The change of variable formula for the Riemann-Stieltjes integral

Alberto Torchinsky

April 1, 2019

Abstract

We consider general formulations of the change of variable formula for the Riemann-Stieltjes integral, including the case when the substitution is not invertible.

1 Introduction.

This note concerns general formulations of the change of variable, or substitution, formula for the Riemann-Stieltjes integral. A prototype of our results is the following,

Substitution Formula. Let \( \varphi \) be a bounded, Riemann integrable function defined on an interval \( I = [a, b] \) that does not change sign on \( I \), and let \( \Phi \) be an indefinite integral of \( \varphi \) on \( I \). Let \( \psi \) be a bounded, Riemann integrable function defined on \( \Phi(I) \), the range of \( \Phi \), and let \( \Psi \) be an indefinite integral of \( \psi \) on \( \Phi(I) \).

Then, if a bounded function \( f \) defined on \( \Phi(I) \) is Riemann integrable with respect to \( \Psi \) on \( \Phi(I) \), \( f(\Phi) \psi(\Phi) \) is Riemann integrable with respect to \( \Phi \) on \( I \), and in that case, with \( \mathcal{I} = [\Phi(a), \Phi(b)] \),

\[
\int_{\mathcal{I}} f d\Psi = \int_{I} f(\Phi) \psi(\Phi) d\Phi. \tag{1}
\]

Thus, the substitution formula holds when the Riemann-Stieltjes integral is computed with respect to an arbitrary function \( \Psi \), and the substitution \( \Phi \) is invertible. Together with the change of variable formula established below, which holds when \( \Psi \) is monotone, or a difference of monotone functions, and \( \Phi \) is not necessarily invertible, these formulas constitute the main results in
this note. Though it seems that these formulas should be known, they are not present in the classic or standard literature in the area.

We will begin by introducing the necessary definitions and notations. Fix a closed finite interval $I = [a, b] \subset \mathbb{R}$, and let $\Phi$ be a continuous monotone (increasing) function defined on $I$. For a partition $\mathcal{P} = \{I_k\}$ of $I$, where $I_k = [x_{k,l}, x_{k,r}]$, and a bounded function $f$ on $I$, let $U(f, \Phi, \mathcal{P})$ and $L(f, \Phi, \mathcal{P})$ denote the upper and lower Riemann sums of $f$ with respect to $\Phi$ on $I$ along $\mathcal{P}$, i.e.,

$$U(f, \Phi, \mathcal{P}) = \sum_k (\sup_{I_k} f) (\Phi(x_{k,r}) - \Phi(x_{k,l})),$$

and

$$L(f, \Phi, \mathcal{P}) = \sum_k (\inf_{I_k} f) (\Phi(x_{k,r}) - \Phi(x_{k,l})).$$

respectively, and set

$$U(f, \Phi) = \inf_{\mathcal{P}} U(f, \Phi, \mathcal{P}), \quad \text{and} \quad L(f, \Phi) = \sup_{\mathcal{P}} L(f, \Phi, \mathcal{P}).$$

We say that $f$ is Riemann integrable with respect to $\Phi$ on $I$ if $U(f, \Phi) = L(f, \Phi)$, and in this case the common value is denoted $\int_I f d\Phi$, the Riemann integral of $f$ with respect to $\Phi$ on $I$.

When $\Phi(x) = x$ one gets the usual Riemann integral on $I$, and $\Phi$ is omitted in the above notations. And, throughout this note, when it is clear from the context, integrable means Riemann integrable with respect to $\Phi(x) = x$, and Riemann-Stieltjes integrable means integrable with respect to a general $\Phi$.

The following are working characterizations of integrability [3], [6], [16]. A bounded function $f$ defined on $I$ is Riemann integrable with respect to $\Phi$ on $I$ iff, given $\varepsilon > 0$, there is a partition $\mathcal{P}$ of $I$, which may depend on $\varepsilon$, such that

$$U(f, \Phi, \mathcal{P}) - L(f, \Phi, \mathcal{P}) \leq \varepsilon. \quad (2)$$

Furthermore, a sequential characterization holds, to wit, (2) is equivalent to the existence of a sequence of partitions $\{\mathcal{P}_n\}$ of $I$ such that

$$\lim_n (U(f, \Phi, \mathcal{P}_n) - L(f, \Phi, \mathcal{P}_n)) = 0,$$

and in this case

$$\lim_n U(f, \Phi, \mathcal{P}_n) = \lim_n L(f, \Phi, \mathcal{P}_n) = \int_I f d\Phi. \quad (3)$$
Also, integrability can be characterized in terms of oscillations. Recall that, given a bounded function \( f \) defined on \( I \) and an interval \( J \subset I \), the oscillation \( \text{osc} (f, J) \) of \( f \) on \( J \) is defined as \( \text{osc} (f, J) = \sup_J f - \inf_J f \).

Then, a bounded function \( f \) is Riemann integrable with respect to \( \Phi \) on \( I \) iff, given \( \varepsilon > 0 \), there is a partition \( P = \{ I_k \} \) of \( I \), which may depend on \( \varepsilon \), such that

\[
\sum_k \text{osc} (f, I_k) \left( \Phi(x_{k,r}) - \Phi(x_{k,l}) \right) \leq \varepsilon. \tag{4}
\]

And, a sequential characterization holds, to wit, (4) is equivalent to the existence of a sequence of partitions \( \{ P_n \} \) of \( I \) consisting of the intervals \( P_n = \{ I^n_k \} \) with \( I^n_k = [x^n_{k,l}, x^n_{k,r}] \), such that

\[
\lim_n \sum_k \text{osc} (f, I^n_k) \left( \Phi(x^n_{k,r}) - \Phi(x^n_{k,l}) \right) = 0. \tag{5}
\]

These characterizations do not necessarily hold if \( \Phi \) fails to be monotone. Moreover, note that if (2) holds for a partition \( P \), it also holds for partitions \( P' \) finer than \( P \). Invoking (36), this observation applies to other concepts as well, including (3), (4), and (5).

