+h} |f(t)-\alpha|\,dt - |f(x)-\alpha|\right| \\
&\qquad \leqslant \left|\frac{1}{h}\int_x^{x+h} |f(t)-\alpha|\,dt - \frac{1}{h}\int_x^{x+h} |f(t)-\beta|\,dt\right| + \\
&\qquad\quad + \left|\frac{1}{h}\int_x^{x+h} |f(t)-\beta|\,dt - |f(x)-\beta|\right| + \big||f(x)-\beta|-|f(x)-\alpha|\big| \\
&\qquad \leqslant |\beta-\alpha|+\epsilon+|\beta-\alpha|,
\end{aligned}$$

which may be made as small as we please, by choice first of β and then of ϵ. Hence $|f(x)-\alpha|$ is also the derivative of its indefinite integral for all irrational α, if x is not a point of E. This proves the theorem.

We may, in particular, take $\alpha = f(x)$. Hence

$$\int_0^h |f(x+t)-f(x)|\,dt = o(h)$$

as $h \to 0$, for almost all values of x. The set where this holds is called the *Lebesgue set.*

All points of continuity are of course included in the Lebesgue set.

The interest of the Lebesgue set lies in the fact that many theorems which hold at all points of continuity are also found to hold at all points of the Lebesgue set, and so almost everywhere. We shall have examples of this in the chapter on Fourier series.

We note finally that if the modulus sign is omitted from the formula, the α disappears, and the result reduces to the previous theorem.

11.7. Absolutely continuous functions. *A function $f(x)$ is said to be absolutely continuous in an interval (a, b) if, given ϵ, we can find δ such that*

$$\sum_{\nu=1}^{n} |f(x_\nu+h_\nu)-f(x_\nu)| \leqslant \epsilon$$

for every set of non-overlapping intervals $(x_\nu, x_\nu+h_\nu)$ such that $\sum h_\nu \leqslant \delta$.

An absolutely continuous function is continuous, since we can take the above sum to consist of one term only.

An absolutely continuous function is of bounded variation, since its total variation over an interval of length δ is at most ϵ, and consequently its total variation over (a, b) is at most $(b-a)\epsilon/\delta$.

On the other hand, there are continuous functions of bounded variation which are not absolutely continuous. An example of such a function will be given in § 11.72.

11.71. *A necessary and sufficient condition that a function should be an integral is that it should be absolutely continuous.*

If $F(x)$ is the integral of $f(x)$,

$$\sum_{\nu=1}^{n} |F(x_\nu+h_\nu)-F(x_\nu)| \leqslant \sum_{\nu=1}^{n} \int_{x_\nu}^{x_\nu+h_\nu} |f(x)|\,dx = \int_E |f(x)|\,dx,$$

where E denotes the set of intervals $(x_\nu, x_\nu+h_\nu)$. The right-hand side tends to zero with $\sum h_\nu$, in the sense of the above definition, by § 10.73 (v). Hence $F(x)$ is absolutely continuous.

To prove the converse we require the following lemma.

If $\phi(x)$ is absolutely continuous in (a, b), and $\phi'(x) = 0$ almost everywhere, then $\phi(x)$ is a constant.

Let E be the set where $\phi'(x) = 0$. Every point x of E is the left-hand end-point of arbitrarily small intervals $(x, x+h)$, such that

$$|\phi(x+h)-\phi(x)| < \epsilon h.$$

By the lemmas of § 11.41, we can select a finite set S of these intervals which do not overlap, and which contain all E except a set of measure δ, and so all (a, b) except a set of measure δ.

Let x_1, x_2,... be the end-points of the intervals of S, and let $\sum_1$ denote a summation over the intervals of S, and $\sum_2$ over the complementary intervals. Then

$$|\phi(b)-\phi(a)| \leqslant \sum_1 |\phi(x_{\nu+1})-\phi(x_\nu)| + \sum_2 |\phi(x_{\nu+1})-\phi(x_\nu)|.$$

Now $\quad \sum_1 |\phi(x_{\nu+1})-\phi(x_\nu)| < \epsilon \sum_1 (x_{\nu+1}-x_\nu) < \epsilon(b-a).$

Also $\sum_2 (x_{\nu+1}-x_\nu) < \delta$, and so, by the property of absolute continuity,

$$\sum_2 |\phi(x_{\nu+1})-\phi(x_\nu)|$$

tends to zero with δ. Hence, making $\delta \to 0$,

$$|\phi(b)-\phi(a)| \leqslant \epsilon(b-a).$$

Making $\epsilon \to 0$, it follows that $\phi(b) = \phi(a)$; and similarly $\phi(x) = \phi(a)$ for every value of x.

Suppose now that $F(x)$ is any absolutely continuous function. Then it is continuous and of bounded variation, and we may write

$$F(x) = F_1(x) - F_2(x),$$

where F_1 and F_2 are continuous non-decreasing functions. By the lemma of § 11.54, $F_1'(x)$ and $F_2'(x)$ are integrable, and hence so is $F'(x)$. Hence

$$\int_a^x F'(t)\,dt$$

is absolutely continuous, and so also is

$$\phi(x) = F(x) - \int_a^x F'(t)\,dt.$$

But $\phi'(x)=0$ almost everywhere. Hence, by the lemma, $\phi(x)$ is a constant, i.e.

$$F(x)-\int_a^x F'(t)\,dt = F(a).$$

Thus $F(x)$ is the integral of $F'(x)$.

11.72. A continuous increasing function which is not an integral.* We can define a function of this type by means of Cantor's ternary set (§ 10.291).

Let a_n always take the values 0 or 2, and let $b_n = \frac{1}{2}a_n$, so that b_n is always 0 or 1. If

$$x = \cdot a_1 a_2 a_3 \ldots (3)$$

is a point of Cantor's set E, we define

$$f(x) = \cdot b_1 b_2 b_3 \ldots (2)$$

(in the scale of 2).

At the ends of an interval δ_{pk}, $f(x)$ therefore has the values

$$\cdot b_1 \ldots b_m 0111 \ldots (2), \qquad \cdot b_1 \ldots b_m 1000 \ldots (2),$$

and these are equal. We define $f(x)$ throughout the interval δ_{pk} to be equal to its value at the end-points.

