THEORY OF INTEGRATION

HENSTOCK

BUTTERWORTHS

THEORY OF INTEGRATION

RALPH HENSTOCK, M.A., PH.D.

Reader in Pure Mathematics

The Queen's University, Belfast

LONDON

BUTTERWORTHS

1963

THEORY OF INTEGRATION

BUTTERWORTHS MATHEMATICAL TEXTS

CONSULTANT EDITOR P. T. LANDSBERG

In Production

FACTOR ANALYSIS AS A STATISTICAL METHOD
(LAWLEY AND MAXWELL)

OPTIMAL TRAJECTORIES FOR SPACE NAVIGATION
(LAWDEN)

In Preparation

THE THEORY OF RE-ENTRY INTO THE ATMOSPHERE

STATISTICAL PROPERTIES OF LANGUAGE AND MUSIC

TOPOLOGY

MATRIX MODELS IN ECONOMICS

A LINEAR REGRESSION ANALYSIS

MEASURE THEORY

THEORY OF SATELLITE ORBITS IN THE PRESENCE OF AN ATMOSPHERE

THE STOCHASTIC APPROACH TO CHEMICAL KINETICS

THEORY OF LIQUIDS

LATTICE STATISTICS

INFORMATION THEORY OF STOCHASTIC CHAINS

PREFACE

It gives me great pleasure to acknowledge the advice of Professor S. Verblunsky, who has read the text, and also the following lectures and notes—

(i) Lectures on Lebesgue integration by J. C. Burkill (1942), R. G. Cooke (1944), A. S. Besicovitch (1947).

(ii) Notes on Stieltjes integration by P. Dienes (1944).

(iii) Lectures on statistics and stochastic processes by J. O. Irwin (1943), J. Wishart (1947), M. S. Bartlett (1947).

Most of the book has been given in lecture form at the Queen's University, Belfast. Two research students, Mr. James McGrotty and Mr. Chan Kai Meng, have read the proofs. Nevertheless, I am solely responsible for errors and omissions that remain.

THE QUEEN'S UNIVERSITY
BELFAST

RALPH HENSTOCK

ENGLAND BUTTERWORTH & CO. (PUBLISHERS) LTD.
London 88 Kingsway, W.C.2

AFRICA BUTTERWORTH & CO. (AFRICA) LTD.
DURBAN 33/35 Beach Grove

AUSTRALIA BUTTERWORTH & CO. (AUSTRALIA) LTD.
SYDNEY 6-8 O'Connell Street
MELBOURNE 430 Bourke Street
BRISBANE 240 Queen Street

CANADA BUTTERWORTH & CO. (CANADA) LTD.
TORONTO 1367 Danforth Avenue, 6

NEW ZEALAND BUTTERWORTH & CO. (NEW ZEALAND) LTD.
WELLINGTON 49/51 Ballance Street
AUCKLAND 35 High Street

U.S.A. BUTTERWORTH INC.
WASHINGTON, D.C. 7235 Wisconsin Avenue, 14

Suggested U.D.C. number: 517·3

PRINTED BY OFFSET IN GREAT BRITAIN BY
WILLIAM CLOWES AND SONS LTD, LONDON AND BECCLES

CONTENTS

CHAPTER PAGE

Preface v

1. HISTORICAL INTRODUCTION . . . 1
1. General introduction—2. Areas, and the differential and integral calculus—3. Riemann, Riemann–Stieltjes and Burkill integration—4. A new approach to Newton's and Riemann's integrals—5. The Pollard–Getchell integral—6. Lebesgue and Radon integration—7. Special and general Denjoy integrals—8. Perron, Ward, and Burkill's approximate Perron integrals—9. Extensions to 1958—10. The N-, N-variational, and Riemann-complete integrals

2. THE RIEMANN-COMPLETE INTEGRAL . 13
11. Notation—12. Countability—13. The real line—14. Open sets—15. Limit-points and closed sets—16. Left-complete and right-complete families—17. The integration of point functions—18. Functions of intervals—19. Distributive functional—20. Order relation—21. Additivity in elementary sets—22. Two integrability results—23. Stieltjes transformation

