6

THE RIEMANN-STIELTJES INTEGRAL

The present chapter is based on a definition of the Riemann integral which depends very explicitly on the order structure of the real line. Accordingly, we begin by discussing integration of real-valued functions on intervals. Extensions to complex- and vector-valued functions on intervals follow in later sections. Integration over sets other than intervals is discussed in Chaps. 10 and 11.

DEFINITION AND EXISTENCE OF THE INTEGRAL

6.1 Definition Let $[a, b]$ be a given interval. By a *partition P* of $[a, b]$ we mean a finite set of points $x_0, x_1, \ldots, x_n$, where

$$a = x_0 \le x_1 \le \cdots \le x_{n-1} \le x_n = b.$$

We write

$$\Delta x_i = x_i - x_{i-1} \qquad (i = 1, \ldots, n).$$

Now suppose f is a bounded real function defined on $[a, b]$. Corresponding to each partition P of $[a, b]$ we put

$$M_i = \sup f(x) \qquad (x_{i-1} \le x \le x_i),$$
$$m_i = \inf f(x) \qquad (x_{i-1} \le x \le x_i),$$
$$U(P,f) = \sum_{i=1}^{n} M_i \, \Delta x_i,$$
$$L(P,f) = \sum_{i=1}^{n} m_i \, \Delta x_i,$$

and finally

(1)
$$\overline{\int_a^b} f\,dx = \inf U(P,f),$$

(2)
$$\underline{\int_a^b} f\,dx = \sup L(P,f),$$

where the inf and the sup are taken over all partitions P of $[a, b]$. The left members of (1) and (2) are called the *upper* and *lower Riemann integrals* of f over $[a, b]$, respectively.

If the upper and lower integrals are equal, we say that f is *Riemann-integrable* on $[a, b]$, we write $f \in \mathscr{R}$ (that is, $\mathscr{R}$ denotes the set of Riemann-integrable functions), and we denote the common value of (1) and (2) by

(3)
$$\int_a^b f\,dx,$$

or by

(4)
$$\int_a^b f(x)\,dx.$$

This is the *Riemann integral* of f over $[a, b]$. Since f is bounded, there exist two numbers, m and M, such that

$$m \le f(x) \le M \qquad (a \le x \le b).$$

Hence, for every P,

$$m(b-a) \le L(P,f) \le U(P,f) \le M(b-a),$$

so that the numbers $L(P,f)$ and $U(P,f)$ form a bounded set. This shows that *the upper and lower integrals are defined* for *every* bounded function f. The question of their equality, and hence the question of the integrability of f, is a more delicate one. Instead of investigating it separately for the Riemann integral, we shall immediately consider a more general situation.

6.2 Definition Let α be a monotonically increasing function on $[a, b]$ (since $\alpha(a)$ and $\alpha(b)$ are finite, it follows that α is bounded on $[a, b]$). Corresponding to each partition P of $[a, b]$, we write

$$\Delta\alpha_i = \alpha(x_i) - \alpha(x_{i-1}).$$

It is clear that $\Delta\alpha_i \geq 0$. For any real function f which is bounded on $[a, b]$ we put

$$U(P, f, \alpha) = \sum_{i=1}^{n} M_i \, \Delta\alpha_i,$$

$$L(P, f, \alpha) = \sum_{i=1}^{n} m_i \, \Delta\alpha_i,$$

where M_i, m_i have the same meaning as in Definition 6.1, and we define

$$\overline{\int_a^b} f \, d\alpha = \inf U(P, f, \alpha), \tag{5}$$

$$\underline{\int_a^b} f \, d\alpha = \sup L(P, f, \alpha), \tag{6}$$

the inf and sup again being taken over all partitions.

If the left members of (5) and (6) are equal, we denote their common value by

$$\int_a^b f \, d\alpha \tag{7}$$

or sometimes by

$$\int_a^b f(x) \, d\alpha(x). \tag{8}$$

This is the *Riemann-Stieltjes integral* (or simply the *Stieltjes integral*) of f with respect to α, over $[a, b]$.

If (7) exists, i.e., if (5) and (6) are equal, we say that f is integrable with respect to α, in the Riemann sense, and write $f \in \mathscr{R}(\alpha)$.

By taking $\alpha(x) = x$, the Riemann integral is seen to be a special case of the Riemann-Stieltjes integral. Let us mention explicitly, however, that in the general case α need not even be continuous.

