Introduction

We will approach Riemann-Stieltjes integrals using Riemann-Stieltjes sums instead of the upper and lower sums that appear in Rudin's book. The main reasons are that I want to study Riemann-Stieltjes integrals with "integrators" $\alpha(x)$ that are not monotone, but are "of bounded variation," and (most important) 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.

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 $x\ldots$ We always assume that the "absolute value" is *complete;* Cauchy sequences converge. In this note, you can do fine by thinking of f as being real-valued. 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 here can have different meanings! We say: f takes its values in a complete normed space.

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) A *partition* π of $[a, b]$, namely an ordered, finite set of points x_i, with $a = x_0 < x_1 < \cdots < x_n = b$ (where n is a positive integer that can be any positive integer, and one that we will often write as $n = n_\pi$),

(2) A *selection vector* $\xi = (\xi_1, \ldots, \xi_n)$ that has n_π components that must satisfy $x_{i-1} \leq \xi_i \leq x_i$, for $i = 1, 2, \ldots, n$.

and

(3) An *integrator* $\alpha(x)$, which is a function defined on $[a, b]$ that plays the rôle of the x in $dx\ldots$

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

$$RS(f, \alpha, [a, b], \pi, \xi) := \sum_{i=1}^{n_\pi} f(\xi_i)(\alpha(x_i) - \alpha(x_{i-1})). \tag{4}$$

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

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}. The Δx_i is used in the Riemann-Stieltjes context!

A partition π can be thought of as "dividing" the interval $[a, b]$ into subintervals. We may write $\pi\big|[a, b]$ 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 y_j to denote the other's points and J_j to denote its intervals, etc.

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 written

$$\operatorname{mesh}(\pi) := \max_{1\leq i\leq n_\pi} (x_i - x_{i-1}) = \max_{1\leq i\leq n_\pi} \Delta x_i. \tag{5}$$

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

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

(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* $[a,\, b]$,

$$mesh(\pi) < \delta \Rightarrow |RS(f,\, \pi) - RSI| < \epsilon.$$

We write

$$\int_a^b f\, d\alpha = \int_a^b f(x)\, d\alpha(x) := RSI$$

and we call this 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 $RS(f,\, \alpha,\, \pi)$ that can be formed (using all possible appropriate selection vectors and all possible partitions whose mesh sizes are less than δ), the definition demands that they all lie in the open interval $(RSI - \epsilon,\, RSI + \epsilon)$. When we had $\alpha(x) = x$ this led to a Theorem.

(7) **Theorem:** *If* f *is Riemann integrable on* $[a,\, b]$ *then* f *is bounded on* $[a,\, b]$.

This Theorem has to be modified in the Riemann-Stieltjes context! A simple example: suppose that $[a,\, b]$ is $[0,\, 1]$ and that $\alpha(x) = 0$ if $0 \le x \le c$, where $0 < c < 1$, and $\alpha(x) = 1$ if $c < x \le 1$. Then every function $f(x)$ that is continuous at c is Riemann-Stieltjes integrable on $[0,\, 1]$ with respect to this α. In particular the function that is $1/x$ except at zero, where we define it to be zero, is Riemann-Stieltjes integrable on $[0,\, 1]$ with respect to this α, but f is not bounded. The difference is that when $\alpha(x)$ was just x, we had $\Delta x_i > 0$ for every i. In our example, $\Delta\alpha_i = 0$ unless I_i contains c and some d with $c < d$. What we need (to make the proof we used in the Riemann case work) is that on the set where the function α *"really" varies*, f must be bounded. To make a definition, we will extend the definitions of f and α beyond the interval $[a,\, b]$ by setting them equal to their values at the endpoints. Thus we think of $f(x) = f(a)$ if $x < a$ and $f(x) = f(b)$ if $x > b$, with the same idea used to extend α. We now define the *oscillation of* f *on an interval* U by

(8) $\omega(f,\, U) := \sup_{x,\, y \in U} |f(x) - f(y)|$. We allow the interval to be open or half-open now!