3. VARIATIONAL PROPERTIES OF THE INTEGRAL 39
24. The variational integral—25. Order relation and integrability—26. Variation—27. Inner variation—28. Properties of the variation and inner variation—29. Functions of generalized bounded variation—30. The variation of variationally equivalent functions—31. Special results—32. Sierpinski's lemma and the inner variation—33. Integration by parts

4. DIFFERENTIATION 74
34. Differentiation of VB^* and VBG^* functions—35. The differentiation of variational integrals

5. LIMITS UNDER THE INTEGRAL SIGN . 82
36. Monotone sequences and functions—37. Majorized sequences and functions—38. Absolute integration—39. Limits of step functions—40. An absolute integral is AC^*, and the integrability of a product

6. DOUBLE INTEGRALS AND FUBINI'S THEOREM 100
41. The plane—42. Functions of rectangles—43. The integration of rectangle functions—44. Fubini's theorem—45. Further results

7. THE CAUCHY AND DENJOY EXTENSIONS, AND THE INTEGRAL IN AN INFINITE RANGE 113
46. The Cauchy extension—47. The integral in an infinite range—48. The Denjoy extension

8. CONNECTIONS WITH EARLIER INTEGRALS 121
49. The integrals of Sections 2 to 7—50. The Perron and Ward integrals

9. LINEAR TOPOLOGICAL SPACES, YOUNG'S INEQUALITY AND INTEGRATION . . 127
51. Linear spaces—52. Linear topological spaces—53. Young's inequality—54. Mean convergence—55. The complete normed linear space $L^p (p > 1)$—56. The Riemann-complete integral for functions with values in a linear topological space—57. Functions of bounded variation

10. INTEGRATION IN STATISTICS 148
58. Introduction—59. Classification and relative frequency—60. Probability—61. Probability in three areas—62. Independence—63. The Neyman-Pearson theory of tests of simple hypotheses—64. Generating functions, moments

References 162

Index 165

1
HISTORICAL INTRODUCTION

1. General introduction

Most students begin real variable theory with the differential calculus, proceeding to the integral as the inverse of the derivative. In essentials they use the ideas of Newton and Leibnitz. Some never go beyond this stage; much is done in physics and applied mathematics by using this alone, without recourse to more modern ideas. However, it is becoming increasingly clear to scientists in general, and particularly to those interested in statistics, statistical physics and quantum theory, that such a limitation of the pure mathematics is a restriction, and that the use of better integration tools will result in a greater ease of application. The mathematician or scientist who wishes to go beyond the Newtonian inverse of a derivative could advantageously take for his tools the two new and equivalent integrals of this book. He will find that the usual theorems follow. In particular,

(i) the integral is additive in the functions to be integrated, and in the intervals in which they are integrated (Theorem 19.1, p. 27; Theorem 21.1, p. 31);

(ii) if the convergent sequence $\{f_n(x)\}$ is bounded by M independent of n and x, and if each $f_n(x)$ is integrable, then

$$\int_a^b \lim_{n\to\infty} f_n(x)\,\mathrm{d}x = \lim_{n\to\infty} \int_a^b f_n(x)\,\mathrm{d}x$$

(special case of Theorem 37.1, p. 85);

(iii) if in (ii), $f_n(x)$ is the sum of $g_m(x)$ for $m = 1, 2, \ldots, n$, then

$$\int_a^b \sum_{m=1}^{\infty} g_m(x)\,\mathrm{d}x = \sum_{m=1}^{\infty} \int_a^b g_m(x)\,\mathrm{d}x$$

(iv) $\dfrac{\mathrm{d}}{\mathrm{d}x} \displaystyle\int_a^x f(t)\,\mathrm{d}t = f(x)$

at almost all points x (special case of Theorem 35.1, p. 78);

(v) if, for each y, Y, $f(x, y)$ is integrable and

$$|f(x, Y)-f(x,y)| \leqslant M|Y-y|,$$

where M is independent of x, y, Y, then

$$\frac{\mathrm{d}}{\mathrm{d}y}\int_a^b f(x,y)\,\mathrm{d}x = \int_a^b \frac{\partial f(x,y)}{\partial y}\,\mathrm{d}x$$

(special case of Theorem 37.3, p. 88).