Finally, since \( \Phi \) is continuous and increasing on \( I \), \( \Phi(I) \) is an interval \( \mathcal{I} = [\Phi(a), \Phi(b)] \) with endpoints \( \Phi(a) \) and \( \Phi(b) \). Note that each interval \( J = [y_1, y_2] \subset \mathcal{I} \) is of the form \( [\Phi(x_1), \Phi(x_2)] \), where \( \Phi(x_1) = y_1, \Phi(x_2) = y_2 \), and \( [x_1, x_2] \) is a subinterval of \( I \). Moreover, partitions \( P \) of \( I \) induce a corresponding partition \( Q \) of \( \mathcal{I} \), and, conversely, every partition of \( \mathcal{I} \) can be expressed as \( Q \) for some partition \( P \) of \( I \).

2 The Substitution Formula.

We begin by proving a result that includes the familiar substitution formula [1].

**Proposition.** Let \( \Phi \) be a continuous monotone function defined on \( I \), and \( \Psi \) defined on \( \mathcal{I} = \Phi(I) \). Let \( f \) be a bounded function on \( \mathcal{I} \). Then, \( f \) is Riemann integrable with respect to \( \Psi \) on \( \mathcal{I} \) iff \( f(\Phi) \) is Riemann integrable with respect to \( \Psi(\Phi) \) on \( I \), and in that case we have

\[
\int_{\mathcal{I}} f \, d\Psi = \int_I f(\Phi) \, d\Psi(\Phi). \tag{6}
\]

**Proof.** Specifically, (6) means that, if the integral on either side of the equality exists, so does the integral on the other side and they are equal. To see
this, let the partition $Q = \{I_k\}$ of $\mathcal{I}$ correspond to the partition $P = \{I_k\}$ of $I$ such that $I_k = [\Phi(x_{k,l}), \Phi(x_{k,r})]$, where $I_k = [x_{k,l}, x_{k,r}]$. Then, since $\sup_{I_k} f = \sup_{I_k} f(\Phi)$, it readily follows that

$$U(f, \Psi, Q) = \sum_k \left( \sup_{I_k} f \right) \left( \Psi(\Phi(x_{k,r})) - \Psi(\Phi(x_{k,l})) \right)$$

$$= \sum_k \left( \sup_{I_k} f(\Phi) \right) \left( \Psi(\Phi(x_{k,r})) - \Psi(\Phi(x_{k,l})) \right) = U(f(\Phi), \Psi(\Phi), P),$$

and, similarly, $L(f, \Psi, Q) = L(f(\Phi), \Psi(\Phi), P)$. (6) follows at once from these identities.

Next we consider the particular case of the substitution formula when both $\varphi$ and $\psi$ are of constant sign. In this instance we have,

**Lemma.** Let $\varphi$ be a bounded, Riemann integrable function defined on an interval $I = [a, b]$ that does not change sign on $I$, and let $\Phi$ be an indefinite integral of $\varphi$ on $I$. Let $\psi$ be a bounded, Riemann integrable function defined on $\Phi(I)$, the range of $\Phi$, that does not change sign, and let $\Psi$ be an indefinite integral of $\psi$ on $\Phi(I)$.

Let $f$ be bounded on $\Phi(I)$. Then, $f$ is integrable with respect to $\Psi$ on $\Phi(I)$ iff $f(\Phi)\psi(\Phi)$ is integrable with respect to $\Phi$ on $I$, and in that case, with $\mathcal{I} = [\Phi(a), \Phi(b)]$,

$$\int_{\mathcal{I}} f d\Psi = \int_{I} f(\Phi) \psi(\Phi) d\Phi. \quad (7)$$

**Proof.** It suffices to prove the result when $\varphi, \psi$ are positive. Indeed, if the result holds in this case, when $\varphi$ is negative it follows by replacing $\varphi$ by $-\varphi$, $\psi(x)$ by $\psi(-x)$, and $f(y)$ by $f(-y)$ in (7), and when $\psi$ is negative, by replacing $\psi$ by $-\psi$ in (7).

Now, by assumption we have,

$$\Phi(x) = \Phi(a) + \int_{[a,x]} \varphi, \quad x \in I,$$

and,

$$\Psi(y) = \Psi(\Phi(a)) + \int_{[\Phi(a),y]} \psi, \quad y \in \Phi(I).$$

First, assume that $f(\Phi)\psi(\Phi)$ is Riemann-Stieltjes integrable, and fix $\varepsilon > 0$. Then, for a partition $P = \{I_k\}$ of $I$ with $I_k = [x_{k,l}, x_{k,r}]$ and $I_k = [\Phi(x_{k,l}), \Phi(x_{k,r})]$, pick $\xi_k \in I_k$ such that

$$U(f(\Phi), \Psi(\Phi), P) \leq \sum_k f(\Phi(\xi_k)) \left( \Psi(\Phi(x_{k,r})) - \Psi(\Phi(x_{k,l})) \right) + \varepsilon. \quad (8)$$
Now, the sum on the right-hand side of (8) equals
\[
\sum_{k} f(\Phi(\xi_k)) \int_{I_k} (\psi - \psi(\Phi(\xi_k)))
\]
\[
+ \sum_{k} f(\Phi(\xi_k)) \psi(\Phi(\xi_k))(\Phi(x_{k,r}) - \Phi(x_{k,l})) = A + B,
\]
(9)
say, where clearly, with \(M_f\) a bound for \(f\), \(A \leq M_f \sum_k \text{osc}(\psi, I_k) |I_k|\), and \(B \leq U(f(\Phi)\psi(\Phi), \Phi, \mathcal{P})\). Thus, by (8) and (9),
\[
U(f(\Phi), \Psi(\Phi), \mathcal{P}) \leq M_f \sum_k \text{osc}(\psi, I_k) |I_k| + U(f(\Phi)\psi(\Phi), \Phi, \mathcal{P}) + \varepsilon.
\]
(10)

Applying (10) to \(-f\) gives
\[
-L(f(\Phi), \Psi(\Phi), \mathcal{P}) \leq M_f \sum_k \text{osc}(\psi, I_k) |I_k| - L(f(\Phi)\psi(\Phi), \Phi, \mathcal{P}) + \varepsilon,
\]
which added to (10) yields
\[
U(f(\Phi), \Psi(\Phi), \mathcal{P}) - L(f(\Phi), \Psi(\Phi), \mathcal{P}) \leq 2M_f \sum_k \text{osc}(\psi, I_k) |I_k|
\]
\[
+ (U(f(\Phi)\psi(\Phi), \Phi, \mathcal{P}) - L(f(\Phi)\psi(\Phi), \Phi, \mathcal{P})) + 2\varepsilon.
\]
(11)

