DOI 10.1007/s10958-018-3932-8
Journal of Mathematical Sciences, Vol. 233, No. 1, August, 2018

THE LEIBNIZ DIFFERENTIAL AND THE PERRON–STIELTJES INTEGRAL

E. V. Shchepin

UDC 517.22+517.3+517.518.12+517.518.126

Abstract. We implement Leibniz's idea about the differential as the length of an infinitesimally small elementary interval (a monad) in a form satisfying modern standards of rigor. The concept of sequential differential introduced in this paper is shown to be in good alignment with the standard convention of the integral calculus. As an application of this concept we simplify and generalize the construction of the Perron–Stieltjes integral.

1. Introduction

The very first exposition of mathematical analysis in the famous book by Marquis de L'Hôpital [1] was in essence axiomatic. Among the axioms of the infinitesimal calculus formulated by de L'Hôpital, there is a geometric axiom that asserts that each curved line is a polygonal line with infinitely small rectilinear segments. This axiom fits well with Cavalieri's method of indivisibles (which appeared before the emergence of the differential calculus), who perceived that a rectilinear interval consists not of points, but rather of infinitely small indivisible intervals. These perceptions brought to life the key concept of calculus—the concept of the Leibniz differential and the Leibniz notation for the integral and the derivative (which are based on this concept). Leibniz's wonderful notation was never abandoned since, despite the revision of the entire foundation of the calculus. It was simply given a new sense, which is sometimes quite far from the original meaning. The key concept in Leibniz's infinitesimal calculus (as distinct from Newton's calculus) is the differential, through which both the integral and the derivative are expressed. The Leibniz differential of an independent variable is an infinitely small constant. However, in the modern development of the subject, the differential is a function (a differential form). Differential forms are known to be contravariant (when changing an integration variable), whereas both the differential in Leibniz's sense and the differential put forward in the present paper are covariant.

Leibniz assumed that the points of the real line that represent a variable quantity x are indeed infinitely small intervals of fixed infinitely small length—the so-called monads of an interval. The length of a monad, denoted by dx, was called the differential of a variable x. The differential of a function $f(x)$ with a fixed x meant the variation of a function on the corresponding infinitely small interval $df(x)$.

We put all these intuitive perceptions of Leibniz in a rigorous form in the sense of Cauchy, who interpreted an infinitely small quantity as a sequence tending to zero. It should be noted that the "nonstandard analysis," which proposed a rigorous conception of infinitely small quantities, did not give an adequate interpretation of the Leibniz differential. In the nonstandard analysis, a monad of a standard point, which is defined as the set of infinitely close points, has no length. The concept of the Leibniz differential naturally leads us to the construction of an integral in the sense of Perron. In the present paper, this concept of an integral is given in the form of the integral of infinitesimal distributions. Note that distributions of this kind are also capable of representing generalized functions. This being so, the emerging concept of the integral is to some extent superior to all the existing concepts.

Under this approach, the differential forms are replaced by the infinitesimal forms, which are functions associating number sequences with infinitely small intervals. The integration of Stieltjes-type infinitesimal forms leads exactly to the well-known concept of the Perron–Stieltjes integral [6].

Translated from Fundamentalnaya i Prikladnaya Matematika, Vol. 20, No. 6, pp. 237–258, 2015.

The after-Leibnizian history of the integral, which started with the Riemann integral and ended with the Kurzweil–Henstock integral, eventually led to the concept of the differential [3, 4]—however, this concept was overcomplicated and defined on the basis of the integral. The concept of the Leibniz differential, which we put forward below, rehabilitates the priority of the differential with respect to the integral.

The fairly general concept of the Leibniz integral was successfully implemented by Kolmogorov [2] in a form meeting all the modern requirements of rigor; however, he did not introduce the concept of the differential, even though he used the differential in his notation.

In the present paper, the author's aim was not only to further generalize the concept of an integral, however being concerned only with the integration over intervals on the real line, but rather precisely reproduce Leibniz's intuitive perceptions, which he followed in introducing his remarkable notation. The most important contribution of Leibniz to mathematics was the introduction of the notation $\int f(x)\,dx$ for the integral. At first sight, the introduction of dx seems superfluous—indeed, all the information required for integral evaluation is already contained in $f(x)$. However, the presence of a differential in the integrand allows one to correctly transform the integral when changing the integration variable. The proposed concept of the differential fits perfectly with the well-established system of notation, giving them their original clear and transparent sense, which perfectly matches the modern generally accepted sense.

2. The Structure of the Continuum

What is the continuum made of? Points, infinitely small intervals? A question of this sort was of special relevance for the 18th century mathematicians. An inexperienced reader might think that there is really no difference. This is nothing but philosophy. However, the difference between a point and an interval is as follows: all sizes of a point are zero, whereas infinitely small quantities may well have finite relations. For example, a two-fold increase in all sizes corresponds to doubling the length of an infinitely small interval, while a point will be transformed into the same sizeless point. This is why Leibniz and Euler's intuitive perceptions about the continuum are better matched to the position expressed by the de L'Hôpital's geometric axiom. Later Dedekind put forward an elegant theory of real numbers, which put an end to the historic controversy. It looked like the scales had finally tipped towards points. The concept of the continuum, which we introduce below, rehabilitates the concept of infinitely small intervals, without affecting the entitlement of points. In other words, our answer to the "points-or-intervals" alternative is as follows: both points and intervals.

Infinitesimal Intervals. Curly brackets will be used to denote the sets consisting of the elements inside the brackets. By an *infinitely small* or *infinitesimal* interval we shall mean a sequence of nested intervals $\{[a_k, b_k]\}$[1] whose lengths tend to zero. The length of an infinitely small interval is the sequence of the lengths of these intervals $\{b_k - a_k\}$. So, the length of an infinitely small interval is a sequence that monotonically tends to zero.

Infinitesimal Partitions. By square brackets $[x_0, x_1, \dots, x_n]$ we shall denote the set of intervals of the real line, whose end-points lie in the set $\{x_0, x_1, \dots, x_n\}$ and whose interiors do not contain points of this set. A set of intervals $[x_0, x_1, \dots, x_n]$ is called a *partition* of an interval $[a, b]$ generated by a set $\{x_0, x_1, \dots, x_n\}$ if a and b are, respectively, the smallest and largest elements of this set. If the points of a set $\{x_0, x_1, \dots, x_n\}$ are arranged in an increasing order ($x_i \le x_{i+1}$), then $x_0 = a$, $x_n = b$, and so $[x_0, x_1, \dots, x_n]$ is the set of intervals $\{[x_i, x_{i+1}]\}$.