A few words should be said about the notation. We prefer (7) to (8), since the letter x which appears in (8) adds nothing to the content of (7). It is immaterial which letter we use to represent the so-called "variable of integration." For instance, (8) is the same as

$$\int_a^b f(y) \, d\alpha(y).$$

The integral depends on f, α, a and b, but not on the variable of integration, which may as well be omitted.

The role played by the variable of integration is quite analogous to that of the index of summation: The two symbols

$$\sum_{i=1}^{n} c_i, \qquad \sum_{k=1}^{n} c_k$$

mean the same thing, since each means $c_1 + c_2 + \cdots + c_n$.

Of course, no harm is done by inserting the variable of integration, and in many cases it is actually convenient to do so.

We shall now investigate the existence of the integral (7). Without saying so every time, f will be assumed real and bounded, and α monotonically increasing on $[a, b]$; and, when there can be no misunderstanding, we shall write $\int$ in place of $\int_a^b$.

6.3 Definition We say that the partition P^* is a *refinement* of P if $P^* \supset P$ (that is, if every point of P is a point of P^*). Given two partitions, P_1 and P_2, we say that P^* is their *common refinement* if $P^* = P_1 \cup P_2$.

6.4 Theorem *If P^* is a refinement of P, then*

$$L(P, f, \alpha) \le L(P^*, f, \alpha) \tag{9}$$

and

$$U(P^*, f, \alpha) \le U(P, f, \alpha). \tag{10}$$

Proof To prove (9), suppose first that P^* contains just one point more than P. Let this extra point be x^*, and suppose $x_{i-1} < x^* < x_i$, where x_{i-1} and x_i are two consecutive points of P. Put

$$w_1 = \inf f(x) \qquad (x_{i-1} \le x \le x^*),$$

$$w_2 = \inf f(x) \qquad (x^* \le x \le x_i).$$

Clearly $w_1 \ge m_i$ and $w_2 \ge m_i$, where, as before,

$$m_i = \inf f(x) \qquad (x_{i-1} \le x \le x_i).$$

Hence

$$\begin{aligned} &L(P^*, f, \alpha) - L(P, f, \alpha) \\ &\quad = w_1[\alpha(x^*) - \alpha(x_{i-1})] + w_2[\alpha(x_i) - \alpha(x^*)] - m_i[\alpha(x_i) - \alpha(x_{i-1})] \\ &\quad = (w_1 - m_i)[\alpha(x^*) - \alpha(x_{i-1})] + (w_2 - m_i)[\alpha(x_i) - \alpha(x^*)] \ge 0. \end{aligned}$$

If P^* contains k points more than P, we repeat this reasoning k times, and arrive at (9). The proof of (10) is analogous.

6.5 Theorem $\underline{\int_a^b} f\,d\alpha \le \overline{\int_a^b} f\,d\alpha.$

Proof Let P^* be the common refinement of two partitions P_1 and P_2. By Theorem 6.4,

$$L(P_1, f, \alpha) \le L(P^*, f, \alpha) \le U(P^*, f, \alpha) \le U(P_2, f, \alpha).$$

Hence

$$L(P_1, f, \alpha) \le U(P_2, f, \alpha). \tag{11}$$

If P_2 is fixed and the sup is taken over all P_1, (11) gives

$$\underline{\int} f\,d\alpha \le U(P_2, f, \alpha). \tag{12}$$

The theorem follows by taking the inf over all P_2 in (12).

6.6 Theorem *$f \in \mathscr{R}(\alpha)$ on $[a, b]$ if and only if for every $\varepsilon > 0$ there exists a partition P such that*

$$U(P, f, \alpha) - L(P, f, \alpha) < \varepsilon. \tag{13}$$

Proof For every P we have

$$L(P, f, \alpha) \le \underline{\int} f\,d\alpha \le \overline{\int} f\,d\alpha \le U(P, f, \alpha).$$

Thus (13) implies

$$0 \le \overline{\int} f\,d\alpha - \underline{\int} f\,d\alpha < \varepsilon.$$

Hence, if (13) can be satisfied for every $\varepsilon > 0$, we have

$$\overline{\int} f\,d\alpha = \underline{\int} f\,d\alpha,$$

that is, $f \in \mathscr{R}(\alpha)$.