Let the partition \(\mathcal{P}\) of \(I\) satisfy simultaneously (4) for \(\psi\) and (2) for \(f(\Phi)\psi(\Phi)\) with respect to \(\Phi\) for the \(\varepsilon > 0\) picked at the beginning of the proof; a common refinement of a partition that satisfies (4) for \(\psi\) and one that satisfies (2) for \(f(\Phi)\psi(\Phi)\) with respect to \(\Phi\) will do. Then from (11) it follows that \(U(f(\Phi), \Psi(\Phi), \mathcal{P}) - L(f(\Phi), \Psi(\Phi), \mathcal{P}) \leq 2M_f \varepsilon + \varepsilon + 2\varepsilon\), and, since \(\varepsilon > 0\) is arbitrary, by (2), \(f(\Phi)\) is Riemann-Stieltjes integrable and \(U(f(\Phi), \Psi(\Phi)) = L(f(\Phi), \Psi(\Phi)) = \int_I f(\Phi) \, d\Psi(\Phi)\).

To evaluate the integral in question, let the sequence of partitions \(\{\mathcal{P}_n\}\) of \(I\) satisfy simultaneously (5) for \(\psi\) and (3) for \(f(\Phi)\psi(\Phi)\). Then, given \(\varepsilon > 0\), from (10) it follows that
\[
\int_I f(\Phi) \, d\Psi(\Phi) = U(f(\Phi), \Psi(\Phi)) \leq \limsup_n U(f(\Phi), \Psi(\Phi), \mathcal{P}_n)
\]
\[
\leq \limsup_n M_f \sum_k \text{osc}(\psi, I_k) |I_k| + \limsup_n U(f(\Phi)\psi(\Phi), \Phi, \mathcal{P}_n) + \varepsilon,
\]
\[
= \int_I f(\Phi) \psi(\Phi) \, d\Phi + \varepsilon,
\]
5
which, since ε is arbitrary, gives \( \int_I f(\Phi)d\Psi(\Phi) \leq \int_I f(\Phi)\psi(\Phi)d\Phi \). Furthermore, replacing \( f \) by \(-f\) it follows that \( \int_I f(\Phi)\psi(\Phi)d\Phi \leq \int_I f(\Phi)d\Psi(\Phi) \), and the integrals are equal. Therefore, since \( \Phi \) is continuous, monotone on \( I \), by (6),

\[
\int_I f d\Psi = \int_I f(\Phi)d\Psi(\Phi) = \int_I f(\Phi)\psi(\Phi)d\Phi,
\]

(7) holds, and the conclusion obtains.

To prove the converse, by (6) and the assertion we just proved, it suffices to show that, if \( f(\Phi) \) is integrable with respect to \( \Psi(\Phi) \) on \( I \), \( f(\Phi)\psi(\Phi) \) is integrable with respect to \( \Phi \) on \( I \). Let, then, \( \mathcal{P} = \{I_k\} \) be a partition of \( I \), and, given \( \varepsilon > 0 \), pick \( \xi_k \in I_k \) such that

\[
U(f(\Phi)\psi(\Phi), \Phi, \mathcal{P}) \leq \sum_k f(\Phi(\xi_k))\psi(\Phi(\xi_k)) (\Phi(x_{k,r}) - \Phi(x_{k,l})) + \varepsilon. \tag{12}
\]

Now, there are two types of summands in (12), to wit, those where \( f(\Phi(\xi_k)) > 0 \), and those where \( f(\Phi(\xi_k)) < 0 \). In the former case, we have

\[
f(\Phi(\xi_k)) (\psi(\Phi(\xi_k)) + \inf_{I_k} \psi) (\Phi(x_{k,r}) - \Phi(x_{k,l})) \\
\leq f(\Phi(\xi_k)) \text{osc} (\psi, I_k) |I_k| + f(\Phi(\xi_k)) (\inf_{I_k} \psi) |I_k|,
\]

where the first term is bounded by \( M_f \text{osc} (\psi, I_k) |I_k| \), and the second by

\[
f(\Phi(\xi_k)) \int_{[\Phi(x_{k,l}), \Phi(x_{k,r})]} \psi = f(\Phi(\xi_k)) (\Psi(\Phi(x_{k,r})) - \Psi(\Phi(x_{k,l}))) \\
\leq \left( \sup_{I_k} f(\Phi) \right) (\Psi(\Phi(x_{k,r})) - \Psi(\Phi(x_{k,l}))).
\]

Along similar lines, since \( \int_{I_k} \psi \leq (\sup_{I_k} \psi) (\Phi(x_{k,r}) - \Phi(x_{k,l})) \), in the latter case we have

\[
f(\Phi(\xi_k))\psi(\Phi(\xi_k)) (\Phi(x_{k,r}) - \Phi(x_{k,l})) \\
= (-f(\Phi(\xi_k))) (-\psi(\Phi(\xi_k)) \pm \sup_{I_k} \psi) (\Phi(x_{k,r}) - \Phi(x_{k,l})) \\
\leq M_f \text{osc} (\psi, I_k) |I_k| - f(\Phi(\xi_k)) (\psi(\Phi(x_{k,r}) - \Phi(x_{k,l})) \\
\leq M_f \text{osc} (\psi, I_k) |I_k| + (\sup_{I_k} f(\Phi)) (\Psi(\Phi(x_{k,r})) - \Psi(\Phi(x_{k,l}))).
\]
Whence, combining these estimates it follows that
\[
\sum_k f(\Phi(\xi_k)) \psi(\Phi(\xi_k)) \left(\Phi(x_{k,r}) - \Phi(x_{k,l})\right) \\
\leq M_f \sum_k \text{osc} (\psi, I_k) |I_k| + U(f(\Phi), \psi(\Phi), \mathcal{P}),
\]
and by (12),
\[
U(f(\Phi)\psi(\Phi), \Psi(\Phi), \mathcal{P}) \leq M_f \sum_k \text{osc} (\psi, I_k) |I_k| + U(f(\Phi), \psi(\Phi), \mathcal{P}) + \varepsilon.
\]
Now, as in (11) it follows that
\[
U(f(\Phi)\psi(\Phi), \Phi, \mathcal{P}) - L(f(\Phi)\psi(\Phi), \Phi, \mathcal{P})) \leq 2 M_f \sum_k \text{osc} (\psi, I_k) |I_k| \\
+ (U(f(\Phi), \Psi(\Phi), \mathcal{P}) - L(f(\Phi), \Psi(\Phi), \mathcal{P})) + 2 \varepsilon,
\]
and so, since \(\varepsilon > 0\) is arbitrary, picking an appropriate \(\mathcal{P}\), by (2) we conclude that \(f(\Phi)\psi(\Phi)\) is Riemann-Stieltjes integrable, and the proof is finished.