If a set of points is denoted by one letter, say P, then by $[P]$ we will denote the partition of an interval generated by this set. If one set of points is contained in another one, $P_1 \subset P_2$, and if both P_1 and P_2 generate partitions $[P_1]$ and $[P_2]$ of the same interval, then the second partition is called a *refinement* of

[1] It would be more correct to write $\{[a_k, b_k]\}_{k=1}^{\infty}$ or at least $\{[a_k, b_k]\}_k$, but our convention is that the presence of a subscript k without more specific information in an expression in curly brackets means that the range of its variation is $\mathbb{N}$, the set of positive integers.

the first one. A sequence of partitions $[P_1], \ldots, [P_n], \ldots$ of increasing refinement of an interval $[a, b]$ is called a *refinement* of this interval. A nested sequence of intervals $\{I_k\}$ is called a *monad* of a refinement $[P_1], \ldots, [P_n], \ldots$ if $I_k \in [S_k]$ for any k. A refinement of an interval is called *infinitesimal* if all its monads are infinitely small. For a given interval, nonterminal points from the union $\bigcup_{k=1}^{\infty} P_k$ are called *cutting points* for a refinement $[P_1], \ldots, [P_n], \ldots$. For each cutting point x its *order* is defined as the minimum of k for which $x \in P_k$.

Together with partitions of intervals, we can also consider partitions of the entire line or noncompact intervals. Moreover, one may also consider partitions with infinite number of points of a fixed order. For example, we let P_0 agree with the set of integers $\mathbb{Z}$ and define P_k as the set decimal fractions of the form $n/10^k$, where $n \in \mathbb{Z}$. In this case, the sequence $\{[P_k]\}$ forms an infinitesimal partition of the real line, which is called the *decimal* partition.

Points and Monads. For given infinitesimal partition on an interval, boldface Latin lower case letters will be used to denote monads of this partition; the corresponding normal-weight letters will denote the points. Under this convention, the same letter will be used to denote a point and the monad that has this point as a limit. For example, for the monad $\mathbf{x}$ we denote by x its limit point. In general, a monad $\mathbf{x}$ is not uniquely determined by a point x. However, the uniqueness here is secured if a limit point x is internal for all intervals representing the monad of a sequence. If a limit point of a monad is a cutting point of a partition, then there are exactly two monads of the partition whose limit is this point: the left and right ones.

For an infinitesimal partition $[P_1], [P_2], \ldots$ of an interval $[a, b]$ and a point $x \in [a, b]$, we let $x + 0$ denote the sequence in which the kth term is the point from P_k that is nearest to x, distinct from x and lying to the right of x. We denote by $x - 0$ the sequence in which the kth term is the point from P_k that is nearest to x, distinct from this point, and lying to the left of x.

This convention provides the means for denoting various infinitesimal intervals that have x as their limit point. Namely, given $x - 0 = \{a_k\}$ and $x + 0 = \{b_k\}$, we have three infinitesimal intervals $\{[a_k, x]\}$, $\{[x, b_k]\}$, and $\{[a_k, b_k]\}$, which will be written for brevity as $[x - 0, x]$, $[x, x + 0]$, and $[x - 0, x + 0]$, respectively.

If a point x is a cutting point of a partition, then $[x - 0, x]$ and $[x, x + 0]$ agree, starting from some number (of order x), with the left and right monads of the partition that have x as their limit point. The corresponding monads, which are called *one-sided*, will be denoted, respectively, as $\mathbf{x} \pm 0$. If x is neither a cutting point of a partition nor a terminal point for the interval, then $[x - 0, x + 0]$ is the unique monad of the partition that has x as its limit point. Such a monad is called *two-sided*. In this case, the infinitesimal intervals $[x - 0, x]$ and $[x, x + 0]$ will be called *half-monads*.

The monads of the decimal partition of the unit interval $[0, 1]$ on the real line are in a one-to-one correspondence with the infinite decimal fractions; the points with nonunique decimal notation as exactly the limit points for pairs of one-sided monads.

3. Sequential Distributions

Monadic Distributions. By a *monadic sequential distribution* we shall mean a mapping assigning with any monad of some refinement a number[1] sequence.

A most important example of a monadic distribution is furnished by the *sequential differential of length*, which associates with any monad $\mathbf{x}$ its length—the sequence (written $d\mathbf{x}$) of lengths of the intervals that it contains.

The *sequential*[2] *differential of a function $f(x)$* on a monad $\mathbf{x} = \{[a_k, b_k]\}$ of some refinement (written $df(\mathbf{x})$) is defined as the sequence of differences $\{f(b_k) - f(a_k)\}$. However, when speaking of the differential

[1] In the present paper, we consider only real numbers.

[2] In what follows we shall drop the adjective "sequential," because in the present paper there will be no other distributions and differentials.

of a function at a point, there is an ambiguity due to an ambiguous correspondence between points and monads. This, in turn, gives us the notion of the right and left differentials at cutting points of a refinement.

Any number function $f(x)$ on a refined interval generates in a natural way a monadic distributions by associating with a monad $\mathbf{x}$ a constant sequence of values of a function at the limit point of the monad. In other words, the distribution generated by the function $f(x)$ can be described by the formula $f(\mathbf{x})_k = f(x)$ for all k.

However, for discontinuous functions at ordinary discontinuities (if such points are cutting points of a refinement) it is more natural to define the value of a function at the left and right monads of a point of discontinuity as, respectively, its limits on the left and right; that is,

$$f(\mathbf{x}+0) = \lim_{t\to x+0} f(t) \quad\text{and}\quad f(\mathbf{x}-0) = \lim_{t\to x-0} f(t).$$

For example, for infinite decimal fractions, which represent monads of a decimal refinement of the real line, it is reasonable to define the integer part as the number appearing before the decimal point. (For instance, we assume in this way that the integer part of the monad $0.9999\ldots$ is zero.)

A monadic distribution is also generated by any sequence of functions $\{f_k(x)\}$. For example, with the monad $\mathbf{x}$ one may associate the sequence $\{f_k(x)\}$. This being so, since sequences of functions are capable of representing generalized functions, the language of monadic distributions is applicable to the description of calculations involving generalized functions.

