INTEGRATION LECTURE NOTES

by E B Davies

1. Introduction

If f is a real-valued function on a bounded interval $[a,b]$ the integral $\int_a^b f(x)\,\mathrm{d}x$ is intended to be a measure of the area under the graph of the function. If the function takes both positive and negative values then one defines the integral to be the difference of the areas of A and B where

$$A = \{(x,y) : \ a \le x \le b \ \text{ and } 0 < y < f(x)\}$$
$$B = \{(x,y) : \ a \le x \le b \ \text{ and } 0 > y > f(x)\}.$$

The above idea of integral transfers the problem of giving a precise definition of integral to that of giving a precise definition of area, which turns out to be no easier. We shall not attempt to define integration from first axioms in these notes, but concentrate on the properties that a successful integration procedure should have, and then develop the theory from there.

The first issue is that one cannot hope to integrate every conceivable function, obtaining a well-defined real number as its integral. In some cases the integral is infinite, in others both sets A and B have infinite areas, and there is no sensible meaning to $\infty - \infty$, while in yet others the function may be so irregular that it is not clear how to start to define its integral.

In these notes we only consider the integrals of piecewise continuous functions on a bounded interval $[a,b]$. Piecewise continuous functions $f : [a,b] \to \mathbb{R}$ are functions for which there exist numbers

$$a = a_0 < a_1 < \cdots < a_n = b$$

such that f is continuous on each interval (a_{k-1}, a_k) and the limits

$$\lim_{x \to a_k - 0} f(x)\,, \qquad \lim_{x \to a_k + 0} f(x)$$

exist for all relevant k. In other words, this means that f has a jump discontinuity of size $f(a_k + 0) - f(a_k - 0)$ at each point a_k, but is otherwise continuous. The space PC of such functions contains the space $\mathcal{S}$ of step functions and the space C of continuous functions on $[a,b]$ (recall that a step function f is a piecewise continuous functions in the sense of our definition, which takes constant values on each interval (a_{k-1}, a_k)).

It is evident from the definition that if f and g are piecewise continuous functions on $[a,b]$ and $\alpha, \beta \in \mathbb{R}$ then the combination $(\alpha f + \beta g)(x) = \alpha f(x) + \beta g(x)$ is also piecewise continuous. The set of jump points of $\alpha f + \beta g$ is just the union of the jump points of f and of g, or possibly a smaller set if the new function has a jump of size zero between two consecutive intervals. We summarize this by saying that PC is a vector space of functions. It is not finite-dimensional because it contains the polynomials of all orders.

If we write

$$I(f) := \int_a^b f(x)\,\mathrm{d}x$$

to abbreviate the notation, then $I : PC \to \mathbb{R}$ is a real valued function on PC, assigning the value $I(f) \in \mathbb{R}$ to each piecewise continuous function $f \in PC$. Three properties of this integral are of importance.

Proposition 1.

(1) The map $I :\ PC \to \mathbb{R}$ is linear, that is, $I(\alpha f+\beta g) = \alpha I(f)+\beta I(g)$ for all $\alpha, \beta \in \mathbb{R}$ and $f, g \in PC$.

(2) If $f \in PC$ and $f \geq 0$ in the sense that $f(x) \geq 0$ for all $x \in [a, b]$ then $I(f) \geq 0$.

(3) If $f(x) = c$ for all $x \in [a, b]$ then $I(f) = c(b - a)$.

(4) If $a \leq b \leq c$ then $I_{a,c}(f) = I_{a,b}(f) + I_{b,c}(f)$, where the subscripts denote the interval of integration.

All of the other properties of integrals can be proved from the above five statements, which we take as **axioms**, in the same way as we assumed the least upper bound property for real numbers, rather than proving it by constructing the reals from first principles. We need to know that $I(f) = 0$ if $a = b$, but this follows from hypothesis (3).

Elementary but important consequences of Proposition 1 are:

(5) If $f \geq g$ in the natural sense then $I(f) \geq I(g)$.

(6) $|I(f)| \leq I(|f|)$ for all $f \in PC$.

(7) If $|f(x)| \leq c$ for all $x \in [a, b]$ then $|I(f)| \leq c(b - a)$.

(8) If $|f(x) - g(x)| \leq \varepsilon$ for all $x \in [a, b]$ then $|I(f) - I(g)| \leq \varepsilon(b - a)$.

Remark. If $\varepsilon > 0$ is small, then (8) can be rephrased as: if f and g are uniformly very close to each other then their integrals are almost equal.

Proofs.

(5) By (1) and (2), $I(f) - I(g) = I(f - g) \geq 0$.