Conversely, suppose $f \in \mathscr{R}(\alpha)$, and let $\varepsilon > 0$ be given. Then there exist partitions P_1 and P_2 such that

$$U(P_2, f, \alpha) - \int f\,d\alpha < \frac{\varepsilon}{2}, \tag{14}$$

$$\int f\,d\alpha - L(P_1, f, \alpha) < \frac{\varepsilon}{2}. \tag{15}$$

We choose P to be the common refinement of P_1 and P_2. Then Theorem 6.4, together with (14) and (15), shows that

$$U(P, f, \alpha) \le U(P_2, f, \alpha) < \int f\, d\alpha + \frac{\varepsilon}{2} < L(P_1, f, \alpha) + \varepsilon \le L(P, f, \alpha) + \varepsilon,$$

so that (13) holds for this partition P.

Theorem 6.6 furnishes a convenient criterion for integrability. Before we apply it, we state some closely related facts.

6.7 Theorem

(*a*) *If* (13) *holds for some* P *and some* ε, *then* (13) *holds* (*with the same* ε) *for every refinement of* P.

(*b*) *If* (13) *holds for* $P = \{x_0, \ldots, x_n\}$ *and if* s_i, t_i *are arbitrary points in* $[x_{i-1}, x_i]$, *then*

$$\sum_{i=1}^{n} |f(s_i) - f(t_i)| \,\Delta\alpha_i < \varepsilon.$$

(*c*) *If* $f \in \mathscr{R}(\alpha)$ *and the hypotheses of* (*b*) *hold, then*

$$\left| \sum_{i=1}^{n} f(t_i)\,\Delta\alpha_i - \int_a^b f\, d\alpha \right| < \varepsilon.$$

Proof Theorem 6.4 implies (*a*). Under the assumptions made in (*b*), both $f(s_i)$ and $f(t_i)$ lie in $[m_i, M_i]$, so that $|f(s_i) - f(t_i)| \le M_i - m_i$. Thus

$$\sum_{i=1}^{n} |f(s_i) - f(t_i)| \,\Delta\alpha_i \le U(P, f, \alpha) - L(P, f, \alpha),$$

which proves (*b*). The obvious inequalities

$$L(P, f, \alpha) \le \sum f(t_i)\,\Delta\alpha_i \le U(P, f, \alpha)$$

and

$$L(P, f, \alpha) \le \int f\, d\alpha \le U(P, f, \alpha)$$

prove (*c*).

6.8 Theorem *If f is continuous on* $[a, b]$ *then* $f \in \mathscr{R}(\alpha)$ *on* $[a, b]$.

Proof Let $\varepsilon > 0$ be given. Choose $\eta > 0$ so that

$$[\alpha(b) - \alpha(a)]\eta < \varepsilon.$$

Since f is uniformly continuous on $[a, b]$ (Theorem 4.19), there exists a $\delta > 0$ such that

(16) $$|f(x) - f(t)| < \eta$$

if $x \in [a, b]$, $t \in [a, b]$, and $|x - t| < \delta$.

If P is any partition of $[a, b]$ such that $\Delta x_i < \delta$ for all i, then (16) implies that

(17) $$M_i - m_i \le \eta \qquad (i - 1, \dots, n)$$

and therefore

$$U(P, f, \alpha) - L(P, f, \alpha) = \sum_{i=1}^{n} (M_i - m_i)\, \Delta\alpha_i$$

$$\le \eta \sum_{i=1}^{n} \Delta\alpha_i = \eta[\alpha(b) - \alpha(a)] < \varepsilon.$$

By Theorem 6.6, $f \in \mathscr{R}(\alpha)$.

6.9 Theorem *If f is monotonic on $[a, b]$, and if α is continuous on $[a, b]$, then $f \in \mathscr{R}(\alpha)$. (We still assume, of course, that α is monotonic.)*

Proof Let $\varepsilon > 0$ be given. For any positive integer n, choose a partition such that

$$\Delta\alpha_i = \frac{\alpha(b) - \alpha(a)}{n} \qquad (i = 1, \dots, n).$$

This is possible since α is continuous (Theorem 4.23).