Another way of generation of a monadic distribution for the conventional and generalized function $\alpha(x)$ is as follows: to a monad $\mathbf{x} = \{[a_k, b_k]\}$ one assigns a sequence of the averaged values of $\alpha(x)$ on the monad intervals, that is, by the formula

$$\alpha(\mathbf{x}) = \left\{ \frac{1}{b_k - a_k} \int_{a_k}^{b_k} \alpha(x)\, dx \right\}. \tag{1}$$

Arithmetic operations with monadic distributions are determined termwise for the corresponding sequences.

So, for any pair of functions $f(x)$, $g(x)$ on a refined interval, the monadic distributions $f(\mathbf{x})\,dg(\mathbf{x})$ and $g(\mathbf{x})\,df(\mathbf{x})$ are defined.

Function Linearity of the Definite Differential. Consider an interval $[\mathbf{a}, \mathbf{b}]$ with fixed infinitesimal partition. It is easily seen that, for any pair of functions $f(x)$, $g(x)$ on $[a, b]$ and any constants λ, μ, for any $\mathbf{x} \in [\mathbf{a}, \mathbf{b}]$ the differential of the linear combination of these functions is equal to the linear combination of their differentials (the linearity of the differential)

$$d\big(\lambda f(\mathbf{x}) + \mu g(\mathbf{x})\big) = \lambda\, df(\mathbf{x}) + \mu\, dg(\mathbf{x}). \tag{2}$$

Comparison of Differentials. A sequence having only a finite number of nonpositive terms is called *eventually nonnegative.* A monadic distribution is called *eventually nonnegative* if so are all its values.

Lemma 3.1. *If differential of a function $f(x)$ on some refinement of an interval $[a, b]$ is eventually nonnegative, then $f(b) \geq f(a)$.*

Proof. Assume on the contrary that $f(b) < f(a)$. Consider the partition of the interval $[a, b]$ from the given refinement. Let $a = x_0 < x_1 < \cdots < x_n = b$ be the sequence of end-points of the intervals from this partition. We have $f(x_n) < f(x_0)$, and hence, the inequalities $f(x_{k+1}) \geq f(x_k)$ cannot hold for all $k < n$. Therefore, there exists an interval $[a_1, b_1]$ of the first partition, for which the value of the function at the right end-point is smaller than the value at the left end-point. The second partition of the interval refines the first one, and hence, a similar argument shows that inside the interval $[a_1, b_1]$ there is an interval $[a_2, b_2]$ of the second partition, for which the value at the right end-point is smaller than the value at the left end-point. Continuing in this way, we define the monad $\{[a_k, b_k]\}$ with negative sequence of differences $\{f(b_k) - f(a_k)\}$, which however contradicts the assumption about eventual nonnegativity of the differential. $\square$

A number sequence $\{x_k\}$ will be said to *eventually majorize* another sequence $\{y_k\}$ (written $\{x_k\} \succ \{y_k\}$) if the inequality $x_k > y_k$ holds for all sufficiently large k. We say that one monadic distribution eventually majorizes another one if the value of the first one on any monad eventually majorizes the value of the second distribution on the same monad.

Theorem 3.1 (comparison)**.** *If* $df(\mathbf{x}) \succ dg(\mathbf{x})$ *for all monads of a given infinitesimal partition on an interval* $[a, b]$*, then* $f(b) - f(a) \geq g(b) - g(a)$

Proof. Indeed, the differential of the difference $d\big(f(x) - g(x)\big)$ in this case is nonnegative. Hence, $f(b) - g(b) \geq f(a) - g(a)$ by Lemma 3.1. $\square$

4. Integration of Distributions

The Upper Integral. The set of monads of an interval $[a, b]$ with fixed infinitesimal partition will be denoted by $[\mathbf{a}, \mathbf{b}]$. The upper integral

$$^*\!\!\int_{\mathbf{a}}^{\mathbf{b}} \Phi(\mathbf{x})$$

of a distribution $\Phi(\mathbf{x})$, $\mathbf{x} \in [\mathbf{a}, \mathbf{b}]$, is defined as the infimum of the differences $f(b) - f(a)$ of functions, whose differentials eventually majorize $\Phi(\mathbf{x})$. From Theorem 3.1 it follows that, for any function $f(x)$ on an interval $[a, b]$,

$$^*\!\!\int_{\mathbf{a}}^{\mathbf{b}} df(\mathbf{x}) = f(b) - f(a). \tag{3}$$

Monotonicity of the Integral. As a direct consequence of the definition of the upper integral, we have the following rule for integration of inequalities:

$$\text{if } \ \Phi(\mathbf{x}) \prec \Psi(\mathbf{x}) \ \text{ for all } \ \mathbf{x} \in [\mathbf{a}, \mathbf{b}], \ \text{ then } \ ^*\!\!\int_{\mathbf{a}}^{\mathbf{b}} \Phi(\mathbf{x}) \leq {}^*\!\!\int_{\mathbf{a}}^{\mathbf{b}} \Psi(\mathbf{x}). \tag{4}$$

Sublinearity of the Upper Integral. Let Φ_1 and Φ_2 be two infinitesimal distributions on the same infinitesimal partition of an interval $[\mathbf{a}, \mathbf{b}]$. Then

$$^*\!\!\int_{\mathbf{a}}^{\mathbf{b}} \big(\Phi_1(\mathbf{x}) + \Phi_2(\mathbf{x})\big) \leq {}^*\!\!\int_{\mathbf{a}}^{\mathbf{b}} \Phi_1(\mathbf{x}) + {}^*\!\!\int_{\mathbf{a}}^{\mathbf{b}} \Phi_2(\mathbf{x}). \tag{5}$$

Homogeneity of the Upper Integral. Let λ be an arbitrary positive constant and $\Phi(\mathbf{x})$, $\mathbf{x} \in [\mathbf{a}, \mathbf{b}]$, be any distribution. Then

$$^*\!\!\int_{\mathbf{a}}^{\mathbf{b}} \lambda\Phi(\mathbf{x}) = \lambda {}^*\!\!\int_{\mathbf{a}}^{\mathbf{b}} \Phi(\mathbf{x}). \tag{6}$$

The Lower Integral. The *lower integral* of a distribution can be defined from the following equality:

$$ {}_*\int_{\mathbf{a}}^{\mathbf{b}} \Phi(\mathbf{x}) = -{}^*\int_{\mathbf{a}}^{\mathbf{b}} -\Phi(\mathbf{x}). \tag{7}$$