The function $f(x)$ is non-decreasing. In proving this it is sufficient to consider points x of E, since $f(x)$ is constant in the intervals of CE. Let

$$x' = \cdot a'_1 a'_2 \ldots (3), \qquad x'' = \cdot a''_1 a''_2 \ldots (3)$$

be points of E, $x'' > x'$. Then there is a suffix n such that $a'_m = a''_m$ $(m < n)$, $a'_n < a''_n$. Hence

$$f(x') = \cdot b'_1 \ldots b'_{n-1} b'_n \ldots (2) \leqslant \cdot b''_1 \ldots b''_{n-1} b''_n \ldots (2) = f(x'').$$

The function $f(x)$ is continuous. We have to prove that $f(x') \to f(x)$ as $x' \to x$, and again it is sufficient to consider points x, x' of E. Let

$$x = \cdot a_1 a_2 \ldots (3), \qquad x' = \cdot a'_1 a'_2 \ldots (3).$$

If now $x' \to x$, there will be a value of n, which tends to infinity as $x' \to x$, such that $a_m = a'_m$ $(m < n)$. Hence

$$f(x)-f(x') = \cdot 00\ldots 0 b_n \ldots - \cdot 00 \ldots 0 b'_n \ldots \to 0.$$

On the other hand,

$$\int_0^1 f'(x)\,dx \neq f(1)-f(0).$$

* A detailed discussion of this function is given by Hille and Tamarkin (1),

For the right-hand side is 1, since $f(1) = \cdot 111...(2) = 1$, $f(0) = 0$; but $f(x)$ is constant in the intervals δ_{pk}, so that $f'(x) = 0$ in the interior of any of these intervals. Hence $f'(x) = 0$ almost everywhere, and the left-hand side is 0.

It follows that $f(x)$ is not the integral of its differential coefficient, and so is not absolutely continuous. It is easy to see this directly. Consider the sum

$$\sum |f(\beta_k) - f(\alpha_k)|$$

taken over the intervals (α_k, β_k) which remain after the pth step of removing intervals δ_{pk}. It is equal to

$$\sum \{f(\beta_k) - f(\alpha_k)\} = f(1) - f(0) = 1.$$

But $$\sum (\beta_k - \alpha_k) = 1 - \frac{1}{3} - \frac{2}{9} - \ldots - \frac{2^{p-1}}{3^p} = \left(\frac{2}{3}\right)^p,$$

which tends to zero as $p \to \infty$. Hence $f(x)$ is not absolutely continuous.

11.8. Integration of a differential coefficient. If $f(x)$ has a differential coefficient almost everywhere, or even everywhere, in an interval (a, b), the formula

$$\int_a^x f'(t)\,dt = f(x) - f(a) \qquad (a \leqslant x \leqslant b) \tag{1}$$

is not necessarily true. It may fail in one or other of two ways. Consider, for example, the function

$$f(x) = x^2 \sin\frac{1}{x^2} \quad (x > 0), \qquad f(0) = 0,$$

already referred to in § 10.7. Here

$$f'(x) = 2x \sin\frac{1}{x^2} - \frac{2}{x}\cos\frac{1}{x^2} \quad (x > 0), \qquad f'(0) = 0,$$

so that $f'(x)$ exists everywhere; but, as we saw in § 10.7, it is not integrable in the Lebesgue sense, so that (1), on the Lebesgue theory, has no meaning.

If we can imagine a function with this kind of singularity distributed everywhere in an interval, we shall obtain some idea of the nature of the problem of integrating a differential coefficient. The problem has been solved by means of the Denjoy integral. This is a highly general type of non-absolutely convergent integral, and it would take us too far to discuss its properties here. The result is that, *if* $f'(x)$ *exists everywhere, the formula* (1) *is true, the integral being a Denjoy integral.*

If we do not assume that $f'(x)$ exists everywhere, but merely almost everywhere, the formula (1) may break down still more completely. The integral on the left may exist as a Lebesgue integral, but be unequal to the right-hand side. We have already had an example of this in § 11.72—in fact an example where $f'(x) = 0$ almost everywhere, without $f(x)$ being a constant.

In order to obtain the formula (1), the integral being a Lebesgue integral, we have therefore to impose further conditions on $f(x)$ or on $f'(x)$. There are several theorems, varying in difficulty according to what is assumed. Their common feature is that we suppose that $f'(x)$ exists everywhere. The example of § 11.72 shows that no set of conditions which is merely given almost everywhere is sufficient.

11.81. *If* $f'(x)$ *exists everywhere and is bounded, then* 11.8 (1) *is true.*

If $|f'(x)| \leqslant M$, then (*P.M.* § 125) there is a number θ between 0 and 1 such that

$$\left|\frac{f(x+h)-f(x)}{h}\right| = |f'(x+\theta h)| \leqslant M. \tag{1}$$

Hence $\{f(x+h)-f(x)\}/h$ converges boundedly to $f'(x)$, and the proof is now the same as that of 11.53 (1) (with $f(x)$ instead of $F(x)$).

Alternatively, we may observe that it follows from (1) that

$$\sum_{\nu=1}^{n} |f(x_\nu+h_\nu)-f(x_\nu)| \leqslant M \sum_{\nu=1}^{n} h_\nu.$$

Hence $f(x)$ is absolutely continuous, and the required result follows from § 11.71.

11.83. *If* $f(x)$ *is any function such that* $f'(x)$ *is finite everywhere and is integrable, then* 11.8 (1) *is true.*

This evidently shows in particular that 11.8 (1) holds if $f(x)$ is of bounded variation and $f'(x)$ is finite everywhere; for if $f(x)$ is of bounded variation, $f'(x)$ is integrable (see § 11.54 and example 12 below).

The following proof is substantially that given by Schlesinger and Plessner.† It depends on the two following lemmas.

LEMMA 1. *Let E be any set in (a,b) of measure zero, ϵ a given positive number. Then there is a non-decreasing absolutely continuous function $\chi(x)$ such that $\chi'(x) = +\infty$ in E, and $\chi(b)-\chi(a) < \epsilon$.*

We can include E in a sequence of open sets $O_1 > O_2 > \dots$ such that $m(O_n) < \epsilon_n$, $\epsilon_1+\epsilon_2+\dots = \epsilon$. Let $f_n(x)$ be the characteristic function of the set O_n. Then

$$\int_a^b f_n(t)\,dt = m(O_n) < \epsilon_n.$$

Let
$$\phi_n(x) = f_1(x)+f_2(x)+\dots+f_n(x).$$

Then $\phi_n(t)$ is non-decreasing as $n \to \infty$ for every t, and

$$\int_a^x \phi_n(t)\,dt < \epsilon_1+\epsilon_2+\dots+\epsilon_n \leqslant \epsilon.$$

Hence by § 10.82 $\phi_n(t)$ tends to a finite limit $\phi(t)$ almost everywhere, and

$$\lim_{n\to\infty} \int_a^x \phi_n(t)\,dt = \int_a^x \phi(t)\,dt = \chi(x),$$

say.