As before, we will let $\omega_i = \omega_i(f) = \omega(f,\, I_i)$ when I_i is an interval (closed!) of a partition π. But now we need to use oscillations of α as well!

(9) **Definition:** *If* $\alpha(x)$ *is defined for* $x \in [a,\, b]$, *we denote by* $\Omega = \Omega(\alpha,\, [a,\, b])$ *the set of all* $c \in [a,\, b]$ *such that every open interval* U *that contains* c *contains* $x_1 < c < x_2$ *with* $|\alpha(x_1) - \alpha(x_2)| > 0$.
Note: c *can* be a or b because of our extension beyond $[a,\, b]$! For instance, if for all $\delta > 0$ there exists x_2 such that $a < x_2 < a + \delta$ and $|\alpha(a) - \alpha(x_2)| > 0$, then $a \in \Omega(\alpha,\, [a,\, b])$ because for every $x_1 < a$ we have $|\alpha(x_1) - \alpha(x_2)| = |\alpha(a) - \alpha(x_2)| > 0$.

(10) **Exercise:** *Prove that* $\Omega(\alpha,\, [a,\, b])$ *is closed.*

In our example, $\Omega(\alpha,\, [a,\, b]) = \{c\}$.

(11) **Theorem:** *If* f *is Riemann-Stieltjes integrable on* $[a,\, b]$ *with respect to* α *then* f *is bounded on* $\Omega(\alpha,\, [a,\, b])$.

Proof: Suppose not. Then there exists a sequence $\{x_n\}$ in $\Omega := \Omega(\alpha,\, [a,\, b])$ such that $|f(x_n)| > n$. Since $f(x)$ is finite at every point x in Ω, there are infinitely many distinct x_n, and so some subsequence (that we will still denote $\{x_n\}$) converges to a point x^* in Ω. We now choose $\epsilon = 1$ in the definition of Riemann-Stieltjes integrability, and obtain a corresponding $\delta > 0$. We can then construct a partition π_o with mesh size less than δ in such a way that x^* is contained in the *interior* of some interval I_{i_o} of π_o (unless x^* is an endpoint of $[a,\, b]$; in that case, we can, by the **Note,** still use the following argument, with $I_{i_o} = I_1$ or $I_{i_o} = I_{n_\pi}$). We know that every neighborhood of x^* contains infinitely many of the x_n. Now we will refine π_o. We know that $\mathrm{Int}(I_{i_o})$ contains points $\hat{x}_1 < x^* < \hat{x}_2$ with $|\alpha(\hat{x}_1) - \alpha(\hat{x}_2)| > 0$. We add these points to π_o, giving us a new partition π, and $\mathrm{mesh}(\pi) < \delta$. We will now call $[\hat{x}_1,\, \hat{x}_2]$, which is an interval of π, $\hat{I}$. Next we pick the components ξ_i of a selection vector ξ in an arbitrary way when $I_i \neq \hat{I}$, and we let $\hat{\xi}$ be some $x_N \in \hat{I}$. Then $|RS(f, \pi, \xi) - RSI| < 1$. We next modify ξ by changing *only* $\hat{\xi} = x_N$ to $\hat{\xi}' := x_M$, where $x_M \in \hat{I}$, and we call the new selection vector ξ'. Then $|RS(f, \pi, \xi') - RSI| < 1$, $RS(f, \pi, \xi') - RS(f, \pi, \xi) = (f(x_M) - f(x_N))(\alpha(\hat{x}_1) - \alpha(\hat{x}_2))$ and

$$RS(f, \pi, \xi') - RSI = RS(f, \pi, \xi) - RSI + (f(x_M) - f(x_N))(\alpha(\hat{x}_1) - \alpha(\hat{x}_2)).$$

By choosing M very large compared to N we can arrange that $|f(x_M) - f(x_N)||\alpha(\hat{x}_1) - \alpha(\hat{x}_2)| > 2$. Then

$$1 > |RS(f, \pi, \xi') - RSI| \geq |RS(f, \pi, \xi') - RS(f, \pi, \xi)| - |RS(f, \pi, \xi) - RSI| > 2 - 1 = 1.$$

The definition of Riemann-Stieltjes integrability is contradicted. Hence f is bounded on $\Omega(\alpha, [a, b])$ if f is Riemann-Stieltjes integrable with respect to α.