(6) Since $f(x) \leq |f(x)|$, (5) implies that $I(f) \leq I(|f|)$. On the other hand, since $-|f| \leq f$, from (1) and (5) it follows that $-I(|f|) = I(-|f|) \leq I(f)$. Finally, the inequalities $-I(|f|) \leq I(f) \leq I(|f|)$ are equivalent to the estimate $|I(f)| \leq I(|f|)$.

(7) By the above, $|I(f)| \leq I(|f|) \leq I(c)$. According to (3), the right hand side coincides with $c\,(b - a)$.

(8) By (1), $I(f) - I(g) = I(f - g)$. Therefore (8) is a special case of (7).

From now on we will use the standard notation $\int_a^b f(x)\,\mathrm{d}x$ rather than our new notation $I(f)$, but we insist on basing our proofs on the properties listed above.

2. Theorems about integration

Our goal in this section is to show that all the standard theorems of integration theory follow from Proposition 1. The most important of these is the fundamental theorem of calculus, which we separate into two parts.

We define the indefinite integral F of a piecewise continuous function f on $[a,b]$ by $F(x) = \int_a^x f(s)\,\mathrm{d}s$, where $x \in [a,b]$ is considered as a variable, so that F is a function of x.

Theorem 2. The indefinite integral $F(x)$ is a continuous function of x on $[a,b]$.

Proof. Let $c \in [a,b]$ and $\varepsilon > 0$. We need to show that there exists $\delta > 0$ such that $|F(x) - F(c)| < \varepsilon$ whenever $|x - c| < \delta$.

If $x > c$ then the axioms (4) implies that

$$|F(x) - F(c)| \;=\; \left|\int_a^x f(s)\,\mathrm{d}s - \int_a^c f(s)\,\mathrm{d}s\right| \;=\; \left|\int_c^x f(s)\,\mathrm{d}s\right|.$$

Now from (7) it follows that $|F(x) - F(c)| \le (x-c)\,\sup|f|$. Thus we can take $\delta = \varepsilon\,(\sup|f|)^{-1}$. This proves the right continuity.

If $x < c$ then similar arguments show that $|F(x) - F(c)| = \left|\int_x^c f(s)\,\mathrm{d}s\right|$ and, consequently, $|F(x) - F(c)| \le (c-x)\,\sup|f|$. This implies the left continuity at c.

The first half of the fundamental theorem of calculus is:

Theorem 3. If f is piecewise continuous on $[a,b]$ and continuous at a particular point $x \in (a,b)$ then F is differentiable at x and $F'(x) = f(x)$.

Proof. If $\delta > 0$ then

$$\delta^{-1}\left(F(x+\delta) - F(x)\right) = \delta^{-1}\left(\int_a^{x+\delta} f(s)\,\mathrm{d}s - \int_a^x f(s)\,\mathrm{d}s\right) = \delta^{-1}\int_x^{x+\delta} f(s)\,\mathrm{d}s.$$

Therefore

$$\begin{aligned}|\delta^{-1}\left(F(x+\delta) - F(x)\right) - f(x)| \;&=\; \left|\delta^{-1}\int_x^{x+\delta} f(s)\,\mathrm{d}s - \delta^{-1}\int_x^{x+\delta} f(x)\,\mathrm{d}s\right| \\ &=\; \delta^{-1}\left|\int_x^{x+\delta}\{f(s) - f(x)\}\,\mathrm{d}s\right| \;\le\; \delta^{-1}\int_x^{x+\delta}|f(s) - f(x)|\,\mathrm{d}s.\end{aligned}$$

Since f is continuous at x, for any $\varepsilon > 0$ there exists $\delta > 0$ such that if $|x - s| < \delta$ then $|f(s) - f(x)| < \varepsilon$. Hence

$$|\delta^{-1}\{F(x+\delta) - F(x)\} - f(x)| \;\le\; \delta^{-1}\int_x^{x+\delta}|f(s) - f(x)|\,\mathrm{d}s \;\le\; \delta^{-1}\int_x^{x+\delta}\varepsilon\,\mathrm{d}s \;=\; \varepsilon.$$

This proves that F is differentiable from the right with right derivative $f(x)$. The proof of the left differentiability is similar.

The proof of our next result uses the following consequence of the mean value theorem.

Lemma 4. If a continuous function $f : [a, b] \to \mathbb{R}$ is differentiable and $f'(x) = 0$ for all $x \in (a, b)$, then f is a constant function.

We are now able to prove the other half of the fundamental theorem of calculus, in which we differentiate before integrating, rather than after.