This function $\chi(x)$ has the required properties. Since it is the integral of a non-negative function it is non-decreasing and absolutely continuous, and

$$\chi(b)-\chi(a) = \int_a^b \phi(t)\,dt < \epsilon.$$

Also
$$\frac{d}{dx}\int_a^x f_\nu(t)\,dt = 1$$

in O_ν, and so, if $\chi_n(x) = \int_a^x \phi_n(t)\,dt$,

$$\chi_n'(x) = \sum_{\nu=1}^{n} \frac{d}{dx}\int_a^x f_\nu(t)\,dt = n$$

† *Lebesguesche Integrale*, pp. 166–74.

in O_n. Hence

$$\frac{\chi(x+h)-\chi(x)}{h} \geqslant \frac{\chi_n(x+h)-\chi_n(x)}{h} > n-\delta$$

for $|h| < h_0(\delta)$ and x in O_n. Hence $D\chi \geqslant n$ for each of the four derivates and x in O_n. Since a point of E belongs to O_n for every n, it follows that $\chi'(x) = +\infty$ in E.

LEMMA 2. *If $f(x)$ is continuous in (a,b), and $D^+f \geqslant 0$ almost everywhere in the interval, and D^+f is nowhere $-\infty$, then $f(x)$ is a non-decreasing function.*

It is sufficient to prove that $f(b) \geqslant f(a)$, since the general result then follows by a similar argument.

Let E be the set of measure zero where $D^+f < 0$. By Lemma 1 there is an absolutely continuous function $\chi(x)$ such that $\chi'(x) = +\infty$ in E, and $\chi(b)-\chi(a) < \epsilon$.

Let $$g(x) = f(x)+\chi(x).$$

Then in E, $D^+g = +\infty$, since $D_+\chi = +\infty$ and D^+f is finite, and $D^+g \geqslant D_+\chi + D^+f$. Also in CE

$$D^+g \geqslant D^+f \geqslant 0$$

since χ is non-decreasing. Hence $D^+g \geqslant 0$ everywhere, and so, by § 11.3, ex. (iv), $g(b) \geqslant g(a)$. Hence

$$f(b)-f(a) \geqslant -\{\chi(b)-\chi(a)\} > -\epsilon,$$

and, making $\epsilon \to 0$, the result follows.

11.84. We can now prove the theorem stated in § 11.83. Let n be any positive number, and let

$$g_n(x) = \min\{f'(x), n\}, \qquad G_n(x) = \max\{f'(x), -n\}.$$

Then $g_n(x) \leqslant f'(x) \leqslant G_n(x)$, and, since $f'(x)$ is integrable, so are $g_n(x)$ and $G_n(x)$. Let

$$f_n(x) = \int_a^x g_n(t)\,dt, \qquad F_n(x) = \int_a^x G_n(t)\,dt.$$

Then $$\lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} F_n(x) = \int_a^x f'(t)\,dt = \phi(x),$$

say. Now $$D^+\{F_n(x)-f(x)\} \geqslant D^+F_n - D^+f.$$

This is almost everywhere equal to $G_n(x)-f'(x)$, i.e.

$$D^+\{F_n(x)-f(x)\} \geqslant 0$$

almost everywhere. Also

$$\frac{F_n(x+h)-F_n(x)}{h} \geqslant \frac{1}{h}\int_x^{x+h} (-n)\,dt = -n,$$

so that $D^+F_n \geqslant -n$, and so $D^+(F_n-f)$ is nowhere $-\infty$. Hence, by Lemma 2, $F_n(x)-f(x)$ is non-decreasing, i.e.

$$F_n(x)-f(x) \geqslant F_n(a)-f(a) = -f(a).$$

Making $n \to \infty$, we obtain

$$\phi(x) \geqslant f(x)-f(a).$$

A similar argument with $f_n(x)$ gives the reversed inequality, and this proves the theorem.

MISCELLANEOUS EXAMPLES

1. For $x = \frac{1}{3}$ the function $f(x)$ of § 11.23 has the derivative $+\infty$.

2. The density of a set E at a point x may be defined as

$$\lim_{h\to 0} \frac{m(EH)}{2h},$$

where H is the interval $(x-h, x+h)$.

Prove that the density of a set is 1 almost everywhere in the set, and 0 almost everywhere outside it.

[Consider the integral of the characteristic function of E.]

3. A set E in $(0, 1)$ is such that, if (α, β) is any interval, then

$$m\{E(\alpha, \beta)\} \geqslant \delta(\beta-\alpha)$$

where $\delta > 0$. Show that $m(E) = 1$.

4. If, as $h \to 0$,

$$\int_a^b |f(x+h)-f(x)|\,dx = o(h),$$

then $f(x)$ is almost everywhere equal to a constant.

[Consider $\int_{x_1}^{x_2} \{f(x+h)-f(x)\}\,dx$. See Titchmarsh (7), where, however, the proof is unnecessarily complicated.]

5. Let α and β be positive numbers, $f(x) = x^{\alpha}\sin x^{-\beta}$ $(0 < x \leqslant 1)$, and $f(0) = 0$. Then $f(x)$ is of bounded variation in $(0, 1)$ if $\alpha > \beta$, but not if $\alpha \leqslant \beta$.

6. A function $f(x)$, defined for $0 \leqslant x < 1$, is absolutely continuous in every interval $(0, \xi)$, where $\xi < 1$, and its total variation in $(0, \xi)$ is bounded as $\xi \to 1$. Show that $f(x)$ tends to a limit as $\xi \to 1$, and that, if we define $f(1)$ to be equal to this limit, then $f(x)$ is absolutely continuous in the whole interval $(0, 1)$.

[The point of this example is that the difference between 'continuity plus bounded variation' and absolute continuity is a property of a whole interval, and cannot be traced to the behaviour of the function in the neighbourhood of any one point.]