Note: From now on, we will usually say "f is Riemann-Stieltjes integrable" instead of "f is Riemann-Stieltjes integrable with respect to α."

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 $\int_a^b f(x)\, d\alpha(x)$. The idea of a Cauchy sequence leads to the following Theorem, which gives an equivalent definition.

(12) **Theorem (Cauchy criterion for Riemann-Stieltjes integrability):** *A function 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' *associated with* π *and* π', *respectively,*

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

Proof: First we suppose that f is Riemann-Stieltjes integrable over $[a, b]$ with respect to α. Then, using $\epsilon/2$ in the definition of Riemann-Stieltjes integrability, we obtain $\delta > 0$ and RSI 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' associated with π and π', respectively,

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

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 sequence of partitions of $[a, b]$. We let π_n denote the partition that divides $[a, b]$ into n equal parts $\left(\pi_n \text{ has points } x_{ni} := a + i\dfrac{b-a}{n}\right)$. Finally we define selection vectors ξ_n by

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

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

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

and it remains to show that if $\pi\big|[a,\, b]$ 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

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

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

Remarks: What we have done so far is inspired by the material in Chapter 2, Section 3 of the book *Measure and Integral,* by Wheeden and Zygmund. There, nothing is said at first about the functions f and α, beside the demand that the integrability definition hold. They point out and prove this, in subsection 4:

If f and α have a discontinuity at the same point, then the Riemann-Stieltjes integral does not exist.

They also show, in Theorem (2.21), that if the Riemann-Stieltjes integral exists, then this integration-by-parts formula holds:

$$\int_a^b f\, d\alpha = -\int_a^b \alpha\, df + f(b)\alpha(b) - f(a)\alpha(a) \tag{13}$$

(in the applications I have far, far back in my mind, the integral on the right would have to be $\int_a^b df\, \alpha$, in order to keep the order of "multiplication" the same). The proof amounts to rearranging the Riemann-Stieltjes sums, adding and subtracting terms in such a way that the ξ_i become partition points and the x_i become selection-vector components when $1 < i < n_\pi$. There are some leftovers, and these turn out to be the "boundary" term $f(x)\alpha(x)\big|_a^b$.

Wheeden and Zygmund state several properties, routine to prove, about Riemann-Stieltjes integrals:

$$\int_a^b f\, d\alpha \text{ is linear in both } f \text{ and } \alpha \tag{14}$$

as long as all the integrals involved exist, and

$$\begin{gathered}\text{if } \int_a^b f\, d\alpha \text{ exists and } a < c < b, \text{ then}\\ \text{both of } \int_a^c f\, d\alpha \text{ and } \int_c^b f\, d\alpha \text{ exist,}\\ \text{and } \int_a^c f\, d\alpha + \int_c^b f\, d\alpha = \int_a^b f\, d\alpha.\end{gathered} \tag{15}$$

Caution! It is possible that both of $\int_a^c f\, d\alpha$ and $\int_c^b f\, d\alpha$ exist, but that $\int_a^b f\, d\alpha$ does not exist! Counterexamples can be constructed using the "common-discontinuity" phenomenon Wheeden and Zygmund discuss in subsection 4.

What has been covered applies to *all* Riemann-Stieltjes integrals, including those that follow! That continuity plays a rôle has already been mentioned.

Functions of bounded variation: definition and properties

In what we do from now on, at least one of f and α will be a function of bounded variation, unless otherwise stated. We will begin by discussing real-valued functions of bounded variation. This material can also be found in *Measure and Integral,* by Wheeden and Zygmund.