The properties of the lower integral are similar to those of the upper integral. In particular, a positive factor can be taken out before the lower integral sign, and moreover, the union theorem for integrals also holds for the lower integral. However, in contrast to the upper integral, the lower integral of the sum of distributions satisfies the opposite inequality:

$$ {}_*\int_{\mathbf{a}}^{\mathbf{b}} \big(\Phi_1(\mathbf{x}) + \Phi_2(\mathbf{x})\big) \geq {}_*\int_{\mathbf{a}}^{\mathbf{b}} \Phi_1(\mathbf{x}) + {}_*\int_{\mathbf{a}}^{\mathbf{b}} \Phi_2(\mathbf{x}). \tag{8}$$

Interval Additivity. If $\{[P_k]\}$ ($\{[Q_k]\}$) is an infinitesimal partition of an interval $[a,b]$ (respectively, $[b,c]$), then, clearly, $\{[P_k \cup Q_k]\}$ is an infinitesimal partition of the interval $[a,c]$, the monads of the partitions being united:

$$[\mathbf{a},\mathbf{b}] \cup [\mathbf{b},\mathbf{c}] = [\mathbf{a},\mathbf{c}]. \tag{9}$$

If (9), then, for any distribution $\Phi(\mathbf{x})$ on $[\mathbf{a},\mathbf{c}]$,

$$ {}^*\int_{\mathbf{a}}^{\mathbf{b}} \Phi(\mathbf{x}) + {}^*\int_{\mathbf{b}}^{\mathbf{c}} \Phi(\mathbf{x}) = {}^*\int_{\mathbf{a}}^{\mathbf{c}} \Phi(\mathbf{x}), \quad {}_*\int_{\mathbf{a}}^{\mathbf{b}} \Phi(\mathbf{x}) + {}_*\int_{\mathbf{b}}^{\mathbf{c}} \Phi(\mathbf{x}) = {}_*\int_{\mathbf{a}}^{\mathbf{c}} \Phi(\mathbf{x}). \tag{10}$$

The Exact Integral. It is easily seen that, for any distribution,

$$ {}_*\int_{\mathbf{a}}^{\mathbf{b}} \Phi(\mathbf{x}) \leq {}^*\int_{\mathbf{a}}^{\mathbf{b}} \Phi(\mathbf{x}). \tag{11}$$

If the upper and lower integrals of a distributions are equal, then the distribution is called *integrable*, and the common value of the upper and lower integrals is called the (*exact*) *integral* of this distribution. The exact integral exists for the differential of any function; it is equal to the difference of the values of the function at the interval terminals.

Hereditarity of Integrability. If (9) holds, then for any distribution $\Phi(\mathbf{x})$ we have that

$$ \int_{\mathbf{a}}^{\mathbf{b}} \Phi(\mathbf{x}) + \int_{\mathbf{b}}^{\mathbf{c}} \Phi(\mathbf{x}) = \int_{\mathbf{a}}^{\mathbf{c}} \Phi(\mathbf{x}) \tag{12}$$

by the interval additivity of one-sided integrals. Moreover, this equality holds *unconditionally*, that is, the existence of the left-hand side implies the existence of the right-hand side, and vice versa. In particular, this is why the integrability property is *hereditary* in subintervals—the integrability on some interval implies that on any subinterval of this interval.

Linearity of the Integral. If distributions $\Phi(\mathbf{x})$ and $\Psi(\mathbf{x})$ are integrable on $[\mathbf{a},\mathbf{b}]$, then, for any constants λ and μ, so is their linear combination $\lambda\Phi(\mathbf{x}) + \mu\Psi(\mathbf{x})$, and moreover,

$$ \int_{\mathbf{a}}^{\mathbf{b}} \big(\lambda\Phi(\mathbf{x}) + \mu\Psi(\mathbf{x})\big) = \lambda \int_{\mathbf{a}}^{\mathbf{b}} \Phi(\mathbf{x}) + \mu \int_{\mathbf{a}}^{\mathbf{b}} \Psi(\mathbf{x}).$$

5. Integration of Infinitesimal Forms

The Indefinite Differential of a Function. The differential of a function, in the way it was treated above, assumed definite values on monads of a given infinitesimal partition of an interval. Moreover, the formula in terms of which the differential of a given monad was defined was independent of a partition. This formula defines a function that associates with any infinitesimal interval $\{[a_k, b_k]\}$ the sequence of differences $\{f(b_k) - f(a_k)\}$. This function of an infinitesimal interval will be called the *indefinite* (*sequential*) *differential of a function* $f(x)$. The differential of a function on an infinitesimal interval $\{[a_k, b_k]\}$ with a limit point x is equal to the sum of its values on the left $\{[a_k, x]\}$ and right $\{[x, b_k]\}$ halves of this interval. Hence, we may restrict the domain of a differential by considering only one-sided infinitesimal intervals. The good point of the one-sided intervals is that they are uniquely characterized by their limit point and length. Thus, the differential of a function $f(x)$ can be defined as a function of two variables: a point x and a sequence of increments dx, the latter having the values

$$f(x + dx) - f(x) \quad \text{or} \quad f(x) - f(x - dx)$$

depending on whether x is the left- or the right-end of this interval. If dx is allowed to assume negative values, then the formula for the differential can be put in a unified way as follows:

$$df(x, dx) = \operatorname{sgn}(dx)\big(f(x + dx) - f(x)\big). \tag{13}$$

However, in what follows, we shall write for brevity $df(x)$ to denote the differential of a function without explicitly indicating its dependence on dx. So, for a fixed x, the differential $df(x)$ is a function of an infinitesimal sequence of increments. In particular, the differential dx of an identity function x is an x-independent function of an increment.

Infinitesimal Forms. By an *infinitesimal form* at a point x we shall mean a function associating with any infinitesimal interval for which x is a boundary point some number sequence. The product $f(x)\,dg(x)$ of the differential of one function of real variable $g(x)$ by another $f(x)$ serves as a basic example of an infinitesimal form. Here, the value of a function $f(x)$ on an infinitesimal interval is defined as its value at the limit point of this interval. Such forms will be also called *Stieltjes* forms.

Infinitesimal forms can be added, multiplied, and divided, and hence any arithmetic expression involving functions and their differentials is an infinitesimal form. Moreover, if an infinitesimal form $D_1(x, dx)$ assumes infinitely small values, then it can be substituted instead of dx in another infinitesimal form $D_2(x, dx)$, thereby forming the superposition $D_2\big(x, D_1(x, dx)\big)$, which is also an infinitesimal form.