We suppose that f is monotonically increasing (the proof is analogous in the other case). Then

$$M_i = f(x_i), \qquad m_i = f(x_{i-1}) \qquad (i = 1, \dots, n),$$

so that

$$U(P, f, \alpha) - L(P, f, \alpha) = \frac{\alpha(b) - \alpha(a)}{n} \sum_{i=1}^{n} [f(x_i) - f(x_{i-1})]$$

$$= \frac{\alpha(b) - \alpha(a)}{n} \cdot [f(b) - f(a)] < \varepsilon$$

if n is taken large enough. By Theorem 6.6, $f \in \mathscr{R}(\alpha)$.

6.10 Theorem *Suppose f is bounded on $[a, b]$, f has only finitely many points of discontinuity on $[a, b]$, and α is continuous at every point at which f is discontinuous. Then $f \in \mathscr{R}(\alpha)$.*

Proof Let $\varepsilon > 0$ be given. Put $M = \sup |f(x)|$, let E be the set of points at which f is discontinuous. Since E is finite and α is continuous at every point of E, we can cover E by finitely many disjoint intervals $[u_j, v_j] \subset [a, b]$ such that the sum of the corresponding differences $\alpha(v_j) - \alpha(u_j)$ is less than ε. Furthermore, we can place these intervals in such a way that every point of $E \cap (a, b)$ lies in the interior of some $[u_j, v_j]$.

Remove the segments (u_j, v_j) from $[a, b]$. The remaining set K is compact. Hence f is uniformly continuous on K, and there exists $\delta > 0$ such that $|f(s) - f(t)| < \varepsilon$ if $s \in K$, $t \in K$, $|s - t| < \delta$.

Now form a partition $P = \{x_0, x_1, \ldots, x_n\}$ of $[a, b]$, as follows: Each u_j occurs in P. Each v_j occurs in P. No point of any segment (u_j, v_j) occurs in P. If x_{i-1} is not one of the u_j, then $\Delta x_i < \delta$.

Note that $M_i - m_i \leq 2M$ for every i, and that $M_i - m_i \leq \varepsilon$ unless x_{i-1} is one of the u_j. Hence, as in the proof of Theorem 6.8,

$$U(P, f, \alpha) - L(P, f, \alpha) \leq [\alpha(b) - \alpha(a)]\varepsilon + 2M\varepsilon.$$

Since ε is arbitrary, Theorem 6.6 shows that $f \in \mathscr{R}(\alpha)$.

Note: If f and α have a common point of discontinuity, then f need not be in $\mathscr{R}(\alpha)$. Exercise 3 shows this.

6.11 Theorem *Suppose $f \in \mathscr{R}(\alpha)$ on $[a, b]$, $m \leq f \leq M$, ϕ is continuous on $[m, M]$, and $h(x) = \phi(f(x))$ on $[a, b]$. Then $h \in \mathscr{R}(\alpha)$ on $[a, b]$.*

Proof Choose $\varepsilon > 0$. Since ϕ is uniformly continuous on $[m, M]$, there exists $\delta > 0$ such that $\delta < \varepsilon$ and $|\phi(s) - \phi(t)| < \varepsilon$ if $|s - t| \leq \delta$ and $s, t \in [m, M]$.

Since $f \in \mathscr{R}(\alpha)$, there is a partition $P = \{x_0, x_1, \ldots, x_n\}$ of $[a, b]$ such that

$$U(P, f, \alpha) - L(P, f, \alpha) < \delta^2. \tag{18}$$

Let M_i, m_i have the same meaning as in Definition 6.1, and let M_i^*, m_i^* be the analogous numbers for h. Divide the numbers $1, \ldots, n$ into two classes: $i \in A$ if $M_i - m_i < \delta$, $i \in B$ if $M_i - m_i \geq \delta$.

For $i \in A$, our choice of δ shows that $M_i^* - m_i^* \leq \varepsilon$.

For $i \in B$, $M_i^* - m_i^* \leq 2K$, where $K = \sup|\phi(t)|$, $m \leq t \leq M$. By (18), we have

$$\delta \sum_{i \in B} \Delta\alpha_i \leq \sum_{i \in B} (M_i - m_i)\, \Delta\alpha_i < \delta^2 \tag{19}$$

so that $\sum_{i \in B} \Delta\alpha_i < \delta$. It follows that

$$\begin{aligned} U(P, h, \alpha) - L(P, h, \alpha) &= \sum_{i \in A} (M_i^* - m_i^*)\, \Delta\alpha_i + \sum_{i \in B} (M_i^* - m_i^*)\, \Delta\alpha_i \\ &\leq \varepsilon[\alpha(b) - \alpha(a)] + 2K\delta < \varepsilon[\alpha(b) - \alpha(a) + 2K]. \end{aligned}$$

Since ε was arbitrary, Theorem 6.6 implies that $h \in \mathscr{R}(\alpha)$.