7. The theorem of § 11.83 remains true if $f'(x) = +\infty$ in an enumerable set.

8. A necessary and sufficient condition that a function should be convex in an interval (a, b), in the sense of § 5.31, is that it should be the integral of a bounded increasing function over any interval interior to (a, b).

9. If $f(x)$ is absolutely continuous, so is $|f(x)|^p$, where $p \geqslant 1$.

10. A necessary and sufficient condition that $f(x)$ should be almost everywhere equal to a function of bounded variation in (a, b) is that as $h \to 0$

$$\int_a^b |f(x+h)-f(x)|\, dx = O(h)$$

[where $f(x) = 0$, say, outside (a, b)].*

[If $f(x)$ is of bounded variation, we have $f(x) = \phi(x)-\psi(x)$, where ϕ and ψ are positive, non-decreasing and bounded in (a, b). Then, if $h > 0$,

$$\int_a^b |f(x+h)-f(x)|\, dx \leqslant \int_a^b \{\phi(x+h)-\phi(x)\}\, dx + \int_a^b \{\psi(x+h)-\psi(x)\}\, dx$$

$$= \int_b^{b+h} \phi(t)\, dt - \int_a^{a+h} \phi(t)\, dt + \int_b^{b+h} \psi(t)\, dt - \int_a^{a+h} \psi(t)\, dt = O(h),$$

so that the condition is necessary.

Suppose now that the condition is satisfied. Let

$$\phi_n(x) = n \int_x^{x+1/n} f(t)\, dt.$$

Then

$$\int_a^b |\phi_n(x+h)-\phi_n(x)|\, dx = n \int_a^b dx \left| \int_{x+h}^{x+h+1/n} - \int_x^{x+1/n} f(t)\, dt \right|$$

$$= n \int_a^b dx \left| \int_0^{1/n} \{f(x+t+h)-f(x+t)\}\, dt \right|$$

$$\leqslant n \int_a^b dx \int_0^{1/n} |f(x+t+h)-f(x+t)|\, dt$$

$$= n \int_0^{1/n} dt \int_a^b |f(x+t+h)-f(x+t)|\, dx = O(h).$$

* Hardy and Littlewood (5), pp. 599–601, and (6), p. 619.

If $(x_\nu, x_\nu+h_\nu)$ is any set of non-overlapping intervals,

$$\sum |\phi_n(x_\nu+h_\nu)-\phi_n(x_\nu)| = \sum \left| \int_{x_\nu}^{x_\nu+h_\nu} \phi_n'(x)\, dx \right|$$

$$\leqslant \sum \int_{x_\nu}^{x_\nu+h_\nu} |\phi_n'(x)|\, dx \leqslant \int_a^b |\phi_n'(x)|\, dx,$$

and, by Fatou's lemma and the above result,

$$\int_a^b |\phi_n'(x)|\, dx \leqslant \underline{\lim} \int_a^b \left| \frac{\phi_n(x+h)-\phi_n(x)}{h} \right| dx = O(1).$$

Hence $$\sum |\phi_n(x_\nu+h_\nu)-\phi_n(x_\nu)| = O(1).$$

But $\phi_n(x) \to f(x)$ almost everywhere. Hence

$$\sum |f(x_\nu+h_\nu)-f(x_\nu)| < A$$

if none of the points x_ν, $x_\nu+h_\nu$ belong to a certain set E of measure zero. If a does not belong to E, it follows as in § 11.4 that $f(x) = f(a)+P(x)-N(x)$ in CE, where $P(x)$ and $N(x)$ are bounded and non-decreasing in CE. In E we can define $P(x)$ as $\lim P(x')$, where $x' \to x$ from below through CE. The result follows without difficulty from this.]

11. In § 11.4 the existence of $f'(x)$ at a point does not imply that of $V'(x)$.

[Consider $f(x) = x^2 \cos x^{-\alpha}$ $(0 < x \leqslant 1)$, $f(0) = 0$, $1 < \alpha < 2$.]

12. In § 11.54 the condition that $\phi(x)$ is continuous can be omitted.

[The proof shows that if α and β are any two points of continuity

$$\int_\alpha^\beta \phi'(x)\, dx \leqslant \phi(\beta)-\phi(\alpha).$$

But for any non-decreasing function points of continuity are everywhere dense. Hence, making $\alpha \to a+0$, $\beta \to b-0$, through such points, we obtain

$$\int_a^b \phi'(x)\, dx \leqslant \phi(b-0)-\phi(a+0).]$$

13. The set consisting of the intervals $\left(\frac{1}{2n+1}, \frac{1}{2n}\right)$, $n = 1, 2,\ldots$, has density $\frac{1}{4}$ at $x = 0$.

14. A convergent series of non-decreasing functions can be differentiated term by term almost everywhere.

Fubini: see Rajchman and Saks (1).

[Let $$u_1(x)+u_2(x)+\ldots+u_n(x) = s_n(x) \to s(x) \quad (a \leqslant x \leqslant b).$$

Then $s(x)$ is non-decreasing; and

$$\frac{s(x+h)-s(x)}{h} = \sum_{n=1}^{\infty} \frac{u_n(x+h)-u_n(x)}{h} \geqslant \sum_{n=1}^{N} \frac{u_n(x+h)-u_n(x)}{h}$$

for every N. Making $h \to 0$, it follows that

$$s'(x) \geqslant \sum_{n=1}^{N} u_n'(x)$$

almost everywhere. Hence $\sum u_n'(x)$ converges almost everywhere, to $\phi(x)$ say, and $\phi(x) \leqslant s'(x)$.

Suppose that the set $E(u,v)$, where $\phi(x) < u < v < s'(x)$, has positive measure μ. Almost everywhere in $E(u,v)$

$$s_n'(x) < u < v < s'(x),$$

so that $$s_n(x+h)-s_n(x) < hu < hv < s(x+h)-s(x)$$

for sufficiently small h. This holds over a finite non-overlapping set of intervals of total length $l > \frac{1}{2}\mu > 0$. Summing over these intervals

$$l(v-u) < \sum \{s(x+h)-s_n(x+h)\}-\{s(x)-s_n(x)\}$$
$$\leqslant \{s(b)-s_n(b)\}-\{s(a)-s_n(a)\}$$

since $s(x)-s_n(x)$ is non-decreasing. Making $n \to \infty$, $l \leqslant 0$, a contradiction. Hence $\phi(x) = s'(x)$ almost everywhere.]

