OBITUARY

JOHN CHARLES BURKILL

John Charles Burkill, born on 1 February 1900, was the only child of Hugh Roberson Burkill (1867–1951) and Bertha (*née* Bourne, 1866–1937). His father came from a family which had farmed in Lincolnshire for generations, whereas his mother came from a background of prosperous farming and building. On neither side was there a strong academic tradition, but Charles was soon to show evidence of intellectual distinction by winning a scholarship to St Paul's school at the age of 14. There he profited fully from the excellent teaching that the school offered and which was reflected not only by his mathematical prowess, which led to a scholarship to Trinity College, Cambridge, in 1918, but also in his ability in classical studies in which he maintained a lifelong interest. He was also a formidable chess player, and had a mischievous sense of humour which he retained, albeit in a more restrained mode, in later life. A striking example of his grasp of the essence of a practical joke is recorded in the story of how, as a boy on a visit to a house-proud aunt, he saw the comic potential of a trail of corn from the chicken run through the front door and upstairs to the bedrooms.

On leaving school in 1918, he joined the Royal Engineers (RE), but was demobilized soon after being commissioned as second lieutenant. However, this early military training was of service in 1939 when he joined the Cambridge University OTC as a second lieutenant and came to command the RE unit with the rank of major. He went up to Trinity in 1919, and stayed on successively as scholar, research student and Smith's Prizeman and fellow until 1924, when he was appointed at an unusually early age to the chair of pure mathematics at Liverpool.

Burkill returned to Cambridge in 1929 to take up a university lectureship and a fellowship, not at his old college but at Peterhouse; and there he stayed for the rest of his life, giving an example of loyalty and devotion to an institution and its people that would be difficult to match. His value and potential as a member of the governing body of the college were soon recognized by his early appointment as a tutor, an office which he held for a large part of his time at Peterhouse, including the war years when, in the absence on leave of the Master, he and a very small group of fellows ensured that the college not only survived but remained a centre of intellectual distinction and sound teaching.

College life was not easy in the years immediately after the war, and Burkill did not retire from the tutorship until 1948, and even served again in 1952 as acting Senior Tutor. His research had inevitably been hampered by a heavy administrative and teaching load, but his release from some of these responsibilities gave him more time, and this is reflected by his substantial output of papers at the time. He was awarded an Adams Prize in 1949, elected a Fellow of the Royal Society in 1953, and served on its Council from 1959 to 1961. He was made Emeritus Reader in Mathematical Analysis on his retirement.

Bull. London Math. Soc. 30 (1998) 85–98

In 1968 it was greatly to the advantage of the college that an amendment to its statutes made it possible for him to be elected Master beyond the normal age of retirement, and he served the college in this capacity until 1973. It was at the time of his election that active student dissatisfaction became a significant element in university affairs, and one of Burkill's many services to the college was to handle this in such a way that there was neither lasting dissension nor the imposition of statutory and bureaucratic involvement of students in all aspects of college government. Another distinctive feature of his mastership was his positive support for the development of graduate studies and research by increasing the number of fellows and graduate students, and this was achieved without weakening the high standards of teaching and pastoral care which he had fostered as tutor. His term as Master was followed by his appointment as editor of the *Mathematical Proceedings of the Cambridge Philosophical Society*, and the journal's high standing when he left it was a tribute to his achievement in an exacting task, which few scholars of his age were willing to contemplate. This work was indeed a sign of his sense of duty and integrity in everything he did, including his fastidious concern for accuracy and economy in the use of words. In respect of the spoken word, this economy became something of a legend. Taciturn is not a sufficiently friendly word to describe his conversational style, because it contained no hint of malice or lack of concern but only an unerring judgement about what was important, and the clearest way of saying it. What is even more important is that his distaste for excessive display of feeling concealed, at first, a truly generous and hospitable nature. He was a kindly man, and shared with his wife Greta a rare perception of the problems and needs of others, and any account of his life would be incomplete without reference to the remarkable qualities which she brought to their partnership.

Greta was the daughter of Adolf Braun, a distinguished journalist in pre-1914 Germany. Her mother was Russian and brought her to England, where she completed her education at school and at Newnham. Although she was herself neither Jewish nor a refugee, her early life had given her a deep and sympathetic understanding of people persecuted for their race, politics or religion, and she became a leading figure in the organization set up to rescue refugees from Hitler. She and Charles together, with their combined experience of international affairs and academic life, were particularly effective in helping many gifted scholars to escape and to contribute to the intellectual life of this country. They did this not only by good organization, but also by the example they set in taking young scholars into their own home and virtually adopting them.

Many of the refugees who came to Britain through the efforts of the Burkills were either mature scholars or research students whose work had been disrupted, and this must have been a major factor in stimulating, and extending beyond the college precincts, their interest in the welfare of graduate students in general. Cambridge was not a comfortable place for scholars without a firm college connection, and the provision of basic amenities for them was a pressing need. The founding of the Graduate Centre and the Cambridge Graduate Society was largely due to their joint efforts, for they made a powerful team, he with his grasp of practicalities and procedures and she with her formidable crusading zeal. In superficial ways, few couples could have seemed more different, but there was a real harmony in their partnership to enable them to do so much good.

https://doi.org/10.1112/S0024609397003767 Published online by Cambridge University Press

1. *Integration and differentiation*

Burkill's work is all in the theory of functions of a real variable, with its main emphasis on theories of differentiation and integration. This was a particularly active area of research in the early decades of this century, after the pioneering work of Lebesgue, Borel and their contemporaries in establishing the concepts of measure and the Lebesgue integral associated with it. These continue to play a central role in modern mathematical analysis, and provide a reference by which further developments can be compared and understood, and it may be useful to give a very brief account of some of the concepts associated with them.