We are now ready to prove the substitution formula stated in the introduction. In this case \(\psi\) is allowed to changes sign on \(\Phi(I)\).

**Proof, Substitution Formula.**

It suffices to prove the result when \(\varphi\) is positive, for when \(\varphi\) is negative on \(I\) the result follows by a direct proof, or simply by replacing \(\varphi\) by \(-\varphi\), \(\psi(x)\) by \(\psi(-x)\), and \(f(y)\) by \(f(-y)\) in (1). Let \(f\) be integrable with respect to \(\Psi\) on \(\Phi(I)\). The idea is to show that \(\int_{\Phi(I)} f d\Psi\) can be approximated arbitrarily close by the Riemann sums of \(f(\Phi)\psi(\Phi)\) with respect to \(\Phi\) on \(I\), and, consequently, \(\int_I f(\Phi)\psi(\Phi) d\Phi\) also exists, and the integrals are equal [2], [11], [17]. To make this argument precise we begin by introducing the partitions used for the approximating Riemann sums. They are based on a partition \(\mathcal{Q}\) of \(\Phi(I)\) defined as follows: given \(\eta > 0\), by (4), there is a partition \(\mathcal{Q} = \{I_k\}\) of \(\Phi(I)\), such that
\[
\sum_k \text{osc} (\psi, I_k) |I_k| \leq \eta^2 |I|.
\]

We first separate the indices \(k\) that appear in \(\mathcal{Q}\) into three classes, the (good) set \(G\), the (bounded) set \(B\), and the (undulating) set \(U\), according
to the following criteria. First, \( k \in G \) if \( \psi \) is strictly positive or negative on \( I_k \). Next, \( k \in B \), if \( k \notin G \) and \( |\psi| \leq \eta \) on \( I_k \). And, finally, \( k \in U \), if \( k \notin G \cup B \). Note that for \( k \in U \) we have \( \text{osc} \left( \psi, I_k \right) \geq \eta \), since \( \psi \) changes signs in \( I_k \) and for at least one point \( \zeta_k \) there, \( |\psi(\zeta_k)| > \eta \).

Recall that to each subinterval \( I_k = [\Phi(x_{k,l}), \Phi(x_{k,r})] \) of \( \Phi(I) \) corresponds an interval \( I_k = [x_{k,l}, x_{k,r}] \), and let \( P \) denote the partition of \( I \) given by \( P = \{I_k\} \).

Now, since \( f \) is integrable with respect to \( \Psi \) on \( \Phi(I) \), \( f \) is integrable with respect to \( \Psi \) on \( I_k \), and if \( k \in G \), since \( \varphi \) and \( \psi \) don’t change sign, by the Lemma, \( f(\Phi)\psi(\Phi) \) is integrable with respect to \( \Phi \) on \( I_k \), and \( \int_{I_k} f \, d\Psi = \int_{I_k} f(\Phi) \psi(\Phi) d\Phi \). Then, by (4), given \( \eta > 0 \), there is a partition \( P^k = \{I_j^k\} \) of \( I_k \) such that

\[
\sum_j \text{osc} \left( f(\Phi)\psi(\Phi), I_j^k \right) |I_j^k| \leq \eta |I_k|.
\]

Moreover, since \( \int_{I_k} f \, d\Psi \leq U(f(\Phi)\psi(\Phi), \Phi, P^k) \), we also have

\[
U(f(\Phi)\psi(\Phi), \Phi, P^k) - \int_{I_k} f \, d\Psi \leq \eta |I_k|.
\]

Hence,

\[
\sum_{k \in G} \sum_j \text{osc} \left( f(\Phi)\psi(\Phi), I_j^k \right) |I_j^k| \leq \eta \sum_{k \in G} |I_k|,
\]

and

\[
\sum_{k \in G} \left| \int_{I_k} f \, d\Psi - U(f(\Phi)\psi(\Phi), \Phi, P^k) \right| \leq \eta \sum_{k \in G} |I_k|.
\]

Now, for \( k \in B \cup U \), let \( P^k = \{I_k\} \) denote the partition of \( I_k \) consisting of the interval \( I_k \). Note that, with \( M_{\varphi} \) a bound for \( \varphi \),

\[
|\Phi(x_{k,r}) - \Phi(x_{k,l})| = |I_k| \leq \int_{[x_{k,l}, x_{k,r}]} |\varphi| \leq M_{\varphi} |I_k|,
\]

and, with \( M_\psi \) a bound for \( \psi \),

\[
\left| \int_{I_k} f \, d\Psi \right| \leq M_f |\Psi(\Phi(x_{k,r})) - \Psi(\Phi(x_{k,l}))| \leq M_f M_\psi M_{\varphi} |I_k|.
\]

Thus, by (16),

\[
\text{osc} \left( f(\Phi)\psi(\Phi), I_k \right) |I_k| \leq 2 M_f M_\psi M_{\varphi} |I_k|,
\]

(18)
and by (17), for \( \xi_k \in I_k \),
\[
\left| \int_{I_k} f d\Psi - f(\Phi(\xi_k))\psi(\Phi(\xi_k))(\Phi(x_k,r) - \Phi(x_k,l)) \right| \leq 2MfM\psi M\varphi |I_k|,
\]
and so, picking \( \xi_k \in I_k \) appropriately, we get
\[
\left| \int_{I_k} f d\Psi - U(f(\Phi)\psi(\Phi), \Phi, P^k) \right| \leq 3MfM\psi M\varphi |I_k|.
\]  
(19)

Now, if \( k \in B \), \( M\psi \leq \eta \), and, therefore, by (18),
\[
\sum_{k \in B} \text{osc} \left( f(\Phi)\psi(\Phi), I_k \right) |I_k| \leq 2MfM\varphi \eta \sum_{k \in B} |I_k|
\]  
(20)

and by (19),
\[
\sum_{k \in B} \left| \int_{I_k} f d\Psi - U(f(\Phi)\psi(\Phi), \Phi, P^k) \right| \leq 3MfM\varphi \eta \sum_{k \in B} |I_k|.
\]  
(21)