Integration of an Infinitesimal Form. For any infinitesimal partition of an interval $[\mathbf{a}, \mathbf{b}]$, the infinitesimal form $D(x, dx)$ generates the monadic distribution $D(\mathbf{x})$ by the formula

$$D(\mathbf{x}) = D(x, d[x - 0, x]) + D(x, d[x, x + 0]), \tag{14}$$

where $d[x - 0, x]$ and $d[x, x + 0]$ denote the lengths of the corresponding half-monads.

In particular, the indefinite differential $df(x)$ of a function generates the definite differential—this is the monadic distribution $df(\mathbf{x})$, whose integral equals the difference of the values of the function $f(x)$ at the end-points of the interval for any infinitesimal partition of this interval. The following definition is motivated by this remark.

An infinitesimal form $D(x, dx)$ is called *Leibniz integrable* on an interval $[a, b]$ if, for any infinitesimal partition of $[a, b]$, the monadic distribution $D(\mathbf{x})$ generated by it is integrable and all the integrals obtained for various partitions are equal. The common value of the integrals of these distributions is denoted by

$$\int_a^b D(x, dx)$$

and is called the *definite integral* of an infinitesimal form $D(x, dx)$ over the interval $[a, b]$.

The *upper integral* of an infinitesimal form $D(x, dx)$ over an interval $[a, b]$ is, by definition, the supremum of the upper integrals over all monadic distributions generated by it. Similarly the *lower integral* of an infinitesimal form $D(x, dx)$ over an interval $[a, b]$ is the infimum of the lower integrals over all monadic distributions generated by it. Clearly, this definition preserves all the conventional relations between one-sided integrals.

Comparison of Infinitesimal Forms. An infinitesimal form $D_1(x, dx)$ will be said to *eventually majorize* an infinitesimal form $D_2(x, dx)$ at a point x (written $D_1(x, dx) \succ D_2(x, dx)$) if, for any infinitely small sequence dx, we have the eventual majorization $D_1(x, dx) \succ D_2(x, dx)$.

The majorization relation of forms $D_1(x, dx) \succ D_2(x, dx)$, when satisfied for all $x \in [a, b]$, implies the majorization relation for the monadic distributions $D_1(\mathbf{x}) \succ D_2(\mathbf{x})$ generated by them. In turn, the latter majorization, as is seen directly from the definition of the integral of a distribution, implies the corresponding inequality for the upper, lower, and exact integrals over an interval on which $D_1(x, dx) \succ D_2(x, dx)$. Hence, we have

$$\int_a^b D_1(x, dx) \geq \int_a^b D_2(x, dx).$$

Linearity and Additivity. The linearity and interval additivity of the integral for infinitesimal forms are similar to the corresponding properties of the integrals of distributions and directly follow from these properties.

6. Transformation of the Integral

Integration by Parts. The following identity is easily verified for any functions:

$$f(x)\, dg(x) = g(x)\, df(x) + d\big(f(x)g(x)\big) + df(x)\, dg(x). \tag{15}$$

As an immediate consequence of this identity, we have the following theorem for the Leibniz integral.

Theorem 6.1. *Let*

$$\int_a^b |df(x)\, dg(x)| = 0.$$

Then the Stieltjes form $f(x)\, dg(x)$ is integrable on an interval $[a, b]$ if and only if the form $g(x)\, df(x)$ is integrable and

$$\int_a^b f(x)\, dg(x) = f(b)g(b) - f(a)g(a) - \int_a^b g(x)\, df(x). \tag{16}$$

For the classical Riemann–Stieltjes integral, the corresponding theorem has the same statement, but without the condition on the integral of the product of differentials. The reason is that the mere existence of the Riemann–Stieltjes integral

$$\int_a^b f(x)\, dg(x)$$

already implies that the Leibniz integral

$$\int_a^b df(x)\, dg(x)$$

is zero—this can be readily checked by representing an integral as the limit of integral sums, as will be done in the last section of the paper. Assume that $f(x)$ is continuous (that is, if $df(x, dx)$ is infinitely small with infinitely small dx, $x \in [a,b]$) and that $g(x)$ has finite variation (that is, $\int_a^b |dg(x)| < \infty$)—these are the standard conditions implying the existence of the Riemann–Stieltjes integral. Then the equality

$$\int_a^b |df(x)\, dg(x)| = 0 \tag{17}$$

follows, first, on integrating the eventual inequality

$$|df(x)\, dg(x)| \prec \varepsilon |dg(x)|,$$

which holds for any $\varepsilon > 0$, and second, by making $\varepsilon \to 0$.

Change of the Integration Variable.

Theorem 6.2. *Let $x(t)$ be an increasing continuous mapping, $x(t_0) = x_0$ and $x(t_1) = x_1$. Then*

$$^*\!\!\int_{x_0}^{x_1} D(x, dx) = {}^*\!\!\int_{t_0}^{t_1} D\big(x(t), dx(t, dt)\big).$$

Proof. We consider an infinitesimal partition of the interval $[\mathbf{t}_0, \mathbf{t}_1]$. Its image under $x(t)$ is an infinitesimal partition of the image interval $[\mathbf{x}_0, \mathbf{x}_1]$.

Given $\varepsilon > 0$, we fix a function $g(x)$ such that, for all $\mathbf{x} \in [\mathbf{x}_0, \mathbf{x}_1]$,

$$D(\mathbf{x}) \prec dg(\mathbf{x}) \quad \text{and} \quad {}^*\!\!\int_{\mathbf{x}_0}^{\mathbf{x}_1} D(\mathbf{x}) + \varepsilon > g(x_1) - g(x_0).$$

Hence, for all $\mathbf{t} \in [\mathbf{t}_1, \mathbf{t}_2]$,

$$D\big(x(\mathbf{t}), dx(\mathbf{t}, d\mathbf{t})\big) \prec dg\big(x(\mathbf{t})\big) \quad \text{and} \quad {}^*\!\!\int_{\mathbf{t}_0}^{\mathbf{t}_1} D\big(x(\mathbf{t})\big) + \varepsilon > g\big(x(t_1)\big) - g\big(x(t_0)\big).$$

Since ε is arbitrary and since $g\big(x(t_1)\big) - g\big(x(t_0)\big) = g(x_1) - g(x_0)$, it follows that the integrals agree:

$$^*\!\!\int_{\mathbf{x}_0}^{\mathbf{x}_1} D(\mathbf{x}) = {}^*\!\!\int_{\mathbf{t}_0}^{\mathbf{t}_1} D\big(x(\mathbf{t})\big).$$

Therefore, the sets of numbers representing the upper integrals of the above forms over various infinitesimal partitions of the corresponding intervals are equal. Hence, their suprema are also equal. In other words, the integral

$$^*\!\!\int_{t_0}^{t_1} D\big(x(t), dx(t, dt)\big)$$

is equal to the integral

$$^*\!\!\int_{x_0}^{x_1} D(x, dx). \qquad \square$$

Replacement of the Differential.