15. If $f'(x)$ is finite everywhere, and equal to a continuous function almost everywhere, it is equal to it everywhere.

16. Show that $$\lim_{\delta \to 0} \int_{\delta}^{1} \frac{f(x+t)-f(x-t)}{t}\,dt$$

exists for every x if $f(x)$ is Weierstrass's non-differentiable function.

Show that the limit does not exist at $x = 0$ if $f(x)$ is the continuous function 0 ($x \leqslant 0$), $1/\log(1/x)$ ($x > 0$).

[This limit exists almost everywhere if $f(x)$ is any integrable function. See Titchmarsh, *Fourier Integrals*, Theorem 105.]

CHAPTER XII

FURTHER THEOREMS ON LEBESGUE INTEGRATION

12.1. In this chapter we adopt a slightly more practical point of view than in the two preceding ones. We have carried the general theory of definite and indefinite integrals as far as we shall require it, and we shall now prove a number of theorems which are useful in the manipulation of integrals.

12.11. Integration by parts. The formula of integration by parts in the Lebesgue theory is, of course, the same as the ordinary one: if $G(x)$ is an indefinite integral of $g(x)$, then

$$\int_a^b f(x)g(x)\,dx = [f(x)G(x)]_a^b - \int_a^b f'(x)G(x)\,dx.$$

The formula holds if $g(x)$ is any integrable function, and $f(x)$ is an integral.

The proof depends on the fact that *the product of two absolutely continuous functions is absolutely continuous.* For let $\phi(x)$ and $\psi(x)$ be absolutely continuous in (a, b), and let M and M' be the upper bounds of $|\phi(x)|$ and $|\psi(x)|$. Let $(x_\nu, x_\nu+h_\nu)$ be a set of non-overlapping intervals in (a, b). Then

$$\begin{aligned}\sum &|\phi(x_\nu+h_\nu)\psi(x_\nu+h_\nu)-\phi(x_\nu)\psi(x_\nu)|\\ &=\sum |\phi(x_\nu+h_\nu)\{\psi(x_\nu+h_\nu)-\psi(x_\nu)\}+\psi(x_\nu)\{\phi(x_\nu+h_\nu)-\phi(x_\nu)\}|\\ &\leqslant M\sum |\psi(x_\nu+h_\nu)-\psi(x_\nu)|+M'\sum |\phi(x_\nu+h_\nu)-\phi(x_\nu)|.\end{aligned}$$

The last two sums tend to zero with $\sum h_\nu$, and so $\phi(x)\psi(x)$ is absolutely continuous.

In the given formula, $f(x)$ and $G(x)$ are absolutely continuous, and hence so is $f(x)G(x)$; and

$$\int_a^b \frac{d}{dx}\{f(x)G(x)\}\,dx = [f(x)G(x)]_a^b.$$

But $$\frac{d}{dx}\{f(x)G(x)\} = f'(x)G(x)+f(x)g(x)$$

wherever $f'(x)$ and $G'(x)$ exist, and $G'(x) = g(x)$. Since this is true almost everywhere the result follows.

12.2. Approximation to an integrable function. The following theorem is often useful.

If $f(x)$ is measurable over a finite interval, then, given two positive numbers δ and ϵ, we can define an absolutely continuous function $\phi(x)$ such that $|f-\phi| < \delta$ except in a set of measure less than ϵ.

Suppose first that $f(x)$ is bounded. We may suppose without loss of generality that $f(x) \geqslant 0$. Divide up the interval of variation of $f(x)$ by the scale

$$0,\ \delta,\ 2\delta, \ldots,\ n\delta.$$

Let e_ν be the set where $\nu\delta \leqslant f(x) < (\nu+1)\delta$. Let $\psi_\nu(x) = \nu\delta$ in e_ν, and zero elsewhere. Then the function

$$\psi(x) = \psi_0(x) + \ldots + \psi_{n-1}(x)$$

differs from $f(x)$ by less than δ.

Let E_ν be an open set, including e_ν, of measure less than $m(e_\nu)+\epsilon/3n$. Let S_ν be the sum of a finite number of the intervals of E_ν, such that $m(E_\nu - S_\nu) < \epsilon/3n$. Let $\phi_\nu(x) = \nu\delta$ in S_ν, and zero elsewhere. Then $\phi_\nu = \psi_\nu$ except in a set of measure less than $2\epsilon/3n$; also ϕ_ν is discontinuous at a finite number of points, viz. the ends of the intervals of S_ν. To remove these discontinuities, we join the graph of the function to zero at the end of each interval by a straight line inclined so that the modifications all occur in a set of measure less than $\epsilon/3n$. Thus if ϕ'_ν is the modified function, ϕ'_ν is absolutely continuous, and $\phi'_\nu = \psi_\nu$ except in a set of measure ϵ/n.

Let $$\phi(x) = \phi'_0 + \phi'_1 + \cdots + \phi'_{n-1}.$$

Then $\phi(x)$ is absolutely continuous, and $\phi(x) = \psi(x)$ except in a set of measure ϵ. Hence $\phi(x)$ has the required property.

If $f(x)$ is not bounded, let $\{f(x)\}_k = f(x)$ where $|f(x)| \leqslant k$, and $\{f(x)\}_k = 0$ elsewhere. We can take k so large that $\{f(x)\}_k = f(x)$ except in a set of measure $\frac{1}{2}\epsilon$. By the first part, we can determine $\phi(x)$ so that $|\{f(x)\}_k - \phi(x)| < \delta$ except in a set of measure $\frac{1}{2}\epsilon$. Then $\phi(x)$ has the required property.

Notice that, if $f(x)$ is bounded, $\phi(x)$ can be constructed to lie between the same bounds as $f(x)$.

If $f(x)$ is integrable, we can construct $\phi(x)$ so that, in addition to the above properties,

$$\int_a^b |f(x) - \phi(x)|\, dx < \eta, \tag{1}$$

where η is arbitrarily small. If $f(x)$ is bounded, say $|f(x)| \leqslant M$, then $|\phi(x)| \leqslant M$, and

$$\int_a^b |f(x)-\phi(x)|\, dx \leqslant \delta(b-a)+2\epsilon M,$$

giving the required result. If $f(x)$ is unbounded, we define $\{f(x)\}_k$ as above, and then determine $\phi(x)$ so that $|\{f(x)\}_k-\phi(x)| < \delta$ except in a set of measure $\frac{1}{2}\epsilon/k$. Then

$$\int_a^b |f(x)-\phi(x)|\, dx \leqslant \int_a^b |f(x)-f\{(x)\}_k|\, dx + \int_a^b |\{f(x)\}_k-\phi(x)|\, dx.$$

The first term tends to zero as $k \to \infty$, and the second term does not exceed $\delta(b-a)+\epsilon$. Hence the result.