Finally, since for \( k \in U \) we have \( \text{osc} (\psi, I_k) \geq \eta \), from (13) it follows that
\[
\eta \sum_{k \in U} |I_k| \leq \sum_{k \in U} \text{osc} (\psi, I_k) |I_k| \leq \sum_{k} \text{osc} (\psi, I_k) |I_k| \leq \eta^2 |I|,
\]
and, consequently,
\[
\sum_{k \in U} |I_k| \leq \eta |I|.
\]  
(22)

Whence, by (18) and (22), the \( U \) terms are bounded by
\[
\sum_{k \in U} \text{osc} \left( f(\Phi)\psi(\Phi), I_k \right) |I_k| \leq 2MfM\psi M\varphi \sum_{k \in U} |I_k| \leq 2MfM\psi M\varphi \eta |I|,
\]  
(23)

and, by (19) and (22),
\[
\sum_{k \in U} \left| \int_{I_k} f - U(f(\Phi)\psi(\Phi), \Phi, P^k) \right|
\leq 3MfM\psi M\varphi \sum_{k \in U} |I_k| \leq 3MfM\psi M\varphi \eta |I|.
\]  
(24)

Consider now the partition \( P' \) of \( I \) that consists of the union of all the intervals in the \( P^k \), where each \( P^k \) is defined according as to whether
Given \( \varepsilon > 0 \), pick \( \eta > 0 \) so that \((1 + 2M_fM_{\varphi} + 2M_fM_{\psi}M_{\varphi})\eta |I| \leq \varepsilon \), and note that the above expression is < \( \varepsilon \), and since \( \varepsilon > 0 \) is arbitrary and \( \Phi \) is monotone, (4) corresponding to \( P' \) implies that \( f(\Phi)\psi(\Phi) \) is Riemann-Stieltjes integrable, and \( L(f(\Phi)\psi(\Phi), \Phi) = U(f(\Phi)\psi(\Phi), \Phi) = \int_I f(\Phi)\psi(\Phi) \, d\Phi \).

It remains to compute the integral in question. First, note that

\[
U(f(\Phi)\psi(\Phi), \Phi, P') = \sum_k U(f(\Phi)\psi(\Phi), \Phi, P_k^k). \tag{26}
\]

Moreover, since \( \Phi(b) - \Phi(a) = \sum_k \left( \Phi(x_k^r) - \Phi(x_k^l) \right) \), by the linearity of the integral, taking orientation into account, it follows that \( \int_I f = \sum_k \int_{I_k} f \), [12], [15]. Hence, regrouping according to the sets \( G, B \) and \( U \), gives

\[
\int_I f \, d\Psi = \sum_{k \in G} \int_{I_k} f \, d\Psi + \sum_{k \in B} \int_{I_k} f \, d\Psi + \sum_{k \in U} \int_{I_k} f \, d\Psi, \tag{27}
\]

and, from (26) and (27), it follows that

\[
\left| \int_I f \, d\Psi - U(f(\Phi)\psi(\Phi), \Phi, P') \right| \leq \sum_k \left| \int_{I_k} f \, d\Psi - U(f(\Phi)\psi(\Phi), \Phi, P_k^k) \right|
\]

\[
+ \sum_{k \in B} \left| \int_{I_k} f \, d\Psi - U(f(\Phi)\psi(\Phi), \Phi, P_k^k) \right|
\]

\[
+ \sum_{k \in U} \left| \int_{I_k} f \, d\Psi - U(f(\Phi)\psi(\Phi), \Phi, P_k^k) \right| = s_1 + s_2 + s_3,
\]

say. Now, by (15),

\[
s_1 \leq \sum_{k \in G} \left| \int_{I_k} f \, d\Psi - U(f(\Phi)\psi(\Phi), \Phi, P_k^k) \right| \leq \eta \sum_{k \in G} |I_k| \leq \eta |I|,
\]
and by (21) and (24), $s_2 + s_3 \leq (3M_fM_\varphi + 3M_fM_\psi M_\varphi) \eta |I|$, which combined imply that
\[
\left| \int_I f d\Psi - U(f(\Phi))\psi(\Phi, \Phi', \mathcal{P}') \right| \leq (1 + 3M_fM_\varphi + 3M_fM_\psi M_\varphi) \eta |I|.
\]

Given $\varepsilon > 0$, pick $\eta > 0$ so that $(1 + 3M_fM_\varphi + 3M_fM_\psi M_\varphi) \eta |I| \leq \varepsilon$, and note that
\[
\left| \int_I f d\Psi - U(f(\Phi))\psi(\Phi, \Phi', \mathcal{P}') \right| \leq \varepsilon. \tag{28}
\]

Also, since $U(f(\Phi)\psi(\Phi, \Phi, \mathcal{P}')) - L(f(\Phi)\psi(\Phi, \Phi, \mathcal{P}'))$ is equal to the left-hand side of (25), from (28) it follows that
\[
\left| \int_I f d\Psi - L(f(\Phi)\psi(\Phi, \Phi, \mathcal{P}')) \right| \leq 2\varepsilon. \tag{29}
\]

Furthermore, since by (28),
\[
\int_I f(\Phi)\psi(\Phi)d\Phi = U(f(\Phi)\psi(\Phi, \Phi)) \leq U(f(\Phi)\psi(\Phi, \Phi', \mathcal{P}')) \leq \int_I f d\Psi + \varepsilon,
\]
and by (29),
\[
\int_I f d\Psi \leq L(f(\Phi)\psi(\Phi, \Phi, \mathcal{P}')) + 2\varepsilon \leq L(f(\Phi)\psi(\Phi, \Phi)) + 2\varepsilon = \int_I f(\Phi)\psi(\Phi)d\Phi + 2\varepsilon,
\]
we conclude that
\[
\left| \int_I f d\Psi - \int_I f(\Phi)\psi(\Phi)d\Phi \right| \leq 2\varepsilon,
\]
which, since $\varepsilon$ is arbitrary, implies that $\int_I f(\Phi)\psi(\Phi)d\Phi = \int_I f d\Psi$. Hence, (1) holds, and the proof is finished. \vspace{10pt}