The theory is developed from the beginning, and only a knowledge of real and complex numbers is assumed. However, a familiarity with the ϵ, δ technique would be an advantage.

The ordinary student is advised to omit all sections that are starred, and also Chapter 8. These topics are intended for the more mature student who is acquainted with the Lebesgue integral at least. To him it may be pointed out that the very elegance and apparent finality of Lebesgue theory has caused mathematical inventiveness to move to other fields. Because of the Riesz representation theorem the integral has been regarded as nothing more than a linear functional. But integration is far more than this, as the present book partially shows. Further, there is no longer any necessity for considering separately the Riemann, Riemann–Stieltjes, Burkill, Pollard–Getchell, Lebesgue, Radon, special Denjoy, Perron and Ward integrals, for all these are included in the Riemann-complete and variational integrals of this book. Measure theory is not needed at the beginning, and the measure properties that are required later are obtained in a far simpler way than usual. Contrary to the general view, complete additivity of the measure is *not* required in order that we should have theorems of Lebesgue type; here we find that finite additivity is enough. In particular, Fatou's lemma is still valid, so that the Riesz–Fischer theorem and the completeness of L^p spaces follow as usual. For the latter see Chapter 9, in which we extend the integral to functions with values in linear topological spaces. We omit all integrals using convergence factors, and for these we refer to the literature (Henstock, (1960b, c; 1961a, b)). For simplicity we also restrict the functions to be functions of points or intervals on the real axis, or of points or rectangles on the plane. The n-dimensional case is a straightforward generalization, and the extension to an abstract space is given in Henstock (1961a, b).

But it is not advisable to restrict ourselves entirely to pure mathematics. There is ample precedent for a pure mathematician to consider in detail some applications to statistics (Chapter 10). This chapter is bound to be inconclusive and to leave much unsaid because of limitation of space, so that I have made an arbitrary selection of topics that illustrate the main theory. The Central Limit and allied theorems are adequately covered in other books, and so are omitted.

The first chapter is divided into two, because of the requirements of two kinds of reader. The beginner can read Sections 2 to 4 and then proceed to Chapter 2. The remaining sections are for the notice of the integration student who has already dealt with Lebesgue integration at least, to explain the relations between the new integrals and the old. We give a very brief summary of the position of integration theory up to, say, 1958. Because of the brevity many illustrious names have to be omitted, and we concentrate on generalizations suitable for problems in trigonometric series.

2. Areas, and the differential and integral calculus

In the simplest case the process of integration is the adding together of areas of non-overlapping elementary figures, and then the taking of some kind of a limit. The Greeks computed many simple areas, the methods being systematized through the years, and culminating in the *method of exhaustions* of Eudoxus (*c*.408–355 B.C.) and Archimedes (*c*.287–212 B.C.). This method was the first crude limit process, and they used the geometry of the figures to fit a sequence of non-overlapping triangles inside each main figure that finally exhausts the area. By this means they found the areas of the circle and sections of parabolas, for example, but could not define a general non-negative polynomial, and so could not compute the area under its curve.

The second approach to integration lies in inverting the result of differentiating a known function. The operation of differentiation was first systematized by I. Newton (1642–1727) and G. W. Leibnitz (1646–1716). To each of a certain class of functions f for which the derivative $Df = \mathrm{d}f/\mathrm{d}x$ exists, say, for x in $a \leqslant x \leqslant b$, we make correspond that derivative, so that we can regard D as an *operator*. It obeys the following rules. If f, g are differentiable functions of x in $a \leqslant x \leqslant b$, and if α, β are constants, then in $a \leqslant x \leqslant b$ we have