Example. If $f(x)$ is integrable over $(a-\epsilon, b+\epsilon)$, then

$$\lim_{h\to 0} \int_a^b |f(x+h)-f(x)|\, dx = 0.$$

12.21. Change of the independent variable. Here again the formula is familiar, but the conditions under which it holds are novel.

If $f(x)$ and $g(x)$ are integrable, $g(x) \geqslant 0$, and $G(x)$ is an indefinite integral of $g(x)$, $a = G(\alpha)$, $b = G(\beta)$, then

$$\int_a^b f(t)\, dt = \int_\alpha^\beta f\{G(x)\}g(x)\, dx,$$

where $f\{G(x)\}g(x)$ is defined as 0 if $g(x) = 0$.

The inverse function of $t = G(x)$, of which α and β are values, is not necessarily one-valued, since $G(x)$ may be constant in some intervals. But if more than one value of x corresponds to a given value of t, these values of x form a closed interval, and we can make the inverse function one-valued by taking x to be, say, the left-hand end-point of the interval.

We next observe that *if $F(x)$ and $G(x)$ are absolutely continuous functions, and $G(x)$ is monotonic, then $F\{G(x)\}$ is absolutely continuous.* For, since F is absolutely continuous,

$$\sum |F\{G(x_\nu+h_\nu)\}-F\{G(x_\nu)\}|$$

tends to zero with

$$\sum |G(x_\nu+h_\nu)-G(x_\nu)|,$$

and, since $G(x)$ is absolutely continuous, this tends to zero with $\sum h_\nu$.

It follows that, if $F(x)$ and $G(x)$ are integrals of $f(x)$ and $g(x)$, then $F\{G(x)\}$ has a finite differential coefficient for almost all values of x, and

$$\int\limits_{\alpha}^{\beta} \frac{d}{dx}[F\{G(x)\}]\, dx = F\{G(\beta)\} - F\{G(\alpha)\} = \int\limits_{a}^{b} f(t)\, dt.$$

The result will now follow if

$$\frac{d}{dx}[F\{G(x)\}] = f\{G(x)\}g(x) \tag{1}$$

for almost all values of x. But this is not obviously true. For

$$\frac{F\{G(x+h)\}-F\{G(x)\}}{h} = \frac{F\{G(x+h)\}-F\{G(x)\}}{G(x+h)-G(x)} \cdot \frac{G(x+h)-G(x)}{h},$$

and the second factor on the right tends to $g(x)$ for almost all values of x, while the first factor tends to $f\{G(x)\}$ for almost all values of $G(x)$; and the difficulty is that the exceptional set of values of $G(x)$, of measure zero, does not necessarily correspond to a set of values of x of measure zero.

Let $f(x)$ be bounded, say $|f(x)| \leqslant M$. Divide the interval (α, β) into sets $E_1, \ldots, E_4$ as follows. In E_1, $G'(x) = g(x) > 0$, and the first factor on the right tends to $f\{G(x)\}$; in E_2,

$$G'(x) = g(x) > 0$$

but the other condition is negatived; in E_3, $G'(x) = g(x) = 0$; in E_4, $G'(x) \neq g(x)$. Clearly (1) holds in E_1; and it holds in E_3, since there

$$\left|\frac{F\{G(x+h)\}-F\{G(x)\}}{h}\right|$$

$$= \left|\frac{1}{h}\int\limits_{G(x)}^{G(x+h)} f(t)\, dt\right| \leqslant M\left|\frac{G(x+h)-G(x)}{h}\right| \to 0$$

and each side of (1) is zero; $m(E_4) = 0$; and we have to prove that $m(E_2) = 0$.

Let $E_{2,n}$ be the part of E_2 in which $G'(x) > 1/n$. Enclose the corresponding t-set in an open set O of measure less than a given ϵ. With each x of $E_{2,n}$ associate an interval $(x, x+h_x)$ such that $G(x+h_x)-G(x) > h_x/n$, and such that the interval $G(x)$, $G(x+h_x)$ is in O. By Lemma 1 of § 11.41, there is a finite non-overlapping set S of the intervals $(x, x+h_x)$ such that

$$m_e(E_{2,n}) < m(S)+\epsilon = \sum_S h_x+\epsilon.$$

This is less than

$$n \sum_S \{G(x+h_x)-G(x)\}+\epsilon \leqslant nm(O)+\epsilon < (n+1)\epsilon.$$

Hence $m(E_{2,n}) = 0$, and, since E_2 is the outer limiting set of the sets $E_{2,n}$, $m(E_2) = 0$.

Lastly let $f(x)$ be any integrable function. We may suppose without loss of generality that it is positive. Defining $\{f(x)\}_n$ in the usual way, the theorem holds for $\{f(x)\}_n$, and it is sufficient to prove that

$$\lim \int_\alpha^\beta [f\{G(x)\}]_n g(x)\,dx = \int_\alpha^\beta f\{G(x)\}g(x)\,dx.$$

But $$\int_\alpha^\beta [f\{G(x)\}]_n g(x)\,dx = \int_a^b \{f(t)\}_n\,dt \leqslant \int_a^b f(t)\,dt.$$

The result therefore follows from the convergence theorem of § 10.82 (regarding $f\{G(x)\}g(x)$ as 0 if $f\{G(x)\} = \infty$, $g(x) = 0$).

12.3. The second mean-value theorem. *If $f(x)$ is integrable over (a, b), and $\phi(x)$ is positive, bounded, and non-increasing, then*

$$\int_a^b f(x)\phi(x)\,dx = \phi(a+0)\int_a^\xi f(x)\,dx,$$

where ξ is some number between a and b.