Broad ideas about differentiation and integration go back to Newton and Leibniz, as do requirements about the formal relationship between them. The natural starting point is the definition of the *derivative* $F'(x)$ of a point function $F(x)$ as the limit in some sense of $h^{-1}[F(x+h)-F(x)]$ as $h \to 0$. Then integration, regarded as the inverse of differentiation, is any operation on a function $f(x)$ which produces a *primitive* or *integral* $F(x)$ with the property that $F'(x) = f(x)$; and we then write $F(x) = \int f(u)\,du$. This descriptive concept of integration is incomplete until we specify the precise definition of a derivative and the interpretation of the equality sign. It is also deficient in a more practical way in that it provides no method, other than organized guesswork, for actually finding the primitive of a given function.

The traditional alternative approach which remedies this is first to develop and make precise the concept of the *area* of a set of points in the plane (or *volume* in three or more dimensions) and to define $\int_a^b f(x)\,dx$ *constructively* to be the area of the set of points S bounded by the x-axis, the lines $x = a$, $x = b$, and the graph $y = f(x)$. The conclusion that these two definitions of an integral are, under appropriate conditions, equivalent is the *fundamental theorem of the calculus*, and is central to any theory of differentiation and integration. Since there is no preordained logical structure to any such theory, it is essential to make clear what is being defined and what is deduced.

The best known examples of the constructive approach are due to Riemann and Lebesgue; in spite of apparent similarities in their definition, they are quite distinct in their properties and in their potential for generalization. In each case the definition of area is based on the limit as $h \to 0$ of the sums of approximations to the areas of parts of S obtained by slicing S into sections of width h either vertically (Riemann) or horizontally (Lebesgue). Riemann requires only approximations by rectangles, while Lebesgue depends on the notion of the *measure* of the more complex set of points x for which $f(x) \geqslant y$. Important distinctions arise directly from the differences in the geometry of the constructions. For example, the existence and properties of the Riemann integral are bound up with the metric topology of the real line and the continuity of the integrand, whereas the Lebesgue integral requires only the existence of a measure, and measures can be defined in a great variety of spaces without reference to the nature, or even the existence, of their topological properties. However, the classical definition of a derivative is a topological concept, and we expect to find a fully satisfactory calculus only in cases in which measure and topology are properly related.

A comparison between the two integrals illustrates the important idea of the *scope* of a method of integration as the set of functions which can be integrated. Thus Lebesgue has greater scope than Riemann, but this is not an unconditional advantage since Lebesgue integrability is the weaker constraint and this may necessitate the strengthening of some other condition when it is part of the hypothesis of a theorem.

In the following summary of Burkill's publications, papers have been grouped to reflect his main areas of interest, and numbers refer to the list at the end of this memoir. It is convenient on occasion to retain his notation, using symbols $+$, $-$, $\cdot$ for set operations, and speaking of functions $g(I)$, $F(I)$, $f(x)$, $G(x)$ (as a reminder that the variable may be an interval or a point) despite the normal convention that such symbols should be used only for *values* of the functions g, F, f, G.

2. *Functions of intervals and the Burkill integral* [**1**–**3**]

An open n-dimensional interval I is defined as the set of points $(x_1, x_2, \dots, x_n)$ which satisfy $a_i < x_i < b_i$ $(i = 1, 2, \dots, n)$. The same numbers also define closed or partly closed intervals, and the distinction is generally immaterial, but I is taken to be open unless the contrary is indicated. An *interval function* $g(I)$ is defined over a system of intervals if a real number $g(I)$ is assigned as its value for each I of the system. The main aim of [**1**] is to give a systematic account of the basic properties of interval functions which are *not necessarily additive*. This means that it is not assumed that $g(I) = g(I_1) + g(I_2)$ when I is the union of abutting but non-overlapping intervals I_1, I_2 (including the interior points of their common boundary). The importance of this becomes clear when we note the many cases of nonlinear interval functions such as the elements in Riemann–Darboux sums or the ratio $h^{-1}[f(x+h) - f(x)]$ for the interval $(x, x+h)$.

The elementary properties of interval functions can now be established. For example, an interval function is *bounded* in an open or closed interval R if its values $g(I)$ are bounded for all intervals I in R. If $\omega(\delta, x)$ is the upper bound of $|g(I)|$ for every I in the square with centre x and side δ, then ω decreases with δ and $\omega(x) = \lim \omega(\delta, x)$ as $\delta \to 0$ exists and is called the *oscillation* of $g(I)$ at x. We say that $g(I)$ is *continuous* at x if $\omega(x) = 0$, and is continuous in R if continuous at every point of R. If R is closed, continuity implies *uniform continuity*, in the sense that, given $\varepsilon > 0$, we can define $\delta(\varepsilon) > 0$ so that $|g(I)| \leqslant \varepsilon$ for every I in R with *diameter* $n(I) \leqslant \delta$. A division of an interval R (two dimensions being typical) by lines parallel to one or other axis into a finite number of subintervals I_j is called a *mesh* $\{I_j\}$. We say that $g(I)$ has an integral l if, given $\varepsilon > 0$, we can define $\delta(\varepsilon) > 0$ so that $|\sum g(I_j) - l| < \varepsilon$ for all meshes $\{I_j\}$ with $\max n(I_j) \leqslant \delta(\varepsilon)$. Such a number l is unique if it exists, and is then written $\int_R g(I)$ (or $\int g$ if the context is clear) and is called (but not by himself) the *Burkill integral*. Whether the integral exists or not, the upper and lower Burkill integrals $\overline{\int}$, $\underline{\int}$ always exist as the upper and lower limits of $\sum g(I_j)$ as $\max n(I_j) \to 0$. Two familiar examples of the integral are:

(1) $\int g(I)$ is the total variation of $f(x)$ in one dimension if $g(I_j) = |f(x_j) - f(x_{j-1})|$;

(2) $\int_{(H,K)} g(I)$ is the Lebesgue integral $\int_a^b f(x)\,dx$ if I_j is the interval (y_{j-1}, y_j) and $g(I_j)$ is the measure of the set in which $y_{j-1} \leqslant f(x) < y_j$ and H, K are bounds of $f(x)$.