Theorem 6.3. *If*

$$ {}^*\!\!\int_a^b |g(x)|\,dx < \infty, $$

then, for any differentiable function $f(x)$, the Leibniz integrals are equal:

$$ {}^*\!\!\int_a^b g(x)f'(x)\,dx = {}^*\!\!\int_a^b g(x)\,df(x). $$

Proof. Given a positive ε, we have the following inequalities for the infinitesimal forms:

$$ -\varepsilon\,dx \prec f'(x)\,dx - df(x) \prec \varepsilon\,dx. \tag{18} $$

Hence,

$$ -\varepsilon|g(x)|\,dx \prec g(x)f'(x)\,dx - g(x)\,df(x) \prec \varepsilon|g(x)|\,dx. \tag{19} $$

As a result, we have

$$ g(x)f'(x)\,dx \prec g(x)\,df(x) + \varepsilon|g(x)|\,dx. $$

Integrating this inequality, we obtain

$$ {}^*\!\!\int_a^b g(x)f'(x)\,dx \le {}^*\!\!\int_a^b g(x)\,df(x) + \varepsilon\,{}^*\!\!\int_a^b |g(x)|\,dx. $$

Making $\varepsilon \to 0$ in this inequality, we find that

$$ {}^*\!\!\int_a^b g(x)f'(x)\,dx \le {}^*\!\!\int_a^b g(x)\,df(x). \tag{20} $$

On the other hand, using (19), one may get the opposite inequality:

$$ g(x)\,df(x) \prec g(x)f'(x)\,dx + \varepsilon|g(x)|\,dx. $$

Integrating this inequality and making $\varepsilon \to 0$, we get the reverse inequality to (20). □

By using in combination the theorem on the change-of-variable in an infinitesimal form and the change-of-differential theorem, we arrive at the following formula for the change of variable in an integral of a function:

$$ {}^*\!\!\int_a^b f(x)\,dx = {}^*\!\!\int_\alpha^\beta f\big(x(t)\big)x'(t)\,dt \tag{21} $$

in the case where $x(t)$ is an increasing mapping of an interval $[\alpha,\beta]$ onto $[a,b]$ and

$$ {}^*\!\!\int_a^b |f(x)|\,dx < \infty. $$

In particular, if $f(x)$ is identically 1, then we get the fundamental theorem of calculus (the Newton–Leibniz formula).

Corollary 6.4. *If a function $f(x)$ is differentiable on an interval $[a,b]$, then the Stieltjes form $f'(x)\,dx$ is Leibniz integrable on $[a,b]$ and*

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

7. Integration of Difference Forms

Difference Forms. Of primary interest for integration are infinitesimal forms generated by difference forms. Namely, by a difference form we shall mean a real function $D(x, \Delta x)$ of two real variables, of which the first one is known as a (*base*) *point*, while the second one is called an *increment*. For example, by $\Delta f(x)$ we denote the (indefinite) difference of a function $f(x)$, which is defined as

$$\Delta f(x) = f(x + \Delta x) - f(x).$$

Given a pair of functions $f(x)$, $g(x)$, the Stieltjes difference form $f(x)\Delta g(x)$ is defined. Any difference form $D(x, \Delta x)$ generates the infinitesimal form $D^*(x, dx)$ by the formula

$$D^*(x, dx) = \{D(x, dx_k)\}_{k=1}^{\infty}, \tag{22}$$

where $dx = \{dx_k\}$. So, the infinitesimal Stieltjes form $f(x)\,dg(x)$ is generated by the corresponding difference form.

Lemma 7.1. *If infinitesimal forms $D_1^*(x, dx)$ and $D_2^*(x, dx)$ are generated by difference forms $D_1(x, \Delta x)$ and $D_2(x, \Delta x)$, respectively, and if $D_1^*(x, dx) \prec D_2^*(x, dx)$ at a point x, then there exists $\varepsilon > 0$ such that*

$$D_1(x, \delta) \le D_2(x, \delta)$$

for any δ such that $|\delta| \le \varepsilon$.

Proof. Assuming the contrary, we get a sequence $\{\delta_k\}$ tending to zero for which the reverse inequality holds. From this sequence one may take a monotone decreasing subsequence on which the condition of eventual domination of the infinitesimal forms would be violated. □

A difference form $D_1(x, \Delta x)$ will be said to *infinitesimally majorize* another difference form $D_2(x, \Delta x)$ at a point x, which we write as

$$D_1(x, \Delta x) \succ D_2(x, \Delta x),$$

if there exists $\varepsilon > 0$ such that, for all $\delta \in [-\varepsilon, \varepsilon]$,

$$D_1(x, \delta) \ge D_2(x, \delta). \tag{23}$$

By Lemma 7.1, the infinitesimal majorization for difference forms is equivalent to the eventual majorization of the corresponding infinitesimal forms.

The Perron–Stieltjes Integral of a Difference Form. We define the upper Perron–Stieltjes integral of a difference form $D(x, \Delta x)$ as follows:

$${}^*\!\!\int_a^b D(x, \Delta x) = \inf\{F(b) - F(a) \mid \Delta F(x) \succ D(x, \Delta x)\}, \tag{24}$$

where the infimum is taken over all *majorants* of the form (a function is a majorant of a form if its difference infinitesimally majorizes this form). The lower and exact Perron–Stieltjes integrals are defined in the standard way on the basis of the upper one.