(16) **Definition:** *A function $f : [a,\, b] \to \mathbb{R}$ is a function of bounded variation on $[a,\, b]$ if*

$$V(f,\, [a,\, b]) := \sup_{\pi|[a,b]} \sum_1^{n_\pi} |f(x_i) - f(x_{i-1})| < \infty \text{ and we say that } f \in BV[a,b].$$

To go farther it will be useful to have some more notation. If π is a partition of $[a, b]$ we will write

$$(17)\qquad V_\pi = V_\pi(f,\,[a,\,b]) := \sum_1^{n_\pi} |f(x_i) - f(x_{i-1})|, \ \text{ so that } \ V = \sup_{\pi|[a,b]} V_\pi \ \text{ (here, } f \text{ and } [a,\,b] \text{ are “assumed”).}$$

We can call V_π the “π-variation” of f over $[a, b]$. Since f has a finite value for each $\pi|[a, b]$, V_π is always finite. However, V can be infinite. This is so, for example, if f is the Dirichlet function.

Each V_π pays attention *only* to the absolute value of the *difference* between the values at the opposite ends of an interval of the partition π. We will need to take the *signs* of those differences into account, and they will lead to two new “variations.”

For a real number x we define its *positive part* to be $x^+ := \max\{0,\, x\}$ and we define its *negative part* to be $x^- := \max\{0,\, -x\}$. Both “parts” are non-negative, and we have $x^+ + x^- = |x|$ and $x^+ - x^- = x$.

(18) **Exercise:** *Prove that for all real numbers* x *and* y, $(x+y)^+ \le x^+ + y^+$ *and* $(x+y)^- \le x^- + y^-$. *These are “triangle inequalities!” What can be said about* $(xy)^+$ *and* $(xy)^-$?

We now define the “positive” and “negative” “π-variations” of f over $[a, b]$:

$$(19)\qquad P_\pi = P_\pi(f,\,[a,\,b]) := \sum_1^{n_\pi} (f(x_i) - f(x_{i-1}))^+ \ \text{ and } \ N_\pi = N_\pi(f,\,[a,\,b]) := \sum_1^{n_\pi} (f(x_i) - f(x_{i-1}))^-.$$

(20) **Definition:** *The positive variation,* $P = P(f,\,[a,\,b])$ *and the negative variation* $N = N(f,\,[a,\,b])$ *of* f *over* $[a,\,b]$ *are given by* $P = \sup\limits_{\pi|\,[a,b]} P_\pi$ *and* $N = \sup\limits_{\pi|\,[a,b]} N_\pi$ *respectively.*

For example, if f increases on $[a, b]$, $P_\pi = V_\pi = f(b) - f(a)$ and $N_\pi = 0$. If we look at $f(x) := |x|$ on $[-1,\,1]$ we will always have $0 \le P_\pi \le 1$ and $0 \le N_\pi \le 1$, and $0 \le V_\pi \le 2$.

Because of how x^+ and x^- were defined, we always have (for *any* function)

$$(21)\qquad P_\pi + N_\pi = V_\pi \ \text{ and } \ P_\pi - N_\pi = f(b) - f(a) \ \text{ (telescoping sums!).}$$

(22) *If* τ *is a refinement of* π, *we always have* $O_\pi \le O_\tau$, *where* O *stands for any of the letters* N, P *or* V. This follows from several applications of the triangle inequality.

Some properties of functions of bounded variation

(23) *If* $f \in BV[a,\,b]$ *then* f *is bounded on* $[a,\,b]$. *Proof:* Suppose $a \le x \le b$. Then, if we let $\pi := \{a,\,x,\,b\}$,

$$|f(x)| = |f(x) - f(a) + f(a)| \le |f(a)| + |f(x) - f(a)| + |f(b) - f(x)| = |f(a)| + V_\pi \le |f(a)| + V.$$