$$D(\alpha f+\beta g) = \alpha Df+\beta Dg \tag{2.1}$$

$$D(fg) = (Df)g+f(Dg) \tag{2.2}$$

$$D\{f(g(x))\} = (\mathrm{d}f/\mathrm{d}g)Dg \tag{2.3}$$

$$D\alpha = 0 \tag{2.4}$$

The rule for division is obtained from (2.2); if $f = h/g$ then

$$Dh = (Df)g+f(Dg)$$

$$D(h/g) = Df = \{Dh-(h/g)Dg\}/g$$

A function H of points x *is an indefinite Newton integral* of a known finite function f in $a \leqslant x \leqslant b$, if $DH = f$ in that interval. The functions that Newton integrated are all continuous, but we can ignore that limitation. Then the *definite Newton integral in* $a \leqslant x \leqslant b$ *is* $H(b)-H(a)$. We can write H as

$$H = D^{-1}f = (NL)\int f\,\mathrm{d}x,\ H(b)-H(a) = (NL)\int_a^b f\,\mathrm{d}x$$

where NL stands for Newton–Leibnitz. This definition of the integral is *descriptive*. No method of construction is offered, but we are given its properties so that we can recognize it if it is produced in another way. Because of this we have to prove that if H and H_1 are both indefinite Newton integrals of the same function f in $a \leqslant x \leqslant b$, then

$$H(b)-H(a) = H_1(b)-H_1(a) \tag{2.5}$$

To prove (2.5) we note that by (2.1)

$$D(H-H_1) = f-f = 0$$

so that in particular $H-H_1$ is continuous, and then the mean value theorem gives (2.5).

From (2.1) we obtain the distributivity of the Newton integral, namely,

$$D^{-1}(\alpha f+\beta g) = \alpha D^{-1}f+\beta D^{-1}g \tag{2.6}$$

From (2.2; 2.6) we have the formula for *integration by parts*,

$$D^{-1}(gDf)+D^{-1}(fDg) = fg$$

$$(NL)\int\left(\frac{f\,\mathrm{d}g}{\mathrm{d}x}\right)\mathrm{d}x = fg-(NL)\int\left(\frac{g\,\mathrm{d}f}{\mathrm{d}x}\right)\,\mathrm{d}x \tag{2.7}$$

From (2.3) we have

$$f(g(x)) = (NL)\int \frac{\mathrm{d}f}{\mathrm{d}g}\cdot\frac{\mathrm{d}g}{\mathrm{d}x}\,\mathrm{d}x$$

and replacing $\mathrm{d}f/\mathrm{d}g$ by f_1,

$$(NL)\int f_1(g)\,\mathrm{d}g = (NL)\int f_1\cdot\frac{\mathrm{d}g}{\mathrm{d}x}\,\mathrm{d}x \tag{2.8}$$

the formula for *integration by substitution.*

When we have defined more general integrals we will see that the formulae (2.5; 2.6; 2.7; 2.8) are in some sense still true for them.

The integration of a polynomial in x is now easy, but some simple functions cannot be integrated. It can be proved that if DH exists in $a \leqslant x \leqslant b$, and if γ is a number between $H'(a)$ and $H'(b)$, then there is a ξ in $a \leqslant \xi \leqslant b$ such that $H'(\xi) = \gamma$. It follows that if f is zero for x less than $\frac{1}{2}(a+b)$, and is 1 otherwise, then f does not have a Newton integral in $a \leqslant x \leqslant b$.

