

§1 Introduction

We will approach Riemann-Stieltjes integrals using Riemann-Stieltjes sums instead of the upper and lower sums that appear in Rudin’s book. We will also study Riemann-Stieltjes integrals with “integrators” $\alpha(x)$ that are not monotone, but are “of bounded variation,” for I want to be able to define Riemann-Stieltjes integrals when the values of my functions belong to an infinite dimensional vector space, where upper and lower sums don’t make sense (no such functions appear in this note!). This makes little difference in the case of real-valued functions, since (as we will see later) functions of bounded variation can always be expressed as the difference of two monotone functions. At first, we don’t need “bounded variation,” so that concept’s development will wait until it is needed. In addition to Rudin’s book, I have relied on Chapter 2 of *Measure and Integral*, by Wheeden & Zygmund; Dekker, 1977.

Throughout this note, our functions $f(x)$ will be “finite-valued.” They may be real, complex, or vector-valued. Their values will thus lie in a vector space. They can thus be added pointwise, and multiplied by scalars, and their values always have finite “distance from zero,” denoted $|f(x)|$, which can denote absolute value (as in the cases of real or complex values) or *norm*, such as the length of a vector, or the “ L^p norm” and the “ L^q norm” that appear in Exercise 10(c) of Chapter 6, in case $f(x)$ is actually a function of t for each $x \dots$ We always assume that the “absolute value” is *complete*: Cauchy sequences converge. Just be aware that expressions such as $f(\xi_i)\Delta\alpha_i$ can be interpreted in a vast variety of ways, only demanding that the “product” $f(\xi_i)\Delta\alpha_i$ makes sense and that $|f(\xi_i)\Delta\alpha_i| \leq |f(\xi_i)||\Delta\alpha_i|$; all three of the “absolute values” appearing in the last inequality can have different meanings! **In this note, all Theorems are stated with f real-valued.**

Riemann-Stieltjes sums

A *Riemann-Stieltjes sum* for a function $f(x)$ defined on an interval $[a, b]$ is formed with the help of

- (1.1) A *partition* π of $[a, b]$, namely an ordered, finite set of points x_i , with $a = x_0 < x_1 < \dots < x_n = b$ (where n is a positive integer that can be any positive integer, and one that we will often write as $n = n_\pi$),
- (1.2) A *selection vector* $\xi = (\xi_1, \dots, \xi_n)$ that has $n = n_\pi$ components is *compatible with* π if $x_{i-1} \leq \xi_i \leq x_i$ for $i = 1, 2, \dots, n$; we require that ξ be compatible with π , and
- (1.3) An *integrator* $\alpha(x)$, a function defined on $[a, b]$ that plays the rôle of the x in $dx \dots$

A *Riemann-Stieltjes sum for f over $[a, b]$ with respect to the partition π , using the selection vector ξ (compatible with π), and integrator α* , may be denoted (in greatest detail!) as follows, and it is given by the value of the sum following it:

$$(1.4) \quad RS(f, \alpha, [a, b], \pi, \xi) := \sum_{i=1}^{n_\pi} f(\xi_i)(\alpha(x_i) - \alpha(x_{i-1})).$$

We try to allow context to let us drop some of the items listed inside the $RS(\dots)$.

More notation; the mesh (size) of a partition

In this definition, as in the Riemann-sums definition, we can write $\Delta x_i := x_i - x_{i-1}$ or $\Delta\alpha_i := \alpha(x_i) - \alpha(x_{i-1})$. These are convenient because they are short and suggest the dx or $d\alpha$ in an integral. But they can cause confusion because they leave out the dependence they have on x_{i-1} and x_i . The Δx_i is used in the Riemann-Stieltjes context too!

A partition π can be thought of as “dividing” the interval $[a, b]$ into subintervals. We may write $\pi|[a, b]$ and read this as “ π divides $[a, b]$,” or “partitions $[a, b]$.” We will denote the *intervals of* π by $I_i := [x_{i-1}, x_i]$. When we wish to work with 2 partitions at the same time we will have to distinguish between them somehow, for example we can use x'_j to denote the other’s points, I'_j to denote its intervals and $\Delta\alpha'_j$ to denote $\alpha(x'_j) - \alpha(x'_{j-1})$.

We measure the “fineness” of a partition (this is a crude measure!) using the length of the longest interval in the partition. This number is called the “mesh size” of the partition (think of a window screen), written

$$(1.5) \quad \text{mesh}(\pi) := \max_{1 \leq i \leq n_\pi} (x_i - x_{i-1}) = \max_{1 \leq i \leq n_\pi} \Delta x_i.$$

This definition of mesh size is used [and not $\max_{1 \leq i \leq n_\pi} (\alpha(x_i) - \alpha(x_{i-1}))$] even in the Riemann-Stieltjes context!