(24) *The space* $BV[a,\,b]$ *is a vector space. For all* $c \in \mathbb{R}$ *and all* $f \in BV[a,\,b]$, $V(cf,\,[a,\,b]) = |c|V(f,\,[a,\,b])$. *For all* $f \in BV[a,\,b]$ *and* $g \in BV[a,\,b]$, $V(f,\,[a,\,b]) \le V(f,\,[a,\,b]) + V(g,\,[a,\,b])$; $V(f,\,[a,\,b]) = 0$ *if and only if* f *is constant. Proof:* The second assertion follows from these facts: for all $\pi|\,[a,\,b]$, $V_\pi(cf,\,[a,\,b]) = |c|V_\pi(f,\,[a,\,b])$; $\sup\{|c|x : x \in E\} = |c| \sup\{x : x \in E\} = |c| \sup E$. The first assertion and the first part of the third one follow from the second one and the triangle inequality. Finally, suppose that $V(f,\,[a,\,b]) = 0$ and that $a \le x \le b$. Then, with $\pi := \{a,\,x,\,b\}$, $|f(x) - f(a)| \le |f(x) - f(a)| + |f(b) - f(x)| = V_\pi = 0$. Therefore $f(x) \equiv f(a)$.

(25) *If* $f \in BV[a,\,b]$ *and* $a < c < b$ *then* $f \in BV[a,\,c]$ *and* $f \in BV[c,\,b]$, *and conversely.*
Moreover, $V = V(f,\,[a,\,b]) = V(f,\,[a,\,c]) + V(f,\,[c,\,b])$.

Proof: If $f \in BV[a,\,b]$ and $a < c < b$, let partitions $\sigma|\,[a,\,c]$ and $\tau|\,[c,\,b]$ be given. Then $\pi := \sigma \cup \tau$ is a partition of $[a,\,b]$ so $V_\sigma + V_\tau = V_\pi \le V$, hence $V_\sigma \le V$ and $V_\tau \le V$. Thus $f \in BV[a,\,c]$ and $f \in BV[c,\,b]$. Conversely, suppose that $a < c < b$ and that $f \in BV[a,\,c]$ and $f \in BV[c,\,b]$. Let $\pi|\,[a,\,b]$. Then $\pi_c := \pi \cup \{c\}$ is a refinement of π. Therefore $V_\pi \le V_{\pi_c} = V_\sigma + V_\tau$, where $\sigma := \pi_c \cap [a,\,c]$ and τ is defined similarly. By hypothesis, $V_\pi \le V_{\pi_c} = V_\sigma + V_\tau \le V(f,\,[a,\,c]) + V(f,\,[c,\,b])$. Thus $V(f,\,[a,\,b]) \le V(f,\,[a,\,c]) + V(f,\,[c,\,b]) < \infty$. This proves part of the asserted equality. To show the other inequality, now that we know $V < \infty$ let partitions $\sigma|\,[a,\,c]$ and

$\tau|\,[c,\, b]$ be given. We recall that earlier we had $V_\sigma + V_\tau = V_{\pi_c} \le V$, so $V_\sigma + V_\tau \le V$ whenever $\sigma|\,[a,\, c]$ and $\tau|\,[c,\, b]$ were arbitrary partitions of $[a,\, c]$ and $[c,\, b]$, respectively.

$$\text{Thus} \quad \sup_{\sigma|\,[a,\,c]} (V_\sigma + V_\tau) = V(f,\, [a,\, c]) + V_\tau \le V, \quad \text{and so} \quad \sup_{\tau|\,[c,\,b]} (V(f,\, [a,\, c]) + V_\tau) = V(f,\, [a,\, c]) + V(f,\, [c,\, b]) \le V.$$

Please note that the first inequality holds for an *arbitrary* $\tau|\,[c,\, b]$, making the second one valid.

(26) **Exercise:** *Prove that the equality in* (25) *holds for every function* $f : [a,\, b] \to \mathbb{R}$, *whether* f *is a function of bounded variation or not.*

Motivated by (25), when $f : [a,\, b] \to \mathbb{R}$ and $a \le x \le b$ we can define the three functions

$$V(x) := V(f,\, [a,\, x]), \qquad P(x) := P(f,\, [a,\, x]) \ \text{ and } \ N(x) := N(f,\, [a,\, x]).$$

Each of these is an increasing function of x. Jordan's Theorem asserts that if $f \in BV[a,\, b]$ we can represent f in terms of $P(x)$ and $N(x)$.