3. Riemann, Riemann–Stieltjes and Burkill integration

G. F. B. Riemann (1826–66) gave the following definition of the definite integral of a function f in $a \leqslant x \leqslant b$. Let

$$a = x_0 < x_1 < \ldots < x_n = b \tag{3.1}$$

be a division of $a \leqslant x \leqslant b$ into smaller intervals, let ξ_j be a point of the interval $x_{j-1} \leqslant x \leqslant x_j$, and consider the sum

$$S = \sum_{j=1}^{n} f(\xi_j)(x_j - x_{j-1}) \tag{3.2}$$

The number I *is the definite Riemann integral of f in $a \leqslant x \leqslant b$,* if to each $\epsilon > 0$ there is a $\delta > 0$ such that

$$|S-I| < \epsilon \tag{3.3}$$

whenever

$$x_j - _1 \leqslant \xi_j \leqslant x_j < x_{j-1}+\delta \quad (j = 1, 2, \ldots, n) \tag{3.4}$$

J. G. Darboux (1842–1917) made the following modification when f is real. He replaced $f(\xi_j)$ by the *supremum* (*least upper bound*) *of f in $x_{j-1} \leqslant x \leqslant x_j$*, and obtained an *upper sum.* For a *lower sum* he replaced $f(\xi_j)$ by the *infimum* *(greatest lower bound) of* f *in* $x_{j-1} \leqslant x \leqslant x_j$. If f is non-negative, with a given graph, and if we take a division (3.1) of $a \leqslant x \leqslant b$, then the upper Darboux sum is the sum of the areas of rectangles with bases the intervals $x_{j-1} \leqslant x \leqslant x_j$, and with just sufficient height to include the graph. The lower Darboux sum is the sum of the areas of rectangles with the same bases, but lying just below the graph. When f is real it is clear that for suitable choice of the ξ_j, the S of (3.2) can be taken arbitrarily near to the upper sum, and for another choice, arbitrarily near to the lower sum, so that the Darboux modification does not alter the Riemann integral of a real function. Thus if a real function has a Riemann integral in $a \leqslant x \leqslant b$ it must be bounded there. From this we can show that not every Newton integral is a Riemann integral. For

$$H(x) = x^2.\sin(1/x^2)\ (x \neq 0), H(0) = 0 \tag{3.5}$$

is differentiable everywhere, the derivative being unbounded in the neighbourhood of $x = 0$. However, not every Riemann integral is a Newton integral, for the Riemann integral of the last function of Section 2 exists in $a \leqslant x \leqslant b$, and is equal to $\frac{1}{2}(b-a)$. There is a common region, for the Riemann and Newton integrals of a continuous function exist and are equal. The Riemann integral cannot integrate every bounded function, for if

$$f(x) = \begin{cases} 1\ (x\ \textit{rational}) \\ 0\ (x\ \textit{irrational}) \end{cases} \tag{3.6}$$

then any upper Darboux sum is $b-a$, while any lower Darboux sum is 0. Thus f does not have a Riemann integral (nor a Newton integral).

The Riemann method has been modified in many ways. T. J. Stieltjes (1856–94) used another function g, replacing $x_j - x_{j-1}$ in (3.2) by

$$g(x_j) - g(x_j - _1)$$

The resulting integral is now called the *Riemann–Stieltjes integral.* J. C. Burkill replaced $f(\xi_j)(x_j - x_{j-1})$ in (3.2) by a function $h(x_{j-1}, x_j)$ of the interval from x_{j-1} to x_j, obtaining the *Burkill integral.* It is clear that the Riemann, Riemann–Stieltjes and Burkill integrals are *constructive* when they exist. For we can construct the definite integral I by finding the limit of sums (3.2) for Riemann integrals, or the corresponding sums for the other integrals, for special sequences of divisions (3.1) and special choices of the ξ_j, e.g. we can take $\xi_j = x_j = a+(b-a)j/2^n$ for $j = 0,1,2,...,2^n$ and $n = 1,2,...$