Let ϵ be a positive number less than $\phi(a+0)-\phi(b-0)$. Then there is a point x_1 such that

$$\begin{aligned}\phi(a+0)-\phi(x) &< \epsilon \qquad (a<x<x_1)\\ &\geqslant \epsilon \qquad (x>x_1).\end{aligned}$$

Similarly there are points $x_2, x_3,\ldots$ such that

$$\begin{aligned} \phi(x_{\nu-1}+0)-\phi(x) &< \epsilon \qquad (x_{\nu-1} < x < x_\nu) \\ &\geqslant \epsilon \qquad (x > x_\nu), \end{aligned}$$

so long as $\phi(x_{\nu-1}+0)-\phi(b-0) > \epsilon$. Otherwise we take $x_n = b$. The point b is thus reached in a finite number of steps, since the variation of $\phi(x)$ in each interval $(x_{\nu-1}, x_\nu)$ is at least ϵ.

Let $\psi(x) = \phi(x_\nu+0)$ in each interval $x_\nu \leqslant x < x_{\nu+1}$. Then $0 \leqslant \psi(x)-\phi(x) < \epsilon$ except possibly at the points $a = x_0$, x_1, $x_2,\ldots, b$, and

$$\int_a^b \psi(x)f(x)\,dx = \sum_{\nu=0}^{n-1} \phi(x_\nu+0) \int_{x_\nu}^{x_{\nu+1}} f(x)\,dx.$$

Let $F(x) = \int_a^x f(t)\,dt$; then, if m and M are the lower and upper bounds of $F(x)$, it follows from Abel's lemma (§ 1.131) that

$$m\phi(a+0) \leqslant \int_a^b \psi(x)f(x)\,dx \leqslant M\phi(a+0).$$

But

$$\left|\int_a^b \psi(x)f(x)\,dx - \int_a^b \phi(x)f(x)\,dx\right| \leqslant \epsilon \int_a^b |f(x)|\,dx,$$

which tends to zero with ϵ. Hence, making $\epsilon \to 0$, it follows that

$$m\phi(a+0) \leqslant \int_a^b \phi(x)f(x)\,dx \leqslant M\phi(a+0).$$

Since $F(x)$ is continuous, it takes every value between m and M, and so, at $x = \xi$ say, the value

$$\frac{1}{\phi(a+0)} \int_a^b \phi(x)f(x)\,dx.$$

This proves the theorem.

If $\phi(x)$ is positive and non-decreasing, the corresponding formula is

$$\int_a^b f(x)\phi(x)\,dx = \phi(b-0) \int_\xi^b f(x)\,dx,$$

where $a < \xi < b$.

If $\phi(x)$ is any monotonic function, there is a number ξ between a and b such that

$$\int_a^b f(x)\phi(x)\,dx = \phi(a+0)\int_a^\xi f(x)\,dx + \phi(b-0)\int_\xi^b f(x)\,dx.$$

This is obtained from the previous results by considering $\phi(x)-\phi(a+0)$ or $\phi(x)-\phi(b-0)$.

12.4. The Lebesgue class* L^p. We denote by $L^p(a,b)$ the class of functions $f(x)$ such that $f(x)$ is measurable, and $|f(x)|^p$, where $p>0$, is integrable over (a,b). If it is not necessary to specify the interval, we denote the class by L^p simply. The class L^1 is the class of functions integrable over (a,b), and is denoted simply by L.

We may classify functions defined over any set, or over an infinite interval, in the same way; for example, the function $(1+x)^{-\frac{1}{2}}$ belongs to $L^p(0,\infty)$ if $p>2$.

If $f(x)$ belongs to L^p, and $|g(x)| \leqslant |f(x)|$, then clearly $g(x)$ also belongs to L^p.

Examples. (i) A bounded function belongs to $L^p(a, b)$, where (a, b) is a finite interval, for all values of p.

(ii) If $f(x)$ belongs to $L^p(a, b)$, where (a, b) is a finite interval, then it also belongs to $L^q(a, b)$ for $q<p$.

(iii) If $f(x)$ belongs to $L^p(0, \infty)$ and to $L^q(0, \infty)$, where $p<q$, then it also belongs to $L^r(0, \infty)$ if $p<r<q$.

[Consider separately the sets where $|f(x)| \leqslant 1$ and $|f(x)| > 1$.]

(iv) The sum of two functions of L^p also belongs to L^p.

[For $|f(x)+g(x)|^p \leqslant \max\{2^p|f(x)|^p, 2^p|g(x)|^p\}$.]

(v) The function $\{x\log^2 1/x\}^{-1}$ belongs to $L(0, \frac{1}{2})$, but not to any $L^p(0, \frac{1}{2})$ for $p>1$.

(vi) The function $\{x^{\frac{1}{2}}(1+|\log x|)\}^{-1}$ belongs to $L^2(0, \infty)$, but not to $L^p(0, \infty)$ for any other value of p.

12.41. Schwarz's inequality. *If $f(x)$ and $g(x)$ belong to L^2, then $f(x)g(x)$ belongs to L, and*

$$\left|\int f(x)g(x)\,dx\right| \leqslant \left\{\int |f(x)|^2\,dx \int |g(x)|^2\,dx\right\}^{\frac{1}{2}}.$$

The interval of integration may be finite or infinite.

* See in particular F. Riesz (2).

Since $2|fg| \leqslant f^2+g^2$, fg belongs to L. Hence the integral

$$\int \{\lambda f(x)+\mu g(x)\}^2\,dx$$
$$=\lambda^2 \int \{f(x)\}^2\,dx+2\lambda\mu \int f(x)g(x)\,dx+\mu^2 \int \{g(x)\}^2\,dx$$

exists for all values of λ and μ. It is evidently never negative. But the necessary and sufficient condition that $a\lambda^2+2h\lambda\mu+b\mu^2$ should be never negative is that $h^2 \leqslant ab$, $a \geqslant 0$, $b \geqslant 0$; and this gives the inequality stated.

Examples. (i) The case of equality in the above theorem occurs only if $f(x)/g(x)$ is almost everywhere equal to a constant.

(ii) If $f(x)$ and $g(x)$ belong to L^p, where $p > 2$, then $f(x)g(x)$ belongs to $L^{\frac{1}{2}p}$.

12.42. Hölder's inequality. This is a generalization of Schwarz's inequality.

If $f(x)$ belongs to L^p, and $g(x)$ to $L^{p/(p-1)}$, where $p > 1$, then $f(x)g(x)$ belongs to L, and

$$\left|\int f(x)g(x)\,dx\right| \leqslant \left\{\int |f(x)|^p\,dx\right\}^{1/p}\left\{\int |g(x)|^{p/(p-1)}\,dx\right\}^{1-1/p}. \quad (1)$$

The interval of integration may be finite or infinite.