The main properties of the Burkill integral are as follows.

(i) If $g(I)$ is finitely additive, $\int_R g(I) = g(R)$.

(ii) (General principle of convergence.) The function $g(I)$ is integrable if and only if, given $\varepsilon > 0$, we can define $\delta(\varepsilon) > 0$ so that $|\sum g(I_j) - \sum g(I_k)| < \varepsilon$ for any two meshes with $\max[n(I_j), n(I_k)] \leqslant \delta$.

(iii) If $\int g$ exists, so does $\int cg$ for any constant c, and $\int cg = c \int g$.

(iv) If $g(I) \geqslant 0$, then $\int g \geqslant 0$.

(v) (Mean value theorem.) If $HmI \geqslant g(I) \geqslant KmI$ for I in R, then $HmR \geqslant \int g \geqslant KmR$.

https://doi.org/10.1112/S0024609397003767 Published online by Cambridge University Press

(vi) If $g = g_1 + g_2$ and any two of $\int g_1$, $\int g_2$, $\int g$ exist, then so does the third, and $\int g = \int g_1 + \int g_2$.

(vii) *Schwartz's inequality*:

$$\left\{\overline{\int} g_1 g_2\right\}^2 \leqslant \overline{\int} g_1^2 \overline{\int} g_2^2; \quad \left\{\underline{\int} p_1 p_2\right\}^2 \leqslant \underline{\int} p_1^2 \overline{\int} p_2^2 \quad \text{if } p_1 \geqslant 0,\ p_2 \geqslant 0.$$

(viii) The condition of integrability can be strengthened (with a decrease in scope of the integral) by allowing meshes in which the dividing lines need not extend right across R, and the integral so defined is called the *extended* Burkill integral and written $E\int$. Then if $R = R_1 + R_2$, it is not generally true that integrability over R_1 and R_2 implies integrability over R, and vice versa. But it is true that if g is integrable over R_1, R_2 and R, then

$$\int_R = \int_{R_1} + \int_{R_2}.$$

Also, if g is integrable E over R, it is integrable E over any subinterval of R.

While these properties are analogues of theorems in the traditional integral calculus, it is important to observe that the Burkill integral has an *interval* function as its integrand and is quite distinct in concept from the integrals (Riemann, Lebesgue, Perron, etc.) in which the integrand is a point function.

Some further properties of interval functions are needed to develop the integral. An interval function $g(I)$ is *absolutely continuous* (a.c.) in R if $\sum g(I_j) \to 0$ as $\sum mI_j \to 0$ and the I_j (finite or enumerable) are non-overlapping. The Lipschitz condition $|g(I)| \leqslant KmI$ obviously implies absolute continuity. These properties follow.

(i) If g, g_1, g_2 are a.c., then so are $|g|$ and $g_1 + g_2$.

(ii) If g_1 is a.c. and g_2 is bounded, then $g_1 g_2$ is a.c.

(iii) If $p = \frac{1}{2}[|g| + g]$, $n = \frac{1}{2}[|g| - g]$, so that $g = p - n$, $|g| = p + n$, $p \geqslant 0$, $n \geqslant 0$, and if g is a.c., then so are p and n.

(iv) If g is a.c. in R, then its extended upper and lower integrals are finite.

(v) If g is a.c. in R and $G(I) = \int_I g$ exists for every I in R, then $G(I)$ is a.c. in R.

If X is a measurable set in R so that, for a sequence ε_r decreasing to 0, we can decompose X as $X = J_r + e_r - e_r'$, where J_r is the union of a finite set of intervals and $me_r < \varepsilon_r$, $me_r' < \varepsilon_r$, it is proved that $G(J_r)$ tends to a limit which is independent of the particular sequence ε_r or the particular decomposition of X for any r. This limit is written $G(X) = \int_X g(I)$ and is called the (Burkill) integral of g over X.

(vi) $G(X)$ is a completely additive function of measurable sets in R, so that $G(\sum X_j) = \sum G(X_j)$ for any enumerable disjoint sets X_j in R.

(vii) If g is a.c. and $g(I) \leqslant g(I_1) + g(I_2)$ when $I = I_1 + I_2$, then g is integrable. In one dimension, the weaker condition that g is continuous may replace absolute continuity, although its integral may then be infinite.

3. *Derivatives of interval functions* [**4**, **6**, **7**]

Burkill's calculus is completed by defining the *derivative* of an interval function and relating it to the integral $G(X)$. If $0 < \rho \leqslant 1$, we define $u(\rho, x)$ as the upper limit of $g(I)/mI$ as $mI \to 0$ and $mI/mS \geqslant \rho$, where S is the smallest square with centre x

containing I. The lower limit $l(\rho, x)$ is defined similarly, and since both limits are monotonic in ρ, we can define $u(x)$, $l(x)$, their limits as $\rho \to 0$, as the *upper and lower derivatives* of $g(I)$ at x. If $u(x) = l(x)$, we say that $g(I)$ is *differentiable* at x and $g'(x) = u(x) = l(x)$ is called its *derivative*.