Remark: This theorem suggests the question: Just what functions are Riemann-integrable? The answer is given by Theorem 11.33(*b*).

PROPERTIES OF THE INTEGRAL

6.12 Theorem

(*a*) *If* $f_1 \in \mathscr{R}(\alpha)$ *and* $f_2 \in \mathscr{R}(\alpha)$ *on* $[a, b]$, *then*

$$f_1 + f_2 \in \mathscr{R}(\alpha),$$

$cf \in \mathscr{R}(\alpha)$ *for every constant* c, *and*

$$\int_a^b (f_1 + f_2)\, d\alpha = \int_a^b f_1\, d\alpha + \int_a^b f_2\, d\alpha,$$

$$\int_a^b cf\, d\alpha = c \int_a^b f\, d\alpha.$$

(*b*) *If* $f_1(x) \le f_2(x)$ *on* $[a, b]$, *then*

$$\int_a^b f_1\, d\alpha \le \int_a^b f_2\, d\alpha.$$

(*c*) *If* $f \in \mathscr{R}(\alpha)$ *on* $[a, b]$ *and if* $a < c < b$, *then* $f \in \mathscr{R}(\alpha)$ *on* $[a, c]$ *and on* $[c, b]$, *and*

$$\int_a^c f\, d\alpha + \int_c^b f\, d\alpha = \int_a^b f\, d\alpha.$$

(*d*) *If* $f \in \mathscr{R}(\alpha)$ *on* $[a, b]$ *and if* $|f(x)| \le M$ *on* $[a, b]$, *then*

$$\left| \int_a^b f\, d\alpha \right| \le M[\alpha(b) - \alpha(a)].$$

(*e*) *If* $f \in \mathscr{R}(\alpha_1)$ *and* $f \in \mathscr{R}(\alpha_2)$, *then* $f \in \mathscr{R}(\alpha_1 + \alpha_2)$ *and*

$$\int_a^b f\, d(\alpha_1 + \alpha_2) = \int_a^b f\, d\alpha_1 + \int_a^b f\, d\alpha_2\,;$$

if $f \in \mathscr{R}(\alpha)$ *and* c *is a positive constant, then* $f \in \mathscr{R}(c\alpha)$ *and*

$$\int_a^b f\, d(c\alpha) = c \int_a^b f\, d\alpha.$$

Proof If $f = f_1 + f_2$ and P is any partition of $[a, b]$, we have

(20)
$$L(P, f_1, \alpha) + L(P, f_2, \alpha) \le L(P, f, \alpha) \le U(P, f, \alpha) \le U(P, f_1, \alpha) + U(P, f_2, \alpha).$$

If $f_1 \in \mathscr{R}(\alpha)$ and $f_2 \in \mathscr{R}(\alpha)$, let $\varepsilon > 0$ be given. There are partitions P_j $(j = 1, 2)$ such that

$$U(P_j, f_j, \alpha) - L(P_j, f_j, \alpha) < \varepsilon.$$

These inequalities persist if P_1 and P_2 are replaced by their common refinement P. Then (20) implies

$$U(P, f, \alpha) - L(P, f, \alpha) < 2\varepsilon,$$

which proves that $f \in \mathscr{R}(\alpha)$.

With this same P we have

$$U(P, f_j, \alpha) < \int f_j \, d\alpha + \varepsilon \qquad (j = 1, 2);$$

hence (20) implies

$$\int f \, d\alpha \leq U(P, f, \alpha) < \int f_1 \, d\alpha + \int f_2 \, d\alpha + 2\varepsilon.$$

Since ε was arbitrary, we conclude that

$$\int f \, d\alpha \leq \int f_1 \, d\alpha + \int f_2 \, d\alpha. \tag{21}$$

If we replace f_1 and f_2 in (21) by $-f_1$ and $-f_2$, the inequality is reversed, and the equality is proved.