\section{The Change of Variable Formula.}

The next result corresponds to the case when $\varphi$ is of variable sign, and in this case the substitution is not required to be invertible. Then $\Phi(I)$, the range of $\Phi$, is an interval, but $\Phi(a), \Phi(b)$ are not necessarily endpoints of
this interval. It is important to keep in mind that the Riemann integral is oriented, and that the direction in which the interval is traversed determines the sign of the integral. When \( \Psi(x) = x \), the formula is related to the general formulation by Preiss and Uher \[11\] of Kestelman’s result pertaining the change of variable formula for the Riemann integral \[4\], \[7\]. In fact, the integral on the right-hand side of (1) can be computed as a Riemann integral \[9\], \[17\], to wit,

\[
\int_I f(\Phi) \psi(\Phi) d\Phi = \int_I f(\Phi)\psi(\Phi)\varphi.
\]

Specifically, we have,

**Change of Variable Formula.** Let \( \varphi \) be a bounded, Riemann integrable function defined on an interval \( I = [a,b] \), and let \( \Phi \) be an indefinite integral of \( \varphi \) on \( I \). Let \( \psi \) be a bounded, Riemann integrable function defined on \( \Phi(I) \), the range of \( \Phi \), that does not change sign on \( \Phi(I) \), and let \( \Psi \) be an indefinite integral of \( \psi \).

Let \( f \) be a bounded function defined on \( \Phi(I) \). Then, \( f \) is Riemann integrable with respect to \( \Psi \) on \( \Phi(I) \) iff \( f(\Phi)\psi(\Phi) \) is Riemann integrable with respect to \( \Phi \) on \( I \), and in that case, with \( I = [\Phi(a), \Phi(b)] \),

\[
\int_I f d\Psi = \int_I f(\Phi) \psi(\Phi) d\Phi.
\]  

**Proof.** The proof of the necessity follows along the lines of the substitution formula, and we shall be brief. Note that since replacing \( \psi \) by \( -\psi \) in (30) preserves the identity, it suffices to assume that \( \psi \) is positive. So, suppose that \( f \) is integrable with respect to \( \Psi \) on \( \Phi(I) \), and let the partition \( P \) of \( I \) be defined as follows: given \( \eta > 0 \), by (4), there is a partition \( P = \{I_k\} \) of \( I \), such that

\[
\sum_k \text{osc} (\varphi, I_k) |I_k| \leq \eta^2 |I|.
\]  

Separate the indices \( k \) that appear in \( P \) into three classes, \( G \), \( B \), and \( U \), according to the following criteria. First, \( k \in G \) if \( \varphi \) is strictly positive or negative on \( I_k \). Next, \( k \in B \), if \( k \notin G \) and \( |\varphi| \leq \eta \) on \( I_k \). And, finally, \( k \in U \), if \( k \notin G \cup B \). Note that for \( k \in U \), since \( \varphi \) changes signs in \( I_k \) and for at least one point \( \xi_k \) there, \( |\varphi(\xi_k)| > \eta \), we have \( \text{osc} (\varphi, I_k) \geq \eta \).

Recall that each \( I_k = [x_{k,l}, x_{k,r}] \) in \( P \) corresponds to the (oriented) subinterval \( \mathcal{I}_k = [\Phi(x_{k,l}), \Phi(x_{k,r})] \) of \( \Phi(I) \). Now, since \( f \) is integrable with respect to \( \Psi \) on \( \Phi(I) \), \( f \) is integrable with respect to \( \Psi \) on \( \mathcal{I}_k \), and if \( k \in G \), since \( \varphi \) and \( \psi \) don’t change sign, by the Lemma, \( f(\Phi)\psi(\Phi) \) is integrable \( \Phi \) on \( I_k \), and by (7), \( \int_{\mathcal{I}_k} f d\Psi = \int_{I_k} f(\Phi) \psi(\Phi) d\Phi \). Then, by (4), given
\[ \eta > 0, \text{ for each } k \in G, \text{ there is a partition } P^k = \{ I^k_j \} \text{ of } I_k, \text{ such that } \sum_j \osc (f(\Phi)\psi(\Phi), I^k_j) |I^k_j| \leq \eta |I_k|, \text{ and therefore,} \]

\[
\sum_{k \in G} \sum_j \osc (f(\Phi)\psi(\Phi), I^k_j) |I^k_j| \leq \eta \sum_{k \in G} |I_k|. \tag{32}
\]

Now, for \( k \in B \cup U \), let \( P^k = \{ I_k \} \) denote the partition of \( I_k \) consisting of the interval \( I_k \). Then, as in (18) and (19),

\[
\osc (f(\Phi)\psi(\Phi), I_k) |I_k| \leq 2 M_f M_\psi M_\varphi |I_k|, \tag{33}
\]

and

\[
\left| \int_{I_k} f d\Psi - U(f(\Phi)\psi(\Phi), \Phi, P^k) \right| \leq 3 M_f M_\psi M_\varphi |I_k|. \]

Now, if \( k \in B \), \( M_\varphi \leq \eta \), and, therefore, by (33),

\[
\sum_{k \in B} \osc (f(\Phi)\psi(\Phi), I_k) |I_k| \leq 2 M_f M_\psi \eta \sum_{k \in B} |I_k| \tag{34}
\]

and

\[
\sum_{k \in B} \left| \int_{I_k} f d\Psi - U(f(\Phi)\psi(\Phi), \Phi, P^k) \right| \leq 3 M_f M_\psi \eta \sum_{k \in B} |I_k|. \]

Finally, since for \( k \in U \) we have \( \sum_{k \in U} |I_k| \leq \eta |I| \), by (33), as in (23), the \( U \) terms are bounded by

\[
\sum_{k \in U} \osc (f(\Phi)\psi(\Phi), I_k) |I_k| \leq 2 M_f M_\psi M_\varphi \eta |I|, \tag{35}
\]

and, as in (24),

\[
\sum_{k \in U} \left| \int_{I_k} f - U(f(\Phi)\psi(\Phi), \Phi, P^k) \right| \leq 3 M_f M_\psi M_\varphi \eta |I|.
\]

Consider now the partition \( P' \) of \( I \) that consists of the union of all the partitions \( P^k \), where each \( P^k \) is defined according as to whether \( k \in G, k \in B \), or \( k \in U \). Then, by (32), (34), and (35),

\[
\sum_{k \in G} \sum_j \osc (f(\Phi)\psi(\Phi), I^k_j) |I^k_j| + \sum_{k \in B \cup U} \osc (f(\Phi)\psi(\Phi), I^k) |I^k| \leq (1 + 2 M_f M_\psi + 2 M_f M_\psi M_\varphi) \eta |I|.
\]

13
Given \( \varepsilon > 0 \), pick \( \eta > 0 \) so that \((1+2MfM_\psi+2MfM_\psi M_\varphi)\eta |I| \leq \varepsilon \), and note that the above expression is \( < \varepsilon \), and so, since \( \varepsilon > 0 \) is arbitrary, (4) corresponding to \( \mathcal{P}' \) implies that \( f(\Phi)\psi(\Phi) \) is Riemann-Stieltjes integrable on \( I \) and \( L(f(\Phi)\psi(\Phi),\Phi) = U(f(\Phi)\psi(\Phi),\Phi) = \int_I f(\Phi)\psi(\Phi) \, d\Phi \).