Theorem 7.1. *The upper Perron–Stieltjes integral of a difference form $D(x, \Delta x)$ is equal to the upper Leibniz integral of the associated infinitesimal form $D^*(x, dx)$.*

Proof. If $\Delta F(x) \succ D(x, \Delta x)$ for all $x \in [a, b]$, then $dF(x) \succ D^*(x, dx)$ by Lemma 7.1. Hence, for any infinitesimal partition of $[a, b]$, we have the inequality

$${}^*\!\!\int_{\mathbf{a}}^{\mathbf{b}} D(\mathbf{x}, d\mathbf{x}) \le F(b) - F(a),$$

which, since both $F(x)$ and the partition are arbitrary, implies the inequality for the integrals,

$$ {}^{*}\!\!\int_a^b D(x, dx) \le {}^{*}\!\!\int_a^b D(x, \Delta x). $$

To prove the reverse inequality, we consider the upper Leibniz integral of the infinitesimal form

$$ F(y) = {}^{*}\!\!\int_a^y D^*(x, dx) $$

as a function of the upper limit. For any positive ε, the differential of $F(y) + \varepsilon y$ infinitesimally majorizes $D(x, \Delta x)$, and hence, it will allow one to estimate from above the integral of the difference form in terms of the integral of the infinitesimal form plus $\varepsilon(b-a)$. Making ε tend to zero, we conclude that the integral of the difference form is not greater than the integral of the corresponding infinitesimal form. □

As a corollary of this theorem, we immediately see that the Perron–Stieltjes integrability of a difference form is equivalent to the Leibniz integrability of the infinitesimal form generated by it. Moreover, the Perron–Stieltjes and Leibniz integrals for the corresponding forms are equal.

8. Integral as the Limit of Integral Sums

The Kurzweil–Henstock Integral. Eventually, the Riemann and Riemann–Stieltjes integral sums were found to be adequate for the purposes of the theory of integral of functions of one variable. The problem of convergence of integral sums to the integral was solved with the proposal of the variable-scale Kurzweil–Henstock filter, which secured this convergence. At first sight, the problem of integration of non-Stieltjes forms, like the form $df(x)\,dg(x)$, does not call for the use of a markup on partitions, and hence it becomes unclear what filter should be used. It turns out that a markup is also required in this situation. Even though it has no effect on the integral sums, it is capable of controlling the sizes of the intervals in a partition in order to ensure the convergence of integral sums.

A function associating with each interval of a partition some (marked) point of this interval will be called a *markup* of the partition. By a *marked partition* we shall mean a pair consisting of a partition and its markup. A marked point in an interval I will be denoted by I^*.

A *scale* (see [5]) on some set is any nonnegative function on this set. A positive scale is a scale not assuming zero values. A scale $\delta_1(x)$ is said to be smaller than another scale $\delta_2(x)$ if the inequality $\delta_1(x) < \delta_2(x)$ is satisfied for all x from the set under consideration.

A *scale* of a marked partition of an interval $[a, b]$ is a function defined on this interval that vanishes outside the marked points of this partition and assumes at a marked point the value equal to the maximum distance of this marked point from the end-points of the interval in which it is marked. The scale of a marked partition $[P]^*$ will be denoted by $[P]^*(x)$.

We define the value of a difference form $D(x, \Delta x)$ on an interval $I = [c, d]$ containing a base point x as follows:

$$ D(x, I) = D(x, d - x) + D(x, c - x). \tag{25} $$

Let $[P]$ be a partition of an interval $[a, b]$ with markup $*$. The *integral sum* over this marked partition for a difference form $D(x, \Delta x)$ is defined as

$$ \sum_{I \in [S]} D(I^*, I). \tag{26} $$

A difference form $D(x, \Delta x)$ is said to be *Kurzweil–Henstock integrable* to an integral I if, for any $\varepsilon > 0$, there exists a positive scale $\delta(x)$ such that the integral sum of this form over any marked partition of smaller scale differs from I by less than ε.

The Topology on Scales. Now we define a topology on the set of all scales as follows: this topology is generated by open sets of the form $\{\delta'(x) \mid \delta'(x) < \delta(x)\}$, where $\delta(x)$ is an arbitrary positive scale.

We let $0(x)$ denote the zero scale (that is, the function that is identically zero). In this case, the Kurzweil–Henstock integral is defined as the limit of the integral sums:

$$\int_a^b D(x, \Delta x) = \lim_{[P]^*(x)\to 0(x)} \sum_{I\in[P]} D(I^*, I). \tag{27}$$

Let $Q*$ be some set from $[a, b]$ consisting of nonoverlapping intervals with marked points. We say that $Q*$ is contained in a marked partition $[P]^*$ of this interval (written $Q^* \subset [P]^*$) if $[P]$ contains Q and if the markups from $[P]^*$ and Q^* coincide on all intervals from Q. By $Q^*(x)$ we denote the scale of Q^*; that is, the function that differs from zero only at marked points from Q^*.

Lemma 8.1. *If a form $D(x, \Delta x)$ is Kurzweil–Henstock integrable on $[a, b]$, then, for any set of nonoverlapping marked intervals Q^*,*

$$\lim_{[P]^*(x)\to Q^*(x)} \sum_{I\in[P]} D(I, I^*) = \sum_{I\in[P]\setminus Q} \int_I D(x, \Delta x) + \sum_{I\in Q^*} D(I, I^*),$$

where the limit is taken over the set $\{P \mid [P]^* \supset Q^*\}$.

Proof. First of all, one verifies that the Kurzweil–Henstock integrability is hereditary in subintervals. The further argument is an exercise in expanding definitions. □

From the Leibniz Integral to the Kurzweil–Henstock Integral.

Lemma 8.2. *If $D_1^*(x, dx) \prec D_2^*(x, dx)$ for any $x \in [a, b]$, then for any $\varepsilon > 0$ there exists a positive scale $\delta(x)$ such that the integral sum over any marked partition of smaller scale for the difference form $D_1(x, \Delta x)$ is not greater than that for the form $D_2(x, \Delta x)$.*

Proof. For any point x, by the condition $D_1^*(x, dx) \prec D_2^*(x, dx)$ and Lemma 7.1 there exists $\delta(x)$ such that

$$D_1(x, \beta) \le D_2(x, \beta) \tag{28}$$

for any $x \in [a, b]$ with $|\beta| < \delta(x)$. In this case, each term of the integral sum of the form D_1 will not exceed the corresponding term of the form D_2. Hence, the corresponding inequality is also satisfied for the sums. □

Theorem 8.1. *If an infinitesimal form $D^*(x, dx)$ is Leibniz integrable, then the difference form $D(x, \Delta x)$ that generates it is Kurzweil–Henstock integrable, the integrals agreeing*

$$\int_a^b D(x, \Delta x) = \int_a^b D^*(x, dx).$$

Proof. Assume that $dF(x) \succ D^*(x, dx) \succ dG(x)$ are ε-close majorant and minorant of the form under consideration (they exist because the form is integrable). By Lemma 8.2, we may find a pair of positive scales $\delta_1(x)$ and $\delta_2(x)$, of which the first one, for partitions of smaller scale, secures that an estimate from above for the Riemann–Stieltjes integral sums is not greater than the Leibniz integral $\int_a^b f(x)\, dg(x)$ plus ε, while the second one gives an estimate from below by $\int_a^b f(x)\, dg(x) - \varepsilon$. But since ε is arbitrary, this implies the convergence of the integral sums to the Leibniz integral when the scale tends to zero. □

From Kurzweil–Henstock to Perron–Stieltjes. The proof of the converse theorem follows exactly the same lines as that of Theorem 9.12 in [5] on the equivalence of the Perron and Kurzweil–Henstock integrals, the only difference being that instead of the form $f(x)\Delta x$ one takes an arbitrary difference form. We give the proof for the convenience of the reader.