The proofs of the other assertions of Theorem 6.12 are so similar that we omit the details. In part (*c*) the point is that (by passing to refinements) we may restrict ourselves to partitions which contain the point c, in approximating $\int f \, d\alpha$.

6.13 Theorem *If $f \in \mathscr{R}(\alpha)$ and $g \in \mathscr{R}(\alpha)$ on $[a, b]$, then*

(*a*) $fg \in \mathscr{R}(\alpha)$;

(*b*) $|f| \in \mathscr{R}(\alpha)$ *and* $\left| \int_a^b f \, d\alpha \right| \leq \int_a^b |f| \, d\alpha.$

Proof If we take $\phi(t) = t^2$, Theorem 6.11 shows that $f^2 \in \mathscr{R}(\alpha)$ if $f \in \mathscr{R}(\alpha)$. The identity

$$4fg = (f + g)^2 - (f - g)^2$$

completes the proof of (*a*).

If we take $\phi(t) = |t|$, Theorem 6.11 shows similarly that $|f| \in \mathscr{R}(\alpha)$. Choose $c = \pm 1$, so that

$$c \int f \, d\alpha \geq 0.$$

Then

$$\left| \int f \, d\alpha \right| = c \int f \, d\alpha = \int cf \, d\alpha \leq \int |f| \, d\alpha,$$

since $cf \leq |f|$.

6.14 Definition The *unit step function* I is defined by

$$I(x) = \begin{cases} 0 & (x \leq 0), \\ 1 & (x > 0). \end{cases}$$

6.15 Theorem *If $a < s < b$, f is bounded on $[a, b]$, f is continuous at s, and $\alpha(x) = I(x - s)$, then*

$$\int_a^b f\, d\alpha = f(s).$$

Proof Consider partitions $P = \{x_0, x_1, x_2, x_3\}$, where $x_0 = a$, and $x_1 = s < x_2 < x_3 = b$. Then

$$U(P, f, \alpha) = M_2, \qquad L(P, f, \alpha) = m_2.$$

Since f is continuous at s, we see that M_2 and m_2 converge to $f(s)$ as $x_2 \to s$.

6.16 Theorem *Suppose $c_n \geq 0$ for $1, 2, 3, \ldots$, Σc_n converges, $\{s_n\}$ is a sequence of distinct points in (a, b), and*

$$\alpha(x) = \sum_{n=1}^{\infty} c_n I(x - s_n). \tag{22}$$

Let f be continuous on $[a, b]$. Then

$$\int_a^b f\, d\alpha = \sum_{n=1}^{\infty} c_n f(s_n). \tag{23}$$

Proof The comparison test shows that the series (22) converges for every x. Its sum $\alpha(x)$ is evidently monotonic, and $\alpha(a) = 0$, $\alpha(b) = \Sigma c_n$. (This is the type of function that occurred in Remark 4.31.)

Let $\varepsilon > 0$ be given, and choose N so that

$$\sum_{N+1}^{\infty} c_n < \varepsilon.$$

Put

$$\alpha_1(x) = \sum_{n=1}^{N} c_n I(x - s_n), \qquad \alpha_2(x) = \sum_{N+1}^{\infty} c_n I(x - s_n).$$

By Theorems 6.12 and 6.15,

$$\int_a^b f\, d\alpha_1 = \sum_{i=1}^{N} c_n f(s_n). \tag{24}$$

Since $\alpha_2(b) - \alpha_2(a) < \varepsilon$,

$$\left| \int_a^b f\, d\alpha_2 \right| \leq M\varepsilon, \tag{25}$$

where $M = \sup|f(x)|$. Since $\alpha = \alpha_1 + \alpha_2$, it follows from (24) and (25) that

$$\left|\int_a^b f\,d\alpha - \sum_{i=1}^{N} c_n f(s_n)\right| \leq M\varepsilon. \tag{26}$$

If we let $N \to \infty$, we obtain (23).