Theorem 5. Suppose $f : [a, b] \to \mathbb{R}$ is continuous and has a continuous derivative on (a, b) with finite limits $\lim_{x\to a+0} f'(x)$ and $\lim_{x\to b-0} f'(x)$ at the end points. Then

$$f(x) = f(a) + \int_a^x f'(s)\,\mathrm{d}s\,, \qquad \forall x \in [a, b].$$

Proof. Let

$$G(x) = f(a) + \int_a^x f'(s)\,\mathrm{d}s \;-\; f(x)\,.$$

It follows from Theorem 3 that $G'(x) = 0$ for all $x \in (a, b)$. Since $G(a) = 0$, Lemma 4 now establishes that G is identically zero.

Other standard integration procedures can now be proved easily, for example the procedure for integrating by parts.

Theorem 6. If functions $f, g : [a, b] \to \mathbb{R}$ satisfy the conditions of Theorem 5 then

$$\int_a^b f(s)g'(s)\mathrm{d}s = f(b)g(b) - f(a)g(a) - \int_a^b f'(s)g(s)\,\mathrm{d}s.$$

Proof. If we rewrite the formula as

$$f(b)g(b) = f(a)g(a) + \int_a^b (fg)'(s)\,\mathrm{d}s$$

then it becomes a special case of Theorem 5.

The next theorem justifies the process of integration by substitution.

Theorem 7. Let $f : [a, b] \to \mathbb{R}$ be a strictly increasing function satisfying the conditions of Theorem 5. If $f(a) = c$, $f(b) = d$ and g is a continuous function on the interval $[c, d]$ then

$$\int_c^d g(y)\,\mathrm{d}y = \int_a^b g\left(f(x)\right) f'(x)\,\mathrm{d}x.$$

Proof. Consider the function

$$G(s) = \int_c^{f(s)} g(y)\,\mathrm{d}y - \int_a^s g\left(f(x)\right) f'(x)\,\mathrm{d}x.$$

Direct calculations show that $G(a) = 0$ and $G'(s) = 0$ for all s. Therefore G is identically zero and $G(b) = 0$.

Example. The integral $\int_a^b f(x^2)\,x\,\mathrm{d}x$ can be rewritten in the form $\frac{1}{2}\int_a^b f(g(x))g'(x)\,\mathrm{d}x$ with $g(x) = x^2$. Applying Theorem 7, we obtain

$$\int_a^b f(x^2)\,x\,\mathrm{d}x \;=\; \frac{1}{2}\int_{a^2}^{b^2} f(x)\,\mathrm{d}x\,.$$

Theorem 8. Let $a(s)$ and $b(s)$ be differentiable functions of s taking values in an interval $[A, B]$ such that $a(s) \le b(s)$, and let $g[A, B] \to \mathbb{R}$ be a continuous function of x. Then

$$\frac{\mathrm{d}}{\mathrm{d}s}\int_{a(s)}^{b(s)} g(x)\,\mathrm{d}x = g(b(s))b'(s) - g(a(s))a'(s)$$

for all s.

Proof. If we define

$$F(s) := \int_A^s g(x)\,\mathrm{d}x$$

then the integral equals $F(b(s)) - F(a(s))$ and this is a restatement of the rule

$$\frac{\mathrm{d}}{\mathrm{d}s}[F(b(s)) - F(a(s))] = F^{'}(b(s))b^{'}(s) - F^{'}(a(s))a^{'}(s).$$

The differentiability of the LHS and the validity of the formula are standard results in the theory of differentiation.

Definition. If $x < y$ then we define $\int_y^x f(s)\,\mathrm{d}s = -\int_x^y f(s)\,\mathrm{d}s$. This is just a convenient agreement, under which all previous statements about integrals of the form $\int_y^x$ remain valid without the assumption $y < x$.

3. Integrating and differentiating power series

Let $f(x) = \sum_{n=0}^{\infty} c_n\,x^n$, and let $\hat{R}$ be the radius of convergence. Recall that that the function f defined by the power series is continuous on $(-\hat{R}, \hat{R})$ and the power series converges uniformly on every closed subinterval interval $[a, b] \subset (-\hat{R}, \hat{R})$, that is, there exists a sequence ε_m such that $\varepsilon_m \to 0$ as $m \to \infty$ and

$$\left| f(x) - \sum_{n=0}^{m} c_n\,x^n \right| \;\le\; \varepsilon_n\,, \qquad \forall x \in [a, b]\,.$$

Theorem 9. We have $\int_0^x f(y)\,\mathrm{d}y = \sum_{n=0}^{\infty} c_n\,(n+1)^{-1}\,x^{n+1}$, where the series in the right hand side has the same radius of convergence $\hat{R}$.