Let E be the set where $|g(x)| \leqslant |f(x)|^{p-1}$. Then

$$|f(x)g(x)| \leqslant |f(x)|^p$$

in E; hence $f(x)g(x)$ is integrable over E. In the complementary set CE, $|f(x)| < |g(x)|^{1/(p-1)}$. Hence

$$|f(x)g(x)| < |g(x)|^{p/(p-1)}$$

in CE; hence $f(x)g(x)$ is integrable over CE, and so over the whole interval considered.

This argument can be used to obtain an inequality similar to (1), but with a factor 2 on the right-hand side. Let

$$I = \int_a^b |f(x)|^p\,dx, \qquad J = \int_a^b |g(x)|^{p/(p-1)}\,dx.$$

Then

$$\left|\int_a^b fg\,dx\right| \leqslant \int_E |fg|\,dx + \int_{CE} |fg|\,dx$$
$$\leqslant \int_E |f|^p\,dx + \int_{CE} |g|^{p/(p-1)}\,dx \leqslant I+J. \quad (2)$$

If we replace $f(x)$ and $g(x)$ in this inequality by

$$(J/I)^{(p-1)/p^2}f(x), \qquad (I/J)^{(p-1)/p^2}g(x),$$

respectively, the left-hand side is unchanged, and each term on the right-hand side is replaced by $I^{1/p}J^{1-1/p}$. Hence

$$\left|\int fg\,dx\right| \leqslant 2I^{1/p}J^{1-1/p}. \tag{3}$$

The inequality (1) can be deduced from the well-known inequality

$$x^m - 1 < m(x-1) \qquad (x>1,\ 0<m<1). \tag{4}$$

Putting $x = a/b$ $(a>b)$, and multiplying by b,

$$a^m b^{1-m} < b + m(a-b).$$

Putting $m=\alpha$, $1-m=\beta$, so that $\alpha+\beta=1$, this takes the form

$$a^\alpha b^\beta < a\alpha + b\beta, \tag{5}$$

and since this is symmetrical it holds if a and b are any unequal positive numbers. If $a=b$ it becomes an equality.

Using (5), we have, if $F(x)\geqslant 0$, $G(x)\geqslant 0$,

$$\int_a^b \left(\frac{F(x)}{\int_a^b F(t)\,dt}\right)^\alpha \left(\frac{G(x)}{\int_a^b G(t)\,dt}\right)^\beta dx \leqslant \int_a^b \left(\frac{\alpha F(x)}{\int_a^b F(t)\,dt} + \frac{\beta G(x)}{\int_a^b G(t)\,dt}\right) dx$$

$$= \alpha+\beta = 1,$$

i.e.

$$\int_a^b \{F(x)\}^\alpha \{G(x)\}^\beta\,dx \leqslant \left\{\int_a^b F(x)\,dx\right\}^\alpha \left\{\int_a^b G(x)\,dx\right\}^\beta.$$

Finally, putting $\alpha = 1/p$, $F(x) = |f(x)|^p$, and $G(x) = |g(x)|^{p/(p-1)}$, the result (1) follows.*

Example. The case of equality occurs only if $|f(x)|^p/|g(x)|^{p/(p-1)}$ is almost everywhere equal to a constant.

12.421. Hölder's inequality for sums. This is

$$|\sum a_n b_n| \leqslant (\sum |a_n|^p)^{1/p} (\sum |b_n|^{p/(p-1)})^{1-1/p}.$$

The proof is similar to that of the integral inequality. We have

$$\sum \left\{\left(\frac{A_n}{\sum A_n}\right)^\alpha \left(\frac{B_n}{\sum B_n}\right)^\beta\right\} \leqslant \sum \left(\alpha\frac{A_n}{\sum A_n} + \beta\frac{B_n}{\sum B_n}\right)$$

$$= \alpha+\beta = 1,$$

i.e.

$$\sum A_n^\alpha B_n^\beta \leqslant (\sum A_n)^\alpha (\sum B_n)^\beta,$$

and writing $\alpha = 1/p$, $A_n = |a_n|^p$, $B_n = |b_n|^{p/(p-1)}$, the result follows.

* This proof is given by Hardy (20).

12.43. Minkowski's inequality. *If $f(x)$ and $g(x)$ belong to L^p, where $p > 1$, then*

$$\left\{\int |f(x)+g(x)|^p\,dx\right\}^{1/p} \leqslant \left\{\int |f(x)|^p\,dx\right\}^{1/p} + \left\{\int |g(x)|^p\,dx\right\}^{1/p}. \quad (1)$$

For

$$\int |f+g|^p\,dx \leqslant \int |f|\,.\,|f+g|^{p-1}\,dx + \int |g|\,.\,|f+g|^{p-1}\,dx$$
$$\leqslant \left\{\int |f|^p\,dx\right\}^{1/p}\left\{\int |f+g|^p\,dx\right\}^{1-1/p} +$$
$$+\left\{\int |g|^p\,dx\right\}^{1/p}\left\{\int |f+g|^p\,dx\right\}^{1-1/p}$$

by Hölder's inequality. Dividing each side by

$$\left\{\int |f+g|^p\,dx\right\}^{1-1/p},$$

the result follows.

The corresponding inequality for sums

$$(\sum |a_n+b_n|^p)^{1/p} \leqslant (\sum |a_n|^p)^{1/p} + (\sum |b_n|^p)^{1/p} \quad (2)$$

can be proved in a similar way.

12.44. The integral of a function of L^p. We have seen in the previous chapter that a necessary and sufficient condition that a function should be an integral is that it should be absolutely continuous. There is a corresponding condition that a function should be an integral of a function of the class L^p.

A necessary and sufficient condition that a function $F(x)$ should be the integral of a function of the class L^p, where $p > 1$, is that the sum

$$\sum |F(x_\nu+h_\nu)-F(x_\nu)|^p h_\nu^{1-p},$$

taken over any system of non-overlapping intervals $(x_\nu, x_\nu+h_\nu)$, should be bounded.

If instead of 'should be bounded' we say 'should be bounded and tend to zero with $\sum h_\nu$', the theorem is still true, and in this form it is true for $p = 1$ also, and so includes the theorem on absolute continuity as a particular case. For $p > 1$ the two conditions, one of which appears to be more restrictive than the other, turn out to be equivalent.

To prove that the condition is necessary, suppose that

$$F(x) = F(a) + \int_a^x f(t)\,dt,$$