(27) **Theorem (Jordan):** *A function* $f \in BV[a,\, b]$ *if and only if there exist functions* g *and* h, *both increasing on* $[a,\, b]$, *such that* $f(x) = g(x) - h(x)$ *for* $a \le x \le b$. *If this is the case, then* $P(x) \le g(x) - g(a)$, $N(x) \le h(x) - h(a)$ *and* $f(x) = f(a) + P(x) - N(x)$ *for* $a \le x \le b$.

Proof: Suppose first that $f(t) = g(t) - h(t)$, $t \in [a,\, b]$, where the functions g and h are both increasing on $[a,\, b]$. Let $\pi\,|[a,\, b]$ (later, we will apply this when $x \in [a,\, b]$ and $\pi\,|[a,\, x]$). Then

$$\Delta f_i = f(x_i) - f(x_{i-1}) = \Delta g_i - \Delta h_i \begin{cases} \le & \Delta g_i \\ \ge & -\Delta h_i. \end{cases} \tag{28}$$

Thus $-\Delta h_i \le \Delta f_i \le \Delta g_i$, so $|\Delta f_i| \le \max\{\Delta g_i,\, \Delta h_i\} \le \Delta g_i + \Delta h_i$ for $1 \le i \le n_\pi$. Hence $V_\pi(f) \le V_\pi(g) + V_\pi(h) = g(b) - g(a) + h(b) - h(a) < \infty$, so $f \in BV[a,\, b]$.

Next, we show that $f(x) = f(a) + P(x) - N(x)$ for $a \le x \le b$. But we will do this just by showing it for $x = b$. Then we can use (25) and let each $x \in [a,\, b]$ play the rôle of b. This will show the existence of the functions $g(x)(= f(a) + P(x))$ and $h(x)(= N(x))$. We will use (21) and (22). After that is done, we'll prove the P–g and N–h inequalities.

By the definitions of P, N and V we know there exist sequences $\{\pi_k\}$, $\{\rho_k\}$ and $\{\sigma_k\}$ such that $P_{\pi_k} \to P$, $N_{\pi_k} \to N$ and $V_{\sigma_k} \to V$. Let us define $\tau_k := \pi_k \cup \rho_k \cup \sigma_k$. By (22) $P_{\pi_k} \le P_{\tau_k} \le P$. By the Squeeze Principle $P_{\tau_k} \to P$. Similarly, $N_{\tau_k} \to N$ and $V_{\tau_k} \to V$. By (21) and Limit Theorems

$$P + N = V \ \text{ and } \ P - N = f(b) - f(a) \ \text{ and the second is the same as } \ f(x) = f(a) + P(x) - N(x)$$

when $x = b$. By (24) we can use any $x \in [a,\, b]$ in place of b by restricting our attention to f on $[a,\, x]$.

Now suppose that $f(x)$ is *defined* as the difference of two increasing functions on $[a,\, b]$: $f(t) = g(t) - h(t)$. We will use (28) again, as well as the following observation: $t \mapsto t^+$ is increasing and $t \mapsto t^-$ is decreasing. Therefore, with the help of (28), applied to partitions of $[a,\, x]$, $(\Delta f_i)^+ \le (\Delta g_i)^+ = \Delta g_i$ and $\Delta h_i = (-\Delta h_i)^- \ge (\Delta f_i)^-$. Hence $P_\pi(f,\, [a,\, x]) \le P_\pi(g,\, [a,\, x]) = g(x) - g(a)$. Similarly, $h(x) - h(a) = N_\pi(-h,\, [a,\, x]) \ge N_\pi(f,\, [a,\, x])$. When, in each case, we take the supremum over all $\pi\,|[a,\, x]$, we get $P(x) \le g(x) - g(a)$ and $N(x) \le h(x) - h(a)$. These inequalities "say" that there is no "wasted cancellation" in the formula $f(x) = f(a) + P(x) - N(x)$.

(29) **Problem:** *Prove that if* $f(x) \in BV[a,\, b]$ *and* $f(x)$ *is continuous at* $x_o \in [a,\, b]$ *then so are* $P(x)$, $N(x)$ *and* $V(x)$.