The basic properties of derivatives are then as follows.

(i) The necessary and sufficient condition that $g'(x)$ exists is $u(\rho, x) = l(\rho, x)$ for every ρ in $0 < \rho \leqslant 1$.

(ii) If $g = g_1 + g_2$, then $l_1(x) + l_2(x) \leqslant l(x) \leqslant l_1(x) + u_2(x) \leqslant u(x) \leqslant u_1(x) + u_2(x)$, and if $g_1'(x)$, $g_2'(x)$ exist, then so does $g'(x)$, and $g'(x) = g_1'(x) + g_2'(x)$.

(iii) For any $g(I)$, $u(x)$, $l(x)$, $g'(x)$ are measurable.

(iv) In one dimension, if $g(I) = f(x_j) - f(x_{j-1})$ when I is (x_{j-1}, x_j) and if $\bar{u}(x)$, $\bar{l}(x)$ are upper and lower derivatives of $f(x)$, either on the right or on the left, then $u(x) \geqslant \bar{u}(x)$, $l(x) \leqslant \bar{l}(x)$. The existence of either $f'(x)$ or $g'(x)$ implies the existence of the other, and the two are equal.

(v) If $\partial^2 f/\partial x \partial y$ exists near (x_0, y_0) and has upper and lower limits $M(x_0, y_0)$, $m(x_0, y_0)$ as $x \to x_0$, $y \to y_0$, then $M(x_0, y_0) \geqslant u(x_0, y_0) \geqslant l(x_0, y_0) \geqslant m(x_0, y_0)$. In particular, $g'(x_0, y_0) = \partial^2 f/\partial x \partial y$ at (x_0, y_0) if $\partial^2 f/\partial x \partial y$ is continuous at (x_0, y_0).

These properties are sufficient to establish analogues of the fundamental theorems of calculus.

(i) If $p(I) \geqslant 0$ and $E\int_R p(I)$ is finite, then $u(x)$ and $l(x)$ are finite *almost everywhere* (a.e.) in R. In particular, if $g(I)$ is a.c. in R, then $u(x)$ and $l(x)$ are finite a.e. in R. In one dimension with $p(I_j) = f(x_j) - f(x_{j-1})$, we have the familiar result that $u(x)$, $l(x)$ are finite a.e. if $f(x)$ has bounded variation.

(ii) If $g(I)$ is a.c. in R, then (Lebesgue integrals for $l(x)$, $u(x)$)

$$E\underline{\int} g(I) \leqslant \int u(x)\,dx \leqslant E\overline{\int} g(I).$$

In particular, if $E\int g(I)$ exists, so does $g'(x)$ a.e., and

$$E\underline{\int} g(I) \leqslant \int g'(x)\,dx \quad \text{and} \quad G(X) = \int_X g'(x)$$

for any measurable set X.

(iii) If $g(I)$ is a.c. in R and $l(x) \geqslant 0$ a.e. in R, then

$$\overline{\int} g(I) \geqslant 0.$$

(iv) If $g(I)$ is a.c. in R and $G(I)$ exists for every I in R, and if $l(x) \geqslant 0$ in a measurable set X, then $G(X) \geqslant 0$.

(v) If $g(I)$ is a.c. in R and $G(I)$ exists for every I in R, then

$$\int_X l(x)\,dx \leqslant G(X) \leqslant \int_X u(x)\,dx$$

for any measurable set X. In particular, if $g'(x)$ exists a.e. in X, then

$$G(X) = \int_X g'(x)\,dx.$$

(vi) If R is one-dimensional and $g(I)$ is a.c., then

$$\overline{\int} g(I) = \int u(x)\,dx, \quad \underline{\int} g(I) = \int l(x)\,dx,$$

and, in particular, the existence of either $\int g(I)$ or $\int g'(x)\,dx$ implies the existence of the other, and the two are equal.

(vii) If $\int_R g(I) < \infty$, then the set of points at which $u(x) = \infty$ and $l(\frac{1}{2}, x) > -\infty$ has measure zero.

(viii) If $\int_R g(I)$ exists, then the set of points at which $u(x)$ and $l(x)$ are finite and unequal has measure zero.

4. *The expression of area as an integral*

An important application of the Burkill integral is to simplify and extend the work of W. H. Young and others on the definition of the area of a curved surface. The starting point is an observation on the conditions needed by Young for $\iint J\,du\,dv$ to give a satisfactory expression for the area of the plane set of points bounded by the curve $x = x(u, v)$, $y = y(u, v)$ when (u, v) traces out the boundary of a rectangle R and J is the Jacobian of (x, y) with respect to (u, v). Burkill points out that Young's conditions involve the partial derivatives of x and y, and the separation of x from y and u from v, whereas the natural relationship is between *points* (x, y) and (u, v). This suggests that J is not the most natural tool, and that a modification of it might be used to better effect. The modification which he introduces depends on the notion of the *two-dimensional increment* $\Delta(x, y)$ of (x, y) over the rectangle I in the (u, v) plane which is defined by

$$\Delta(x, y) = \tfrac{1}{2}[x_1 y_2 - x_2 y_1 + x_2 y_3 - x_3 y_2 + x_3 y_4 - x_4 y_3 + x_4 y_1 - x_1 y_4],$$

where the suffices denote corners of I in anti-clockwise direction, and x_1 is the value of x at point 1. In fact, $\Delta(x, y)$ is simply the area of the quadrilateral with vertices (x_1, y_1), (x_2, y_2), (x_3, y_3), (x_4, y_4). It is a function of intervals but is not additive, and this is the point at which Burkill's theory becomes relevant.