Making use of the integral estimates established above, \( \int_I f(\Phi)\psi(\Phi) \, d\Phi \) can be evaluated exactly as in the previous lemma; the computation is left to the reader.

As for the converse, it suffices to prove that, if \( f(\Phi)\psi(\Phi) \) is integrable with respect to \( \Phi \) on \( I \), \( f \) is integrable with respect to \( \Psi \) on \( \Phi(I) \), and then invoke the result we just proved. Let the partition \( \mathcal{P} \) of \( I \) satisfy (31). Since \( \Phi \) is continuous, \( \Phi(I) \) is a closed interval of the form \( [\Phi(x_m),\Phi(x_M)] \) with (possibly non-unique) \( x_m, x_M \) in \( I \). If \( x_m \) or \( x_M \) is an endpoint of (not necessarily the same) interval in \( \mathcal{P} \), proceed. Otherwise, since for an interval \( J = [x_l, x_r] \) and an interior point \( x \) of \( J \), with \( J_l = [x_l, x] \) and \( J_r = [x, x_r] \), we have

\[
\text{osc } (f,J_l) |J_l| + \text{osc } (f,J_r) |J_r| \leq \text{osc } (f,J) |J|, \tag{36}
\]

\( \mathcal{P} \) can be refined so that the endpoint that was not originally included is now an endpoint of two intervals of the new partition, without increasing the right-hand side of (31). For simplicity also denote this new partition \( \mathcal{P} \), note that it contains both \( x_m \) and \( x_M \) at least once as an endpoint of one of its intervals, and define the sets of indices \( G, B, \) and \( U \), associated to \( \mathcal{P} \), as above.

Now, if \( f(\Phi)\psi(\Phi) \) is integrable with respect to \( \Phi \) on \( I \), \( f(\Phi)\psi(\Phi) \) is integrable with respect to \( \Phi \) on \( I_k \), and, if \( k \in G \), since \( \varphi \) is of constant sign, by the lemma, \( f \) is integrable with respect to \( \Psi \) on \( \mathcal{I}_k \) and \( \int_{\mathcal{I}_k} f(\Phi)\psi(\Phi) \, d\Phi = \int_{\mathcal{I}_k} f \, d\Psi \). Then, by (4), given \( \eta > 0 \), there is a partition \( \mathcal{Q}^k = \{\mathcal{I}_j^k\} \) of \( \mathcal{I}_k \), where \( \mathcal{I}_j^k = [\Phi(x_{j,l}^k),\Phi(x_{j,r}^k)] \), such that

\[
\sum_j \text{osc } (f,\mathcal{I}_j^k) |\Psi(\Phi(x_{j,l}^k)) - \Psi(\Phi(x_{j,r}^k))| \leq \eta |I_k| \tag{37}
\]

As for \( k \in B \cup U \), as in (34),

\[
\text{osc } (f,\mathcal{I}_k) |\Psi(\Phi(x_{k,r})) - \Psi(\Phi(x_{k,l}))| \leq 2MfM_\psi M_\varphi |I_k|. \tag{38}
\]

Next, if \( k \in B, M_\varphi \leq \eta, \) and, therefore,

\[
\sum_{k \in B} \text{osc } (f,\mathcal{I}_k) |\Psi(\Phi(x_{k,r})) - \Psi(\Phi(x_{k,l}))| \leq 2MfM_\psi \eta \sum_{k \in B} |I_k|. \tag{38}
\]
Finally, for \( k \in U \), as in (35),
\[
\sum_{k \in U} \text{osc} (f, I_k) |\Psi(\Phi(x_{k,r})) - \Psi(\Phi(x_{k,l}))| \\
\leq 2 M_f M_\psi M_\varphi \sum_{k \in U} |I_k| \leq 2 M_f M_\psi M_\varphi \eta |I|.
\] (39)

Let \( Q' \) denote the collection of subintervals of \( \Phi(I) \) defined by
\[
Q' = \left( \bigcup_{k \in G} \bigcup_{j} \{I^k_j\} \right) \cup \left( \bigcup_{k \in B \cup U} \{I_k\} \right).
\]
Note that the union of the intervals in \( Q' \) is \( \Phi(I) \) and that, by (37), (38), and (39),
\[
\sum_{k \in G} \sum_{j} \text{osc} (f, I^k_j) |\Psi(\Phi(x^{j}_{k,r})) - \Psi(\Phi(x^{j}_{k,l}))| \\
+ \sum_{k \in B \cup U} \sum_{j} \text{osc} (f, I_k) |\Psi(\Phi(x_{k,r})) - \Psi(\Phi(x_{k,l}))| \\
\leq (1 + 2 M_f M_\psi + 2 M_f M_\psi M_\varphi) \eta |I|.
\] (40)

Consider now the finite set \( \Phi(x_m) = y_1 < y_2 < \cdots < \Phi(x_M) = y_h \), of the endpoints of the intervals in \( Q' \) arranged in an increasing fashion, without repetition. Suppose that the interval \( J \) in \( Q' \) contains the points \( y_{k_1}, \ldots, y_{k_n} \), say, as endpoints or interior points. If they are endpoints, disregard them, otherwise, as in (36), incorporate each as an endpoint of two intervals in a refined \( Q' \) without increasing the right-hand side of (40). Clearly \( Q' \) thus refined contains a partition \( Q'' = \{J_k\} \) of \( \Phi(I) \), which, by (40), satisfies
\[
\sum_{k} \text{osc} (f, J_k) |J_k| \leq (1 + 2 M_f M_\psi + 2 M_f M_\psi M_\varphi) \eta |I|.
\]

Given \( \varepsilon > 0 \), pick \( \eta > 0 \) such that \( (1 + 2 M_f M_\psi + 2 M_f M_\psi M_\varphi) \eta |I| \leq \varepsilon \). Then the sum in (4) corresponding to \( Q'' \) does not exceed an arbitrary \( \varepsilon > 0 \), and, therefore, \( f \) is integrable with respect to \( \Psi \) on \( \Phi(I) \), and the conclusion follows from the first part of the proof.  