6.17 Theorem *Assume α increases monotonically and $\alpha' \in \mathscr{R}$ on $[a, b]$. Let f be a bounded real function on $[a, b]$.*

Then $f \in \mathscr{R}(\alpha)$ if and only if $f\alpha' \in \mathscr{R}$. In that case

$$\int_a^b f\,d\alpha = \int_a^b f(x)\alpha'(x)\,dx. \tag{27}$$

Proof Let $\varepsilon > 0$ be given and apply Theorem 6.6 to α': There is a partition $P = \{x_0, \ldots, x_n\}$ of $[a, b]$ such that

$$U(P, \alpha') - L(P, \alpha') < \varepsilon. \tag{28}$$

The mean value theorem furnishes points $t_i \in [x_{i-1}, x_i]$ such that

$$\Delta\alpha_i = \alpha'(t_i)\,\Delta x_i$$

for $i = 1, \ldots, n$. If $s_i \in [x_{i-1}, x_i]$, then

$$\sum_{i=1}^{n} |\alpha'(s_i) - \alpha'(t_i)|\,\Delta x_i < \varepsilon, \tag{29}$$

by (28) and Theorem 6.7(*b*). Put $M = \sup|f(x)|$. Since

$$\sum_{i=1}^{n} f(s_i)\,\Delta\alpha_i = \sum_{i=1}^{n} f(s_i)\alpha'(t_i)\,\Delta x_i$$

it follows from (29) that

$$\left|\sum_{i=1}^{n} f(s_i)\,\Delta\alpha_i - \sum_{i=1}^{n} f(s_i)\alpha'(s_i)\,\Delta x_i\right| \leq M\varepsilon. \tag{30}$$

In particular,

$$\sum_{i=1}^{n} f(s_i)\,\Delta\alpha_i \leq U(P, f\alpha') + M\varepsilon,$$

for all choices of $s_i \in [x_{i-1}, x_i]$, so that

$$U(P, f, \alpha) \leq U(P, f\alpha') + M\varepsilon.$$

The same argument leads from (30) to

$$U(P, f\alpha') \leq U(P, f, \alpha) + M\varepsilon.$$

Thus

$$|U(P, f, \alpha) - U(P, f\alpha')| \leq M\varepsilon. \tag{31}$$

Now note that (28) remains true if P is replaced by any refinement. Hence (31) also remains true. We conclude that

$$\left| \overline{\int_a^b} f\, d\alpha - \overline{\int_a^b} f(x)\alpha'(x)\, dx \right| \le M\varepsilon.$$

But ε is arbitrary. Hence

$$\overline{\int_a^b} f\, d\alpha = \overline{\int_a^b} f(x)\alpha'(x)\, dx, \tag{32}$$

for *any* bounded f. The equality of the lower integrals follows from (30) in exactly the same way. The theorem follows.

6.18 Remark The two preceding theorems illustrate the generality and flexibility which are inherent in the Stieltjes process of integration. If α is a pure step function [this is the name often given to functions of the form (22)], the integral reduces to a finite or infinite series. If α has an integrable derivative, the integral reduces to an ordinary Riemann integral. This makes it possible in many cases to study series and integrals simultaneously, rather than separately.

To illustrate this point, consider a physical example. The moment of inertia of a straight wire of unit length, about an axis through an endpoint, at right angles to the wire, is

$$\int_0^1 x^2\, dm \tag{33}$$

where $m(x)$ is the mass contained in the interval $[0, x]$. If the wire is regarded as having a continuous density ρ, that is, if $m'(x) = \rho(x)$, then (33) turns into

$$\int_0^1 x^2\, \rho(x)\, dx. \tag{34}$$

On the other hand, if the wire is composed of masses m_i concentrated at points x_i, (33) becomes

$$\sum_i x_i^2\, m_i. \tag{35}$$

Thus (33) contains (34) and (35) as special cases, but it contains much more; for instance, the case in which m is continuous but not everywhere differentiable.

6.19 Theorem (change of variable) *Suppose φ is a strictly increasing continuous function that maps an interval $[A, B]$ onto $[a, b]$. Suppose α is monotonically increasing on $[a, b]$ and $f \in \mathscr{R}(\alpha)$ on $[a, b]$. Define β and g on $[A, B]$ by*

$$\beta(y) = \alpha(\varphi(y)), \qquad g(y) = f(\varphi(y)). \tag{36}$$

Then $g \in \mathscr{R}(\beta)$ and

$$\int_A^B g\, d\beta = \int_a^b f\, d\alpha. \tag{37}$$

Proof To each partition $P = \{x_0, \ldots, x_n\}$ of $[a, b]$ corresponds a partition $Q = \{y_0, \ldots, y_n\}$ of $[A, B]$, so that $x_i = \varphi(y_i)$. All partitions of $[A, B]$ are obtained in this way. Since the values taken by f on $[x_{i-1}, x_i]$ are exactly the same as those taken by g on $[y_{i-1}, y_i]$, we see that

$$U(Q, g, \beta) = U(P, f, \alpha), \qquad L(Q, g, \beta) = L(P, f, \alpha). \tag{38}$$