The upper and lower modified Jacobians $L^*(x, y)$, $L_*(x, y)$ are defined as the upper and lower derivatives of $\Delta(I)$ at (u, v), and if they are equal we say that the *modified Jacobian* $L(x, y)$ exists and takes their common value. Its basic properties are:

(i) $L(x, c) = L(c, y) = 0$ for any constant c;

(ii) $L(x, y) = -L(y, x)$;

(iii) $L(cx, y) = cL(x, y)$, $L(x+y, z) = L(x, z) + L(y, z)$;

(iv) if $x = x(u)$ is independent of v and $y = y(v)$ independent of u, then

$$\mathrm{L} = \frac{\mathrm{dx}}{\mathrm{du}}\frac{\mathrm{dy}}{\mathrm{dv}};$$

(v) L^*, L_*, and L if it exists, are measurable;

(vi) if the partial derivatives of x and y with respect to u and v are continuous at (u_0, v_0), then $L(u_0, v_0)$ exists and has value $J(u_0, v_0)$.

These results can be used to define the area bounded by the closed plane curve $x = x(u)$, $y = y(u)$, $a \leqslant u \leqslant b$, $x(a) = x(b)$, $y(a) = y(b)$. The range (a, b) is divided

https://doi.org/10.1112/S0024609397003767 Published online by Cambridge University Press

into meshes by points $a = u_0 < u_1 < \ldots < u_m = b$, and the corresponding points $A, P_1, \ldots, P_{m-1}, A$ form an inscribed polygon. The interval function $g(I)$ is defined by $g(I_j) = \frac{1}{2}[x(u_{j-1})\,y(u_j) - x(u_j)\,y(u_{j-1})]$ when I is (u_{j-1}, u_j), and the area inside the curve is defined to be

$$\int_{(a,\,b)} g(I)$$

when this exists. A sufficient, but not necessary, condition for this is that the curve is semi-rectifiable in the sense that $x(u)$, $y(u)$ are both continuous and at least one has bounded variation. If $g(I)$ is a.c. and the curve has an area A, then

$$\int_a^b g'(t)\,dt = A,$$

and conversely, if $\int_a^b g'(t)\,dt$ exists and has value A, then the curve has area A. There is an equally satisfactory result when the curve is defined alternatively by $x = x(u,v)$, $y = y(u,v)$ for points (u,v) on the perimeter of R. It is that if every curve in (x,y) which is the image of a parallel subrectangle of R has a definite area, and $\Delta(I,x,y)$ is a.c. in R and $L(x,y)$ exists a.e. in R, then $A = \iint L\,du\,dv$ over R.

A similar appeal to the theory of interval functions can be used to define the area of a *curved* surface S consisting of points (x,y,z), where $x = x(u,v)$, $y = y(u,v)$, $z = z(u,v)$ are continuous in the rectangle R ($a \leqslant u \leqslant b$, $c \leqslant v \leqslant d$). For this, we define interval functions

$$g_1(I) = \Delta(I,y,z), \quad g_2(I) = \Delta(I,z,x), \quad g_3(I) = \Delta(I,x,y),$$
$$G_1(I) = \int g_1, \quad G_2(I) = \int g_2, \quad G_3(I) = \int g_3,$$

over I, and suppose that G_1, G_2, G_3 exist and are finite for every parallel subrectangle in R. This means that the projections on any coordinate plane of the curve on S which is the image of the perimeter of any subrectangle has a definite finite area. Under these conditions, the area of S is defined as the upper integral over R of

$$g(I) = \{G_1^2(I) + G_2^2(I) + G_3^2(I)\}^{1/2},$$

and the following conclusions are deduced.

(i) The area S is finite if and only if the upper integrals of $|G_1|$, $|G_2|$, $|G_3|$ are finite.

(ii) If S is absolutely continuous and $L(y,z)$, $L(z,x)$, $L(x,y)$ exist a.e. in R, then the area of S is

$$\iint \{L^2(y,z) + L^2(z,x) + L^2(x,y)\}^{1/2}\,du\,dv.$$

5. *Approximate differentiation and extensions of the Perron integral* [**3**, **9**, **10**, **14**, **15**, **17**]

A major field of study after the establishment of the Lebesgue integral lay in the search for integrals with greater scope in the range of integrands on which they could operate and greater facility in applications such as the integration of derivatives. These integrands, unlike those in the Burkill integral described above, were point functions and the integrals, including the more familiar ones associated with the

https://doi.org/10.1112/S0024609397003767 Published online by Cambridge University Press

names of Denjoy and Perron, were defined descriptively as primitives $F(x)$ satisfying the basic condition $F'(x) = f(x)$ in some sense.

The starting point for Burkill was the extension of the concept of differentiation to that of *approximate differentiation*, and paper [**9**], written with Haslam-Jones, extends and simplifies some earlier work of Besicovitch.

The upper right λ-derivative $AD^+(f, x, \lambda)$ is defined as the lower bound of numbers a such that the set of points $\xi > x$ for which $f(\xi) - f(x) \geqslant a(\xi - x)$ has upper right density at x less than or equal to λ. Since $AD^+(f, x, \lambda)$ increases as λ decreases to 0, its limit exists and is called the *upper right approximate derivative*, written $AD^+ f(x)$. The other three right or left approximate derivatives, and the upper and lower (two-sided) approximate derivatives $AD^* f$ and $AD_* f$, are defined similarly. The common value of $AD^+ f$ and $AD_+ f$, when they are equal, is called the *right approximate derivative*, and ADf is *the approximate derivative* when all four are equal.