4 Coda.

We close this note with a caveat: not always the most general result is the most useful. By strengthening some assumptions and weakening others in
the change of variable formula, it is possible to obtain a substitution formula that does not follow from this result [5].

Assume that the function Φ is continuous, increasing on \( I = [a, b] \), and differentiable on \((a, b)\) with derivative \( \varphi \geq 0 \); then Φ is uniformly continuous on \( I \), and maps \( I \) onto \( I' = [\Phi(a), \Phi(b)] \). Assume that \( \Psi \) is continuous, increasing on \( I' \), and differentiable on \( (\Phi(a), \Phi(b)) \) with derivative \( \psi \geq 0 \). We will also assume that \( f \) is Riemann integrable, rather than bounded, on \( I \). On the other hand, we will not assume that \( \varphi, \psi \) are bounded. Then, if \( f(\Phi)\psi(\Phi) \) is integrable with respect to Φ on \( I \), the change of variable formula holds.

To see this, consider a partition \( P = \{I_k\}, I_k = [x_{k,l}, x_{k,r}] \), of \( I \), and the corresponding partition \( Q = \{I'_k\} \) of \( I' \), consisting of \( I'_k = [y_{k,l}, y_{k,r}] \), where \( y_{k,l} = \Phi(x_{k,l}) \) and \( y_{n,r} = \Phi(x_{n,r}) \). By the mean value theorem there exist \( \zeta_k \in I_k \) such that

\[
\Psi(y_{k,r}) - \Psi(y_{k,l}) = \psi(\zeta_k)(y_{k,r} - y_{k,l}), \quad \text{all } k,
\]

and with \( \xi_k \in I_k \) such that \( \zeta_k = \Phi(\xi_k) \), all \( k \), it follows that

\[
\sum_k f(\zeta_k)(\Psi(y_{k,r}) - \Psi(y_{k,l})) = \sum_k f(\Phi(\xi_k))\psi(\Phi(\xi_k))(\Phi(x_{k,r}) - \Phi(x_{k,l})),
\]

where the left-hand side is a Riemann sum of \( f \) with respect to \( \Psi \) on \( I' \), and the right-hand side a Riemann sum of \( f(\Phi)\psi(\Phi) \) with respect to \( \Phi \) on \( I \). Since by the uniform continuity of \( \Phi \) it follows that \( \max_k |I_k| \to 0 \) implies \( \max_k |I'_k| \to 0 \), by the integrability assumptions, for appropriate partitions \( P \) the left-hand side above tends to \( \int_I f d\Psi \), and the right-hand side to \( \int_I f(\Phi)\psi(\Phi) d\Phi \). Hence the change of variable formula holds.

This observation applies in the following setting. On \( I = I' = [0, 1] \), with \( 0 < \varepsilon, \eta < 1 \), let \( \Phi(x) = x^{1-\varepsilon}, \varphi(x) = (1-\varepsilon)x^{-\varepsilon} \) for \( x \in (0, 1) \), and \( \Psi(y) = y^{1-\eta}, \psi(y) = (1-\eta)y^{-\eta} \) for \( y \in (0, 1) \); \( \varphi \) and \( \psi \) are unbounded. Then, for an integrable function \( f \) on \( I' \), provided that \( f(\Phi)\psi(\Phi) \) is integrable with respect to \( \Phi \) on \( I \), the change of variable formula holds. For \( f \) we may take a continuous function of order \( x^\beta \) near the origin, where \( \beta \geq \varepsilon/(1-\varepsilon) + \eta \).

References

[1] Apostol, T., *Mathematical analysis, Second edition*, Addison-Wesley Publishing Co., Reading, Mass., 1974.

[2] Bagby, R. J., The substitution theorem for Riemann integrals. *Real Anal. Exchange* 27 (1), (2001/02), 309-314.
[3] Bruckner, A. M., Bruckner, J. B. and Thomson, B. S., *Real Analysis, Second Edition*, ClassicRealAnalysis.com., 2008.

[4] Davies, R. O., An elementary proof of the theorem on change of variable in Riemann integration. *Math. Gaz.* 45, (1961), 23-25.

[5] de Oliveira, O. R. B., Change of variable for the Riemann integral on the real line, *Preprint*, (2019)

[6] Hunter, J. K., *Introduction to Analysis*, Undergraduate Lecture Notes, U C Davis, (2010).

[7] Kestelman, H., Change of variable in Riemann integration, *Math. Gaz.* 45, (1961), 17–23.

[8] L´opez Pouso, R., Riemann integration via primitives for a new proof to the change of variable theorem, (2011), *arXiv:1105.5938v1 [math.CA]*

[9] L´opez Pouso, R., Existence and computation of Riemann-Stieltjes integrals through Riemann integrals, (2011), *arXiv:1107.1996v1 [math.CA]*

[10] Navrátil, J. A note on the theorem on change of variable in a Riemann integral (Czech). *Časopis Pěst. Mat.* 106, (1981), 79-83.

[11] Preiss, D. and Uher, J., A remark on the substitution for the Riemann integral. (Czech) *Časopis Pěst. Mat.* 95, (1970), 345-347.

[12] Robbins, H. E., A note on the Riemann integral. *Amer. Math. Monthly* 50 (10), (1943), 617-618, https://www.jstor.org/stable/2303804.

[13] Sarkhel, D. N. and Výborný, R., A change of variables theorem for the Riemann integral. *Real Anal. Exchange* 22 (1), (1996/97), 390-395.

[14] Thomson, B. S., Characterizations of an indefinite Riemann integral. *Real Anal. Exchange* 35 (2), (2009), 487–492.

[15] Thomson, B. S., On Riemann sums. *Real Anal. Exchange* 37 (1), (2011/2012), 221–242.

[16] Torchinsky, A., *Real variables*, Addison-Wesley Publishing Co., Redwood City, CA, 1988.

[17] Torchinsky, A., The change of variable formula for the Riemann integral, *Preprint* (2019)