INROADS

Real Analysis Exchange
Vol. (), , pp. 309–314

Richard J. Bagby, New Mexico State University, Las Cruces, New Mexico 88003. e-mail: rbagby@nmsu.edu

THE SUBSTITUTION THEOREM FOR RIEMANN INTEGRALS

Abstract

We give a completely elementary proof of the strongest possible version of the change-of-variables formula for the Riemann integration of an integrable function f:

$$\int_a^b f \circ G(x)\, g(x)\ dx = \int_{G(a)}^{G(b)} f(u)\ du$$

whenever g is Riemann integrable over $[a, b]$, $G(x) = G(a) + \int_a^x g(t)\ dt$, and f is Riemann integrable over the image of $[a, b]$ under G. Our arguments do not require f to be real-valued; it can take its values in an arbitrary Banach space.

1 Introduction

There are many versions as well as many proofs of the change-of-variables formula

$$\int_a^b f \circ g(x)\, g'(x)\, dx = \int_{g(a)}^{g(b)} f(u)\ du$$

for Riemann integrals, with the simplest based on the chain rule and the fundamental theorem of calculus. A more sophisticated version can be based on the definition of the Riemann integral:

$$\int_a^b f \circ G(x)\, g(x)\, dx = \int_{G(a)}^{G(b)} f(u)\ du$$

Key Words: Riemann integral, substitution, change-of-variable, Banach space
Mathematical Reviews subject classification: Primary 26A42
Received by the editors June 30, 2000
[0]Keywords:
[0]2000 Mathematics Subject Classification: .

whenever g is Riemann integrable over $[a,b]$, $G(x) = G(a) + \int_a^x g(t)\,dt$, and f is Riemann integrable over the image of $[a,b]$ under G. Note that in this version, G need not be piecewise differentiable. The argument is quite straightforward when g is nonnegative, and is easily extended to the case where g has only finitely many changes of sign. But the general case is considerably more subtle, and it doesn't seem to have been published prior to the recent paper of D. N. Sarkhel and R. Výborný [3]. Somewhat earlier, McShane [2] gave a version with f bounded and Lebesgue integrable instead of Riemann integrable; of course, McShane's theorem only proves that $(f \circ G)\,g$ is Lebesgue integrable.

Here we offer a natural proof of the general change-of-variables formula for Riemann integration. We use only the most basic notions of Riemann integration, with no measure-theoretic concepts or continuity properties of Riemann integrable functions, and no mention of differentiation. Unlike the proof given in [3], our proof allows f to have its values in an arbitrary Banach space. However, our version of the substitution theorem can only use the Riemann integrability of f to establish the Riemann integrability of $(f \circ G)\,g$, while their version includes a partial converse. They also prove that f must be Riemann integrable if $(f \circ G)\,g$ is Riemann integrable with f bounded, and give an example showing that f need not even be improperly Riemann integrable unless its boundedness is explicitly assumed.

2 Riemann Integrability

We use a blend of the familiar Riemann and Darboux conditions for integrability: A function f on $[a,b]$ is Riemann integrable if there is a number $\int_a^b f(x)\,dx$ such that for each $\varepsilon > 0$, we can partition $[a,b] = \bigcup_{k=1}^n [x_{k-1}, x_k]$ in such a way that

$$\left| \int_a^b f(x)\,dx - \sum_{k=1}^n f(x_k^*)\,\Delta x_k \right| < \varepsilon$$

for every choice of $x_k^* \in [x_{k-1}, x_k]$; as usual, $\Delta x_k = x_k - x_{k-1}$. Following Graves [1], we note that this definition of Riemann integrability extends immediately to functions taking values in a Banach space X; in that case the integral and each Riemann sum are elements of X and we use the norm $\|\cdot\|$ in X instead of absolute values.

Given a partition $\mathcal{P}$ of $[a,b]$ and an X-valued function f, we have no notion of upper or lower sums for f, but there is a natural substitute for the interval they normally determine, namely, the closed convex hull of the set of all associated Riemann sums. Let's call this set $C(f;\mathcal{P})$, whether f is real-valued or X-valued. When we refine $\mathcal{P}$ to form a new partition $\mathcal{P}'$, every Riemann sum associated with $\mathcal{P}'$ is a convex combination of Riemann sums associated with $\mathcal{P}$, so of course we have $C(f;\mathcal{P}') \subset C(f;\mathcal{P})$. It then becomes a simple manner to prove that an X-valued function f on $[a,b]$ is Riemann integrable if and only if the diameter of $C(f;\mathcal{P})$ can be made arbitrarily small by choosing the partition $\mathcal{P}$ appropriately, with $\int_a^b f(x)\ dx$ the unique element of X that is common to all the sets $C(f;\mathcal{P})$. Then the basic theory of Riemann integration can be developed along standard lines.