It is plain that ordinary differentiability implies approximate differentiability, but the stronger result that $D^+ f = ADf = D_- f$ a.e. in a set X in which $D^+ f$ is finite is also valid. Similar extensions can be made to other limit processes, and particularly to *approximate continuity*. These ideas are used in [**3**] to give a particularly direct proof of the fundamental theorem of the calculus for the Denjoy integral which, in its restricted form, is known to be equivalent to Perron's.

The study of the possible disposition of derivatives of measurable functions can be extended [**13**] to cases of non-measurable functions by the introduction of the concept of *relative measurability*, whereby a set X is measurable in relation to X_0 if there is a measurable set M such that $X_0 X = X_0 M$. In the same general field, but not directly related to it, is a paper [**12**] on the differentiability of functions of two variables. This completes the theory of Rademacher and Stepanoff by filling in some gaps in the latter's analysis, and goes on to consider monotonic functions in the plane. In complex notation, $f(z)$ is monotonic increasing if $f(z') \geqslant f(z)$ when $z' \geqslant z$ in the sense that $x' \geqslant x$, $y' \geqslant y$. It is then shown that $\limsup |h|^{-1} |f(z+h) - f(z)|$ is finite a.e. when $f(z)$ is monotonic, and that a similar result holds for a function which, in a certain sense, has bounded variation.

Burkill's important contribution to the problem of extending the scope of the Perron integral was to suggest that approximate rather than ordinary continuity might be a more natural property of the indefinite integral to aim for, and to demonstrate this. He uses the usual formulation of the Perron integral, but extends the concepts of major and minor functions by defining a major function $M(x)$ by the conditions that it is approximately continuous, $M(a) = 0$ and $AD_* M(x) > -\infty$, $AD_* M(x) \geqslant f(x)$ a.e. in (a, b). Minor functions $m(x)$ are defined similarly, and we define K, k respectively as the lower bound of $M(b)$ and the upper bound of $m(b)$ for all major and minor functions. Then $K \geqslant k$ and, if they are equal, we say that $f(x)$ has an *approximately continuous Perron integral* (AP) $\int f(x)\,dx$ equal to their common value. The AP integral is then consistent with the ordinary Perron integral and, *a fortiori*, with Riemann and Lebesgue. Other properties of the AP integral, including the approximate continuity of the integral, are established.

This generalization of the Perron integral depends on the replacement of continuity by approximate continuity, but Burkill introduces in a series of papers [**11**, **14**, **15**, **17**] a generalization in a different direction which leads to what he calls the Cesàro–Perron (CP) integral. The essential idea is to replace $Q(x+h)$ in the increment $Q(x+h) - Q(x)$ of a function Q by the arithmetic (Cesàro) mean $C(Q, x, x+h)$ in the interval $(x, x+h)$, and Q is called C-continuous at x if $C(Q, x, x+h) \to Q(x)$ as $h \to 0$. If $Q(x)$ is finite and C-continuous at every point of an interval, it follows that $Q(x)$ is at every point the derivative of its indefinite integral. Since an everywhere finite derivative can be integrated by the restricted Denjoy process, it is appropriate that this, or rather the equivalent ordinary Perron process, should give the sense in which the integral for mean values is understood. The (two-sided) *upper C-derivative* $CD^*Q(x)$ can now be defined as the upper limit as $h \to 0$ of $2h^{-1}[C(Q, x, x+h) - Q(x)]$. The lower C-derivative $CD_*Q(x)$ is defined similarly and, when $CD^*Q = CD_*Q$, Q has *C-derivative CDQ* equal to their common value. The definition of the CP integral can now be completed by the use of major and minor functions as in the case of the ordinary Perron integral. If $f(x)$ is measurable and finite in $[a, b]$, we call $M(x)$ a major function if it is C-continuous, $M(a) = 0$ and, for $a \leqslant x \leqslant b$, $CD_*M(x) > -\infty$, $CD_*M(x) \geqslant f(x)$. A minor function $m(x)$ is defined similarly, and the *Cesàro–Perron integral* CP $\int f(x)\,dx$ exists and has value K if K and k are defined as before and $K = k$. Burkill goes on to establish the basic properties of the CP integral, including its consistency with the ordinary Perron integral.

Two papers [**14**, **17**] are devoted to a further generalization of the Perron integral to a *scale* of C_rP integrals in which r can be any positive real number. This depends on the replacement of the arithmetic mean $C(Q, x, x+h)$, corresponding to the case $r = 1$, by the Cesàro mean of order r defined by

$$C_r(Q, x, x+h) = rh^{-1} \int_x^{x+h} (x+h-t)^{r-1} Q(t)\,dt.$$

In a further paper [**15**], Burkill shows how the Cesàro summability of the Fourier series of a periodic function $f(x)$ is related to the C_rP integrability of $f(x)$. If $f(x)$ is C_rP integrable and $f(x+t) - f(x-t) \to 2s(C, j)$ as $t \to 0$, then the Fourier series of $f(x)$ is summable (C, k) at x to s if $k > j \geqslant r+1$.

The CP integral also provides [**16**] a powerful and elegant approach to the problem which was known to be insoluble in terms of the ordinary Perron integral. This is to express as a Fourier series of a function $f(x)$ any trigonometric series which converges everywhere or, more generally, has finite upper and lower sums. These results are generalized in later papers [**19**] and [**20**] by extending the scope of the CP integral by introducing the *symmetric CP integral*, in which the continuity condition

$$h^{-1} \int_x^{x+h} F(t)\,dt \to F(x) \quad \text{as } h \to 0$$

is replaced by the weaker symmetrical condition

$$h^{-1} \left\{ \int_x^{x+h} F(t)\,dt - \int_{x-h}^{x} F(t)\,dt \right\} \to 0 \quad \text{as } h \to 0.$$

The results on Fourier series can be extended to Fourier integrals.