Since $f \in \mathscr{R}(\alpha)$, P can be chosen so that both $U(P, f, \alpha)$ and $L(P, f, \alpha)$ are close to $\int f\, d\alpha$. Hence (38), combined with Theorem 6.6, shows that $g \in \mathscr{R}(\beta)$ and that (37) holds. This completes the proof.

Let us note the following special case:

Take $\alpha(x) = x$. Then $\beta = \varphi$. Assume $\varphi' \in \mathscr{R}$ on $[A, B]$. If Theorem 6.17 is applied to the left side of (37), we obtain

$$\int_a^b f(x)\, dx = \int_A^B f(\varphi(y))\varphi'(y)\, dy. \tag{39}$$

INTEGRATION AND DIFFERENTIATION

We still confine ourselves to real functions in this section. We shall show that integration and differentiation are, in a certain sense, inverse operations.

6.20 Theorem *Let $f \in \mathscr{R}$ on $[a, b]$. For $a \le x \le b$, put*

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

Then F is continuous on $[a, b]$; furthermore, if f is continuous at a point x_0 of $[a, b]$, then F is differentiable at x_0, and

$$F'(x_0) = f(x_0).$$

Proof Since $f \in \mathscr{R}$, f is bounded. Suppose $|f(t)| \le M$ for $a \le t \le b$. If $a \le x < y \le b$, then

$$|F(y) - F(x)| = \left| \int_x^y f(t)\, dt \right| \le M(y - x),$$

by Theorem 6.12(c) and (d). Given $\varepsilon > 0$, we see that

$$|F(y) - F(x)| < \varepsilon,$$

provided that $|y-x| < \varepsilon/M$. This proves continuity (and, in fact, uniform continuity) of F.

Now suppose f is continuous at x_0. Given $\varepsilon > 0$, choose $\delta > 0$ such that

$$|f(t) - f(x_0)| < \varepsilon$$

if $|t - x_0| < \delta$, and $a \le t \le b$. Hence, if

$$x_0 - \delta < s \le x_0 \le t < x_0 + \delta \qquad \text{and} \qquad a \le s < t \le b,$$

we have, by Theorem 6.12(*d*),

$$\left|\frac{F(t) - F(s)}{t - s} - f(x_0)\right| = \left|\frac{1}{t - s}\int_s^t [f(u) - f(x_0)]\, du\right| < \varepsilon.$$

It follows that $F'(x_0) = f(x_0)$.

6.21 The fundamental theorem of calculus *If $f \in \mathscr{R}$ on $[a, b]$ and if there is a differentiable function F on $[a, b]$ such that $F' = f$, then*

$$\int_a^b f(x)\, dx = F(b) - F(a).$$

Proof Let $\varepsilon > 0$ be given. Choose a partition $P = \{x_0, \ldots, x_n\}$ of $[a, b]$ so that $U(P, f) - L(P, f) < \varepsilon$. The mean value theorem furnishes points $t_i \in [x_{i-1}, x_i]$ such that

$$F(x_i) - F(x_{i-1}) = f(t_i)\, \Delta x_i$$

for $i = 1, \ldots, n$. Thus

$$\sum_{i=1}^n f(t_i)\, \Delta x_i = F(b) - F(a).$$

It now follows from Theorem 6.7(*c*) that

$$\left|F(b) - F(a) - \int_a^b f(x)\, dx\right| < \varepsilon.$$

Since this holds for every $\varepsilon > 0$, the proof is complete.

6.22 Theorem (integration by parts) *Suppose F and G are differentiable functions on $[a, b]$, $F' = f \in \mathscr{R}$, and $G' = g \in \mathscr{R}$. Then*

$$\int_a^b F(x)g(x)\, dx = F(b)G(b) - F(a)G(a) - \int_a^b f(x)G(x)\, dx.$$

Proof Put $H(x) = F(x)G(x)$ and apply Theorem 6.21 to H and its derivative. Note that $H' \in \mathscr{R}$, by Theorem 6.13.