Lemma 8.3 (Saks and Henstock)**.** *If a form $D(x, \Delta x)$ is integrable on $[a, b]$ and a positive scale $\delta(x)$ is such that the inequality*

$$\Big| \sum_{I \in [P]} D(I^*, I) - \int_a^b D(x, \Delta x) \Big| \le \varepsilon \tag{29}$$

holds for any marked partition $[P]^*$ of smaller scale with some $\varepsilon > 0$, then the inequality

$$\Big| \sum_{I \in P_1^*} \Big(D(I^*, I) - \int_I D(x, \Delta x) \Big) \Big| \le \varepsilon \tag{30}$$

also holds for any marked partition $[P]^*$ of scale $< \delta(x)$ and any subset $P_1^* \subset [P]^*$.

Proof. We consider the set $\mathcal{P}_1$ of marked partitions of the interval $[a, b]$ that contains P_1^*. By Lemma 8.1, under the condition $P(x) \to P_1^*(x)$ the limit over this set of the integral sums of the form $D(x, \Delta x)$ equals the mixed sum. But the difference between this sum and the integral of the form over the entire interval $[a, b]$ is exactly the left-hand side of inequality (30). Moreover, this limit may not be greater than ε, because all the integral sums over partitions of smaller scale (than in the hypotheses of the lemma) satisfy inequality (29). □

Lemma 8.4 (Kolmogorov and Henstock)**.** *Assume that a form $D(x, \Delta x)$ is integrable on $[a, b]$ and a scale $\delta(x)$ is such that inequality (29) holds with some $\varepsilon > 0$ for any marked partition $[P]^*$ of this interval of smaller scale. Then the inequality*

$$\sum_{I \in [P]} \Big| D(I^*, I) - \int_a^b D(x, \Delta x) \Big| \le 2\varepsilon \tag{31}$$

also holds for any marked partition $[P]^*$ of this interval of smaller scale.

Proof. We set

$$P_1^* = \Big\{ I \in [P]^* \;\Big|\; D(I^*, I) \ge \int_I D(x, \Delta x) \Big\}$$

and define P_2^* as the complement of P_1^* in $[P]^*$. Applying the Saks–Henstock lemma in turn to P_1^* and P_2^* and adding the resulting inequalities, we get the result required. □

Theorem 8.2. *If a difference form is Kurzweil–Henstock integrable, then it is Perron–Stieltjes integrable.*

Proof. In the proof of this theorem, the integral of a difference form will be understood in the sense of Kurzweil–Henstock. Assume that $D(x, \Delta x)$ is Henstock integrable. Given $\varepsilon > 0$, we find a scale $\delta(x)$ such that the integral sums over partitions of smaller scale differ from the integral by at most ε. Consider the function

$$H(t) = \sup_{[P_t]^*} \sum_{I \in [P_t]^*} \Big\| D(I^*, I) - \int_I D(x, \Delta x) \Big\|, \tag{32}$$

where the supremum is taken over all marked partitions of the interval $[a, t]$ of scale $[P]^*(x) < \delta(x)$. We set $H(a) = 0$. Then $H(x)$ is a nondecreasing function. By the Kolmogorov–Henstock lemma, we have

$$0 \le H(x) \le 2\varepsilon. \tag{33}$$

Next, we fix an arbitrary point $\theta \in [a,b]$. If $\theta \in [x_1, x_2] \subset [\theta - \delta(\theta), \theta + \delta(\theta)]$, then

$$H(x_1) + \left| D(\theta, [x_1, x_2]) - \int\limits_{x_1}^{x_2} D(x, \Delta x) \right| \le H(x_2). \tag{34}$$

Indeed, replacing $H(x_1)$ in this inequality by an arbitrarily close sum of type (32), we get on the left-hand side a sum of this type already for $H(x_2)$.

If we denote by $F(x)$ the integral

$$\int\limits_a^x D(t, \Delta t),$$

then the inequality (34) implies the eventual inequality for infinitesimal forms:

$$|D^*(x, dx) - dF(x)| \prec dH(x).$$

As a result, we have that

$$dF(x) + dH(x) \succ D^*(x, dx) \succ dF(x) - dH(x).$$

So $d\big(F(x) \pm H(x)\big)$ can be looked upon as, respectively, the minorant and the majorant of the form $D^*(x, dx)$ with arbitrarily small difference in view of (33). This, in turn, implies its integrability to $F(b) - F(a)$. □

Thus, the approaches to the definition of a difference form by means of integral sums (the Riemann–Stieltjes–Kurzweil–Henstock approach) and differentials (the Leibniz–Perron–Stieltjes approach) are shown to be equivalent.

This research was carried out with the financial support of the Russian Science Foundation (project 14-50-00005).

REFERENCES

1. G. de L'Hôpital, *Analyse des infiniment petits pour l'intelligence des lignes courbes*, Paris (1968)
2. A. N. Kolmogorov, "Untersuchungen über den Integralbegriff," *Math. Ann.*, **103**, 654–696 (1930).
3. S. Leader, "What is a differential? A new answer from the generalized Riemann integral," *Am. Math. Mon.*, **93**, No. 5, 348–356 (1986).
4. S. Leader, *The Kurzweil–Henstock Integral and Its Differential: A Unified Theory of Integration on* $\mathbb{R}$ *and* $\mathbb{R}^n$, Pure Appl. Math., Vol. 242, Chapman & Hall/CRC (2001).
5. T. P. Lukashenko, V. A. Skvortsov, and A. P. Solodov, *Generalized Integrals* [in Russian], Librokom, Moscow (2011).
6. A. J. Ward, "The Perron–Stieltjes integral," *Math. Z.*, **41**, 578–604 (1936).

E. V. Shchepin
Steklov Mathematical Institute of the Russian Academy of Sciences, Moscow, Russia
E-mail: `scepin@mi.ras.su`