Burkill returns later [**23**] to the idea of a scale of integrals $D\alpha$ with $0 \leqslant \alpha \leqslant 1$ which spans the gap between the Lebesgue integral ($\alpha = 1$) and the restricted Denjoy integral ($\alpha = 0$).

6. *Other topics*

In addition to the work already described, Burkill also produced papers on a wide variety of interesting special topics and problems which are not as strongly related to one another as those in the preceding sections, although they depend for the most part on similar analytical techniques.

(*a*) *Inversion formulae* [**5**, **6**]. The first paper shows how pairs of formulae of the type

$$F(x) = \int H(x,t)\, d\Phi(t), \quad \Phi(t) = \int K(t,u)\, dF(u)$$

arise through a discontinuity integral

$$\int H(x,t) \frac{\partial}{\partial t} K(t,u)$$

with values 0 or 1 according as $u < x$ or $u > x$.

The formulae include Fourier and Hankel transforms, and similar ideas can be applied to the Mellin transform.

(*b*) *Differential properties of Young–Stieltjes integrals* [**18**]. L. C. Young has shown that it is possible to define a Stieltjes integral $F = \int f\, d\phi$ in cases where ϕ is not (as is usual) of bounded variation, provided that a suitable additional constraint is put on the variation of f. The paper establishes the formal differential relationship

$$\frac{d}{d\phi} \int f\, d\phi = f$$

in the precise sense that

$$F(x+h) - F(x) = f(x)\{\phi(x+h) - \phi(x) + h\varepsilon(h)\}, \quad \varepsilon(h) \to 0 \text{ as } h \to 0.$$

(*c*) *Strong and weak convergence* [**7**]. The new results in this paper extend the studies by W. H. Young on the concept of super summability defined by the condition $\int Q\{|f(x)|\}\, dx < \infty$ when $Q(u) = \int_0^u q(u)\, dt$ and $q(u)$ is positive and Lebesgue integrable over every finite interval. The cases $q(u) = u^{p-1}$ with $p \geqslant 1$ give the familiar Lp classes (with $p = 1$ indicating ordinary L integrability) and it is shown that well-known results on strong and weak convergence in Lp can be extended to general Q.

(*d*) *Hobson's convergence theorem for Denjoy integrals* [**8**]. This extends the study by Hobson of the behaviour as $n \to \infty$ of integrals of the type $\int f(t)\Phi(t,x,u)\, dt$ in which $f(t)$ is integrable only in the Denjoy sense. This makes it possible to prove, among other things, that the Fourier series of a D integrable function f is summable a.e. to $f(x)$ by Riesz means of any order greater than one.

(*e*) *The differentiability of multiple integrals* [**21**]. The integral of an L integrable function $f(P)$ is *strongly differentiable* at P_0 if

$$(mI)^{-1} \int_I f(P)\, dP$$

https://doi.org/10.1112/S0024609397003767 Published online by Cambridge University Press

tends to a limit as the diameter of the interval I (containing P_0) tends to 0. The main result is a general theorem on measure from which it is possible to deduce the theorem of Jessen, Marcinkiewicz and Zygmund that the integral of $f(P)$ in k dimensions is strongly differentiable a.e. if $|f|[\log^+|f|]^{k-1}$ is L integrable.

(f) *Rearrangements of functions* [**22**]. The functions f, f^* are rearrangements of one another if the measures of sets in which $f(x) \geqslant y$ and $f^*(x) \geqslant y$ are equal, and the main result is an extension from one to two dimensions of a powerful and interesting inequality of Hardy and Littlewood. The function $\theta(f, x, y)$ is defined in a rectangle R as the upper bound for $0 \leqslant x' < x$, $0 \leqslant y' < y$ of

$$(mI)^{-1} \iint_I f(x, y)\, dx\, dy$$

for a rectangle I in R with (x, y) and (x', y') as its north east and south west corners. The main theorem is that

$$\iint \theta(f, x, y)\, dx\, dy$$

is maximum for all rearrangements f of a given function when f decreases from the south west with contours of the form $\log(ax^{-1})\log(by^{-1}) = k$, $0 < k < \infty$.

(g) *An integral for distributions* [**24**, **26**]. The theory of distribution systematized by L. Schwartz has important applications in mathematics and physics, and different approaches to it are possible. Schwartz himself appeals to the general theory of linear functionals, but this can be avoided [**24**] by using more traditional techniques based on Stieltjes integrals extended in an appropriate way. The same analysis is used effectively in dealing with theorems on Fourier and Mellin transforms.

(h) *Polynomial approximation* [**25**]. The paper deals with the following conjecture of H. Burkill. There is a number K_n, depending only on n, such that, given a continuous function $f(x)$, there is a polynomial $p_{n-1}(x)$ of degree at most $n-1$ for which, for all x in a finite interval I,

$$|f(x) - p_{n-1}(x)| \leqslant K_n \sup |\Delta_n(f)|$$

when $\Delta_n(f)$ is the nth difference (in a sense to be defined) of $f(x)$ with respect to $n+1$ points $h_0, h_1, \ldots, h_n$ of I, and the supremum is taken over all such sets of points.

The theorem was proved by Whitney with Δ_n being the usual nth difference with equally spaced h_i. The theorem is proved here with the much better constant $K_n = 2^{-n}$ provided that the h_i are not required to be equally spread and Δ_n is defined appropriately.