4. A new approach to Newton's and Riemann's integrals

Newton's definite integral can be written as a simple example of the kinds of integrals that we wish to study. Let H be the Newton indefinite integral of a finite function f in $a \leqslant x \leqslant b$. Then $DH = f$ there. Thus, given $\epsilon > 0$, there is a $\delta > 0$, depending on ϵ and x, such that for $0 < |t-x| \leqslant \delta$,

$$\left|\frac{H(t)-H(x)}{t-x} - f(x)\right| \leqslant \epsilon,$$

$$|H(t)-H(x)-f(x)(t-x)| \leqslant \epsilon|t-x| \tag{4.1}$$

Here, all intervals from t to x, where either $x-\delta \leqslant t < x$ or $x < t \leqslant x+\delta$, have this property. Suppose that from these intervals we can construct a division (3.1) of $a \leqslant x \leqslant b$. Then for ξ_j equal to one or other of x_{j-1}, x_j, we use (4.1) and obtain

$$\begin{aligned}|H(b)-H(a)-\sum_{j=1}^{n} f(\xi_j)(x_j-x_{j-1})| \\ = |\sum_{j=1}^{n} \{H(x_j)-H(x_{j-1})-f(\xi_j)(x_j-x_{j-1})\}| \\ \leqslant \sum_{j=1}^{n} |H(x_j)-H(x_{j-1})-f(\xi_j)(x_j-x_{j-1})| \\ \leqslant \sum_{j=1}^{n} \epsilon(x_j-x_{j-1}) = \epsilon(b-a)\end{aligned} \tag{3.2}$$

showing that for the special ξ_j and special divisions, the sum (3.2) tends to $H(b)-H(a)$ as $\epsilon \to 0$. Thus if the construction of divisions is possible, Newton's integral has a definition as an integral of Riemann type. In Chapter 2 we show that the construction is in fact possible. Further, the interval functions $f(\xi_j)(x_j-x_{j-1})$ are of two types,

$$f(x_{j-1})(x_j-x_{j-1}), \quad f(x_j)(x_j-x_{j-1}) \tag{4.2}$$

right-hand and left-hand, where the ξ_j is regarded as fixed, while the other end of the interval lies in a certain neighbourhood of ξ_j.

Again, in Riemann's definition we use interval functions

$$f(\xi_j)(x_j - x_{j-1})(x_{j-1} \leqslant \xi_j \leqslant x_j) \tag{4.2}$$

If $\xi_j = x_{j-1}$ or x_j, we again obtain an interval function (4.2). If $x_{j-1} < \xi_j < x_j$ we put

$$f(\xi_j)(x_j - x_{j-1}) = f(\xi_j)(\xi_j - x_{j-1}) + f(\xi_j)(x_j - \xi_j)$$

a sum of the two kinds in (4.2). Similarly for the Riemann–Stieltjes integral. There is no ξ_j in Burkill integration, so that it does not matter whether an interval is counted as left-hand or right-hand, the h is the same.

This section gives us two ideas that are developed in Chapter 2; namely, the use of intervals from t to x, for $|t-x| \leqslant \delta(x)$, and the use of two interval functions, one 'left-hand' and one 'right-hand'. The beginner can now proceed to Chapter 2.

The rest of Chapter 1 is designed for the student of integration who wishes to connect his previous knowledge with the theory of this book.

*5. The Pollard–Getchell integral

In Riemann–Stieltjes integration S. Pollard (1894–1945) and B. C. Getchell modified (3.4), supposing that for each $\epsilon > 0$, there is a division

$$a = x_0' < x_1' < \ldots < x_p' = b \tag{5.1}$$

with the property that if (3.1) is a subdivision of (5.1) and $x_{j-1} \leqslant \xi_j \leqslant x_j$ for $j = 1,2,\ldots,n$, then (3.3) is true. In this case *I is the Pollard-Getchell integral of f*. See Pollard (1923), Getchell (1935).

Every Riemann–Stieltjes integral is a Pollard–Getchell integral, for we can ensure (3.4) by taking, in (5.1)

$$x'_k - x'_{k-1} < \delta (k = 1, 2, \ldots, p)$$

The Pollard–Getchell integral enables us sometimes to integrate f with respect to g when f and g have common discontinuities and the Riemann–Stieltjes integral does not exist. The Pollard–Getchell modification can be applied to the Burkill integral, and a suitable construction of the divisions (5.1) in this case are given in Henstock (1946 and 1948). Clearly the Pollard–Getchell integral is constructive if suitable divisions can be constructed.

The integrals of Sections 3 and 5 are of Riemann type and