The Riemann-Stieltjes-sum definition of the Riemann-Stieltjes integral

(1.6) **Definition:** A real-valued function $f(x)$ defined on the bounded and closed interval $[a, b]$ is Riemann-Stieltjes integrable on $[a, b]$ with respect to α if there exists a number RSI such that for all $\epsilon > 0$ there exists $\delta > 0$ such that for every partition π of $[a, b]$, and every selection vector ξ compatible with π ,

$$\text{mesh}(\pi) < \delta \Rightarrow |R(f, \pi, \xi) - RSI| < \epsilon. \text{ We write } \int_a^b f d\alpha = \int_a^b f(x) d\alpha(x) := RSI$$

and we call $\int_a^b f d\alpha$ the Riemann-Stieltjes integral of f over $[a, b]$ with respect to α .

If f and α are real-valued and we imagine the set of all numbers $RS(f, \alpha, \pi, \xi)$ that can be formed (using all possible partitions whose mesh sizes are less than δ and all possible appropriate selection vectors) the definition demands that they all lie in the open interval $(RSI - \epsilon, RSI + \epsilon)$.

A difficulty with the definition; the Cauchy criterion for Riemann-Stieltjes integrability

In order to tell whether f is Riemann-Stieltjes integrable we have to know a number that is a candidate for $\int_a^b f(x) d\alpha(x)$. The idea of Cauchy sequences leads to the following Theorem, which gives an equivalent definition.

(1.7) **Theorem (Cauchy criterion for Riemann-Stieltjes integrability):** A function that is defined on $[a, b]$ is Riemann-Stieltjes integrable over $[a, b]$ with respect to α , defined on $[a, b]$, if and only if for all $\epsilon > 0$ there exists $\delta > 0$ such that for all partitions π and π' of $[a, b]$, and for all selection vectors ξ and ξ' compatible with π and π' , respectively,

$$\text{mesh}(\pi) < \delta \text{ and } \text{mesh}(\pi') < \delta \Rightarrow |RS(f, \alpha, \pi, \xi) - RS(f, \alpha, \pi', \xi')| < \epsilon.$$

Proof: First we suppose that f is Riemann-Stieltjes integrable over $[a, b]$ with respect to α . Then, using $\epsilon/2$ in the definition (1.6) of Riemann-Stieltjes integrability, we obtain $\delta > 0$ and RSI (by hypothesis!) such that for all partitions π of $[a, b]$,

$$\text{mesh}(\pi) < \delta \Rightarrow |RS(\pi) - RSI| < \epsilon/2 \text{ (we used context to save some writing).}$$

Now we suppose that π and π' are partitions of $[a, b]$ and that

$$\text{mesh}(\pi) < \delta \text{ and } \text{mesh}(\pi') < \delta.$$

Then for all selection vectors ξ and ξ' compatible with π and π' , respectively,

$$|RS(f, \alpha, \pi, \xi) - R(f, \alpha, \pi', \xi')| \leq |RS(\pi, \xi) - RSI| + |RSI - RS(\pi', \xi')| < \epsilon/2 + \epsilon/2 = \epsilon.$$

This completes half the proof.

Next we suppose that the Cauchy condition, given in the Theorem, is satisfied. We have to find a candidate for $\int_a^b f(x) d\alpha(x)$. We first construct a special sequence of partitions of $[a, b]$. We let π_n denote the partition that divides $[a, b]$ into n equal parts (π_n has points $x_{ni} := a + i \frac{b-a}{n}$). Finally we define selection vectors ξ_n by

$$\xi_{ni} := a + i \frac{b-a}{n}, \quad i = 1, \dots, n \text{ and define } \sigma_n := \sum_{i=1}^n f(\xi_{ni})(\alpha(x_{ni}) - \alpha(x_{n,i-1})),$$

a Riemann-Stieltjes sum ($\sigma_n = RS(f, \alpha, \pi_n, \xi_n)$). Now, given $\epsilon > 0$, we use $\epsilon/2$ in the Cauchy criterion, and obtain $\delta > 0$ such that

$$\text{mesh}(\pi) < \delta \text{ and } \text{mesh}(\pi') < \delta \Rightarrow |RS(f, \alpha, \pi, \xi) - RS(f, \alpha, \pi', \xi')| < \epsilon/2.$$

Then, if n and n' are so large that $(b-a)/n < \delta$ and $(b-a)/n' < \delta$, we have

$$\text{mesh}(\pi_n) < \delta \text{ and } \text{mesh}(\pi_{n'}) < \delta \Rightarrow |\sigma_n - \sigma_{n'}| < \epsilon/2.$$

This means (since ϵ was arbitrary) that $\{\sigma_n\}$ is a Cauchy sequence in our space. Thus we define

$$RSI := \lim_{n \rightarrow \infty} \sigma_n$$

and it remains to show that if $\pi|_{[a, b]}$ then

$$\text{mesh}(\pi) < \delta \Rightarrow |RS(\pi) - RSI| < \epsilon.$$

This is essentially done. We choose the first n such that $\text{mesh}(\pi_n) < \delta$, and we suppose that $\text{mesh}(\pi) < \delta$. Then

$$|RS(\pi) - RSI| \leq |RS(\pi) - \sigma_n| + |\sigma_n - RSI| < \epsilon/2 + \epsilon/2 = \epsilon,$$

since $RS(\pi) - \sigma_n = RS(\pi) - RS(f, \alpha, \pi_n, \xi_n)$. The proof is complete.