(i) *Concavity of discrepancies in inequalities* [**27**]. In the inequality $G \leqslant A$ between geometric and arithmetic means of a set of n numbers, there is a *discrepancy* $\Delta = n(A - G)$, and it is known that Δ is not only non-negative but also super additive in the sense that it is not decreased by the insertion of additional terms in A and G. The paper notes similar results for other inequalities (including Holder, Minkowski and Tchebichoff), and goes on to prove analogues, motivated by discrepancies, of the Hlawka inequality $|x+y+z| - |y+z| - |z+x| - |x+y| \geqslant 0$ for vectors.

ACKNOWLEDGEMENTS. I am greatly indebted to Dr Harry Burkill for his advice and help on all parts of this memoir, and to Professor E. J. Kenney for his appreciation of Charles Burkill's life and work at Peterhouse.

This article appeared in *Biographical Memoirs of Fellows of the Royal Society* 40 (1994) 544–559 and is reprinted by kind permission of the Royal Society. The photograph is the 1968 copyright of David Vicary.

Bibliography

Papers

1. 'Functions of intervals', *Proc. London Math. Soc.* 22 (1923) 275–310.
2. 'The expression of area as an integral', *Proc. London Math. Soc.* 22 (1923) 311–336.
3. 'The fundamental theorem of Denjoy integration', *Proc. Cambridge Philos. Soc.* 21 (1923) 659–663.
4. 'The derivatives of functions of intervals', *Fund. Math.* 5 (1924) 321–327.
5. 'The expression in Stieltjes integrals of the inversion formulae of Fourier and Hankel', *Proc. London Math. Soc.* 25 (1926) 513–524.
6. 'On Mellin's inversion formula', *Proc. Cambridge Philos. Soc.* 23 (1927) 356–360.
7. 'The strong and weak convergence of functions of general type', *Proc. London Math. Soc.* 28 (1928) 493–500.
8. 'On Hobson's convergence theorem for Denjoy integrals', *J. London Math. Soc.* 4 (1929) 127–132.
9. (with U. S. HASLAM-JONES) 'The derivates and approximate derivates of measurable functions', *Proc. London Math. Soc.* 32 (1931) 346–355.
10. 'The approximately continuous Perron integral', *Math. Z.* 34 (1931) 270–278.
11. 'The Cesàro–Perron integral', *Proc. London Math. Soc.* 34 (1932) 314–322.
12. (with U. S. HASLAM-JONES) 'Notes on the differentiability of functions of two variables', *J. London Math. Soc.* 7 (1932) 297–305.
13. (with U. S. HASLAM-JONES) 'Relative measurability and the derivates of non-measurable functions', *Quart. J. Math.* 4 (1933) 233–239.
14. 'The Cesàro–Perron scale of integration', *Proc. London Math. Soc.* 39 (1935) 541–552.
15. 'The Cesàro scales of summation and integration', *J. London Math. Soc.* 10 (1935) 254–259.
16. 'The expression of trigonometrical series in Fourier form', *J. London Math. Soc.* 11 (1936) 43–48.
17. 'Fractional orders of integrability', *J. London Math. Soc.* 11 (1936) 220–226.
18. 'Differential properties of Young–Stieltjes integrals', *J. London Math. Soc.* 23 (1948) 22–28.
19. 'Integrals and trigonometric series', *Proc. London Math. Soc.* 1 (1951) 46–57; Corrigendum, *Proc. London Math. Soc.* 47 (1983) 192.
20. 'Uniqueness theorems for trigonometric series and integrals', *Proc. London Math. Soc.* 1 (1951) 163–169.
21. 'On the differentiability of multiple integrals', *J. London Math. Soc.* 26 (1951) 244–249.
22. 'Rearrangements of functions', *J. London Math. Soc.* 27 (1952) 393–401.
23. (with F. W. GEHRING) 'A scale of integrals from Lebesgue's to Denjoy's', *Quart. J. Math.* 4 (1953) 210–220.
24. 'An integral for distributions', *Proc. Cambridge Philos. Soc.* 53 (1957) 821–824.
25. 'Polynomial approximations to functions with bounded differences', *J. London Math. Soc.* 33 (1958) 157–161.
26. 'Fourier–Stieltjes integrals', *J. Math. Anal.* 43 (1973) 285–292.
27. 'The concavity of discrepancies in inequalities of the means and of Hölder', *J. London Math. Soc.* 7 (1974) 617–626.

Books

Burkill wrote, in addition to his original work, a number of valuable and well-known textbooks, which still rank as standard texts. These all display, as we might expect, not only his mastery of the field but a lucidity and elegance that encourage his readers to appreciate the profound aesthetic quality of good mathematics. The *Lectures on approximation by polynomials* were given during a visit to the Tata Institute, Bombay, in 1959, but have not, unfortunately, found wider circulation.

The Lebesgue integral, Cambridge Tracts in Math. and Math. Phys. 40 (Cambridge University Press, 1951).
The theory of ordinary differential equations (Oliver & Boyd, Edinburgh, 1956; Italian edition 1967).
Lectures on approximation by polynomials (Tata Institute of Fundamental Research, Bombay, 1959).

https://doi.org/10.1112/S0024609397003767 Published online by Cambridge University Press

A first course in mathematical analysis (Cambridge University Press, 1962; Iranian edition 1991).
(with H. BURKILL) *A second course in mathematical analysis* (Cambridge University Press, 1970).
(with H. M. CUNDY) *Mathematical scholarship problems* (Cambridge University Press, 1960).

105 Manor Green Road
Epsom
Surrey KT19 8LW

H. R. PITT