In particular, when an X-valued function f is Riemann integrable over $[a,b]$, it is also Riemann integrable over every subinterval of $[a,b]$, with

$$\int_a^b f(x)\ dx = \int_a^c f(x)\ dx + \int_c^b f(x)\ dx \quad \text{for } a \le c \le b.$$

Every Riemann integrable function is bounded, so

$$F(t) = \int_a^t f(x)\ dx$$

defines a Lipschitz continuous X-valued function F on $[a,b]$.

However, we can prove less about Riemann integrable X-valued functions than we can prove in the real-valued case. The reason is that we have only

$$\operatorname{diam} C(f;\mathcal{P}) \le \sum_{k=1}^{n} \operatorname{diam}\{f(x) : x \in [x_{k-1},x_k]\}\,\Delta x_k$$

in general, but in the real-valued case both quantities equal the difference between the upper and lower sums. Graves [1] defines $f : [0,1] \to L^\infty[0,1]$ by calling $f(x)$ the characteristic function of $[0,x]$; he notes that f is Riemann integrable but discontinuous everywhere. Rather strikingly,

$$\sum_{k=1}^{n} \operatorname{diam}\{f(x) : x \in [x_{k-1},x_k]\}\,\Delta x_k = \sum_{k=1}^{n} \Delta x_k = 1$$

for every partition $\mathcal{P}$ of $[0,1]$, while $\operatorname{diam} C(f;\mathcal{P})$ is $\max_k \Delta x_k$.

3 Proof of The Substitution Formula

First we assume that g is positive and Riemann integrable over $[a,b]$ and that f is a Riemann integrable X-valued function on $[G(a), G(b)]$, where

$$G(t) = G(a) + \int_a^t g(x)\ dx.$$

Any partition $\mathcal{P}$ of $[a,b]$ as $\bigcup_{k=1}^{n}[x_{k-1},x_k]$ produces a corresponding partition $\mathcal{P}'$ of $[G(a),G(b)]$ as $\bigcup_{k=1}^{n}[G(x_{k-1}),G(x_k)]$, and every partition of $[G(a),G(b)]$ can be expressed as $\mathcal{P}'$ for some partition $\mathcal{P}$ of $[a,b]$.

In this case we'll establish the elementary bound

$$\left\|\int_{G(a)}^{G(b)} f(u)\ du-\sum_{k=1}^{n} f\circ G(x_k^*)\,g(x_k^*)\,\Delta x_k\right\|\le \operatorname{diam} C(f;\mathcal{P}')+M\operatorname{diam} C(g,\mathcal{P}) \tag{1}$$

for every choice of $x_k^*\in[x_{k-1},x_k]$, where $M=\sup\{\|f(u)\|: u\in[G(a),G(b)]\}$. With our assumptions, we can make both $C(g;\mathcal{P})$ and $C(f;\mathcal{P}')$ have arbitrarily small diameters by choosing $\mathcal{P}$ appropriately, so (1) implies both the Riemann integrability of $(f\circ G)\,g$ and the substitution formula in this special case.

To prove (1), call $u_k=G(x_k)$ and $u_k^*=G(x_k^*)$. Then

$$\left\|\int_{G(a)}^{G(b)} f(u)\ du-\sum_{k=1}^{n} f(u_k^*)\,\Delta u_k\right\|\le \operatorname{diam} C(f;\mathcal{P}'),$$

while

$$\begin{aligned}\left\|\sum_{k=1}^{n} f(u_k^*)\,\Delta u_k-\sum_{k=1}^{n} f\circ G(x_k^*)\,g(x_k^*)\,\Delta x_k\right\| &= \left\|\sum_{k=1}^{n} f(u_k^*)\left[\Delta u_k-g(x_k^*)\,\Delta x_k\right]\right\|\\ &\le \sum_{k=1}^{n}\|f(u_k^*)\|\,|\Delta u_k-g(x_k^*)\,\Delta x_k|\\ &\le M\operatorname{diam} C(g;\mathcal{P}).\end{aligned}$$

Next we note that the substitution formula must also be valid when g is negative and Riemann integrable over $[a,b]$. Essentially the same argument works, except we call $u_k=G(x_{n-k})$ and $u_k^*=G\left(x_{n-k}^*\right)$ to prove

$$\int_a^b f\circ G(x)\,g(x)\ dx=-\int_{G(b)}^{G(a)} f(u)\ du.$$

To prove the substitution formula in the general case, we use the Riemann integrability of g to partition $[a,b]$ in a way that lets us control the subintervals on which g does not keep a fixed sign. Given $\delta>0$, we first choose a partition $\mathcal{P}_0$ for which

$$\operatorname{diam} C(g,\mathcal{P}_0)=\sum \operatorname{diam}\{g(x): x\in[x_{k-1},x_k]\}\,\Delta x_k<\delta^2(b-a).$$

Then we examine the approximation of $\int_{G(a)}^{G(b)} f(u)\ du$ by arbitrary Riemann sums for $(f \circ G)\, g$ associated with a suitable refinement $\mathcal{P}$ of $\mathcal{P}_0$. The key to our analysis is a decomposition $\{1, ..., n\} = D \cup S \cup L$, where $k \in D$ if g is either strictly positive or strictly negative on $[x_{k-1}, x_k]$, $k \in S$ if $k \notin D$ but $|g(x)| < \delta$ at all points in $[x_{k-1}, x_k]$, and $k \in L$ if $k \notin D \cup S$.