Proof. The radius of convergence of the power series $\sum_{n=0}^{\infty} c_n\,(n+1)^{-1}\,x^{n+1}$ is equal to $\left(\limsup |c_n\,(n+1)^{-1}|^{1/n}\right)^{-1}$ (see Week 4). By definition of the upper limit, there exist positive integers $n_1 < n_2 < n_3 \ldots$ such that

$$\limsup_{n\to\infty} \left|c_n\,(n+1)^{-1}\right|^{1/n} \;=\; \lim_{k\to\infty} \left|c_{n_k}\,(n_k+1)^{-1}\right|^{1/n_k} \;=\; \lim_{k\to\infty} |c_{n_k}|^{1/n_k}\,(n_k+1)^{-1/n_k}\,.$$

By the “powers beat logarithm” theorem (see Numbers and Functions lecture notes), $\frac{\log(n+1)}{n} \to 0$ and, consequently,

$$(n_k+1)^{-1/n_k} \;=\; e^{\frac{-\log(n_k+1)}{n_k}} \;\to\; 1 \quad \text{as } k\to\infty\,.$$

It follows that

$$\lim_{k\to\infty} |c_{n_k}|^{1/n_k}\,(n_k+1)^{-1/n_k} = \lim_{k\to\infty} |c_{n_k}|^{1/n_k} \lim_{k\to\infty}(n_k+1)^{-1/n_k} = \lim_{k\to\infty} |c_{n_k}|^{1/n_k}$$

for any sequence of positive integers $n_1 < n_2 < n_3 \ldots$. Since $\hat{R} = \left(\limsup |c_n|^{1/n}\right)^{-1}$, it follows that the radius of convergence of the series $\sum_{n=0}^{\infty} c_n\,(n+1)^{-1}\,x^{n+1}$ is equal to $\hat{R}$.

It remains to show that $\int_0^x f(y)\,\mathrm{d}y = \sum_{n=0}^{\infty} c_n\,(n+1)^{-1}\,x^{n+1}$ for all $x \in (-\hat{R},\hat{R})$. Assume, for the sake of definiteness, that $x > 0$. If $x < 0$, the equality is proved by applying the same arguments to the interval $[x,0]$.

Let $f_m(y) = \sum_{n=0}^{m} c_n\,y^n$. Since the series is uniformly convergent on the closed subinterval $[0,x] \subset (-\hat{R},\hat{R})$, there exist constants $\varepsilon_m > 0$ such that $|f(y) - f_m(y)| \le \varepsilon_m$ for all $y \in [0,x]$ and $\varepsilon_m \to 0$ as $m \to \infty$.

Clearly, $\int_0^x f_m(y)\,\mathrm{d}y = \sum_{n=0}^{m} c_n\,\frac{x^{n+1}}{n+1}$. Also, from the basic properties of the integral it follows that

$$\left|\int_0^x f(y)\,\mathrm{d}y - \int_0^x f_m(y)\,\mathrm{d}y\right| \le \left|\int_0^x (f(y) - f_m(y))\,\mathrm{d}y\right|$$
$$\le \int_0^x |f(y) - f_m(y)|\,\mathrm{d}y \le \int_0^x \varepsilon_m\,\mathrm{d}y \le \varepsilon_m\,x\,.$$

The right hand side converges to zero as $m \to \infty$, which implies that

$$\sum_{n=0}^{\infty} c_n\,(n+1)^{-1}\,x^{n+1} = \lim_{m\to\infty}\sum_{n=0}^{m} c_n\,\frac{x^{n+1}}{n+1} = \lim_{m\to\infty}\int_0^x f_m(y)\,\mathrm{d}y = \int_0^x f(y)\,\mathrm{d}y\,,$$

that is, $\int_0^x f(y)\,\mathrm{d}y \;=\; \sum_{n=0}^{\infty} c_n\,(n+1)^{-1}\,x^{n+1}$.

Corollary 10. Let $f(x) = \sum_{n=1}^{\infty} c_n\,x^n$ and $g(x) = \sum_{n=1}^{\infty} c_n\,nx^{n-1}$, and let $\hat{R}$ be the radius of convergence of the second series. Then the first series also converges on $(-\hat{R},\hat{R})$ and $g = f'$ for all $x \in (-\hat{R},\hat{R})$.

Proof. Integrating the series $\sum_{n=1}^{\infty} c_n\,ny^{n-1}$ from 0 to x we obtain $\sum_{n=1}^{\infty} c_n\,x^n$. Therefore $\sum_{n=1}^{\infty} c_n\,x^n$ has the same radius of convergence, and the required result is a consequence of the previous theorem and the fundamental theorem of calculus.

4. Improper integrals

This section is not examinable and, therefore, has been removed from the notes.