For $k \in D$, we know that $(f \circ G)\, g$ is Riemann integrable over $[x_{k-1}, x_k]$, with

$$\int_{x_{k-1}}^{x_k} f \circ G(x)\, g(x)\ dx = \int_{G(x_{k-1})}^{G(x_k)} f(u)\ du.$$

Thus for any $\varepsilon > 0$, we can further partition $[x_{k-1}, x_k]$ as $\bigcup_{j=1}^{n_k} [x_{k,j-1}, x_{k,j}]$ in such a way that

$$\left\| \int_{G(x_{k-1})}^{G(x_k)} f(u)\ du - \sum_{j=1}^{n_k} f \circ G\left(x_{k,j}^*\right) g\left(x_{k,j}^*\right) \Delta x_{k,j} \right\| < \varepsilon \Delta x_k$$

for every choice of $x_{k,j}^* \in [x_{k,j-1}, x_{k,j}]$.

When $k \in S \cup L$, we make no further subdivision of $[x_{k-1}, x_k]$, but to unify the notation we call $n_k = 1$ and $[x_{k,0}, x_{k,1}] = [x_{k-1}, x_k]$. Thus we partition $\mathcal{P}$ of $[a, b]$ in the form

$$[a, b] = \bigcup_{k=1}^{n} \bigcup_{j=1}^{n_k} [x_{k,j-1}, x_{k,j}].$$

Since $\int_{G(a)}^{G(b)} f(u)\ du = \sum_{k=1}^{n} \int_{G(x_{k-1})}^{G(x_k)} f(u)\ du$, we have

$$\left\| \int_{G(a)}^{G(b)} f(u)\ du - \sum_{k=1}^{n} \sum_{j=1}^{n_k} f \circ G\left(x_{k,j}^*\right) g\left(x_{k,j}^*\right) \Delta x_{k,j} \right\| \leq \sum_{k=1}^{n} \|\Phi_k\|,$$

where we've called

$$\Phi_k = \int_{G(x_{k-1})}^{G(x_k)} f(u)\ du - \sum_{j=1}^{n_k} f \circ G\left(x_{k,j}^*\right) g\left(x_{k,j}^*\right) \Delta x_{k,j}.$$

By construction, we have $\|\Phi_k\| < \varepsilon \Delta x_k$ for all $k \in D$. To bound the other terms, we'll need bounds for $\|f\|$ and $|g|$; let's assume that $\|f(u)\| \leq M_1$ for all u and $|g(x)| \leq M_2$ for all x.

For $k \in S$ we have $|g(x)| < \delta$ for all $x \in [x_{k-1}, x_k]$. Consequently, we have $|G(x_k) - G(x_{k-1})| \leq \delta \Delta x_k$ and therefore

$$\|\Phi_k\| \leq \left\| \int_{G(x_{k-1})}^{G(x_k)} f(u)\ du \right\| + \left\| f \circ G\left(x_{k,1}^*\right) g\left(x_{k,1}^*\right) \Delta x_k \right\| \leq 2M_1 \delta \Delta x_k.$$

For $k \in L$ we know only that $|g(x)| \leq M_2$, so we get the cruder bound

$$\|\Phi_k\| \leq 2M_1 M_2 \Delta x_k.$$

Here, however, we know that

$$\operatorname{diam}\{g(x) : x \in [x_{k-1}, x_k]\} \geq \delta,$$

so the definition of $\mathcal{P}_0$ shows that the total length of all these subintervals is less than $\delta(b-a)$.

Taken all together, these bounds show that

$$\begin{aligned}\sum_{k=1}^{n} \|\Phi_k\| &\leq \varepsilon \sum_{k\in D} \Delta x_k + 2M_1\delta \sum_{k\in S} \Delta x_k + 2M_1 M_2 \sum_{k\in L} \Delta x_k \\ &< (\varepsilon + 2M_1\delta + 2M_1 M_2 \delta)(b-a).\end{aligned}$$

We can make this arbitrarily small by choosing δ and ε appropriately, so $(f \circ G)\, g$ is indeed Riemann integrable over $[a,b]$ with

$$\int_a^b f \circ G(x)\, g(x)\, dx = \int_{G(a)}^{G(b)} f(u)\, du.$$

References

[1] L. M. Graves, Riemann integration and Taylor's theorem in general analysis, *Trans. Amer. Math. Soc. 29* (1927), 163-177.

[2] E. J. McShane, *Integration*, Princeton University Press, Princeton, 1944.

[3] D. N. Sarkhel and R. Výborný, A change of variables theorem for the Riemann integral, *Real Analysis Exchange* 22 #1 (1996-1997), 390-395.