THE TEACHING OF MATHEMATICS
2025, Vol. XXVIII, No. 2, pp. 100–106
DOI: 10.57016/TM-LTQK8530

THE GENERAL CHANGE OF VARIABLE FORMULA FOR THE RIEMANN INTEGRAL

Amar Sarić

Abstract. The change of variable theorem for functions that are Riemann integrable, i.e. not obligatory continuous or monotonic, is established based on the definition of the integral and using nothing but the fundamentals of the Riemann theory. Specifically, the Lebesgue criterion for Riemann integrability or more advanced theories of integration are not required.

MathEduc Subject Classification: I55
AMS Subject Classification: 97I50
Key words and phrases: Riemann integral; change of variable.

1. Introduction

Standard textbooks, such as those by Apostol [1], Heuser [3] or Shilov [6], typically present in their treatment of Riemann theory only the trivial version of integration by substitution:

$$\int_{g(a)}^{g(b)} f(x)\,dx = \int_a^b f(g(t))g'(t)\,dt$$

which holds for functions f and g provided that g is continuously differentiable on $[a,b]$ and f continuous on $\{x \mid x = g(t), t \in [a,b]\}$. Many authors, e.g. [3], do not point out at all that there are other formulations, while some, e.g. [1], refer the reader to Lebesgue integration. There are a few exceptions, such as Zorich [7], who proves a version of the theorem for Riemann integrable f and g' under the assumption that g is a monotonic function.[1] Of course, it is well-known that the theorem still holds when f and g' are merely Riemann integrable. Nevertheless, one should keep in mind that it might be surprising to students when encountered for the first time, because for instance $f(g(t))$ does not have to be integrable if g is continuous and $f(t)$ Riemann integrable, see [5]. Kestelman [4] is commonly cited as a reference for the general theorem, but his proof relies on Lebesgue's criterion for Riemann integrability. Davies's proof [2] is more direct, yet unfortunately it has not been included in textbooks either. Thus, as a rule, one leaves question unanswered or otherwise introduces the fact that a function is integrable if and

[1]To be exact, Zorich states the theorem for continuous and positive g' and Riemann integrable f, but the proof is general enough and can easily be extended to Riemann integrable functions.

only if it is continuous everywhere except on a null set, which might make the exposition too abstract for some readers many of whom will be first year students, and even then one does not go beyond just mentioning the theorem. Perhaps one could say it does not matter, but that argument is hard to accept when it comes to a tome of several hundred pages. However, since, as will be shown, we can make the size of the set where g cannot be approximated by monotonic functions in a sense arbitrarily small, combining bounds used by Davies [2] with the ideas from Kestelman [4] and Zorich [7] leads to a rather intuitive, although still not necessarily 'easy', proof.

2. Obtaining the formula in an elementary manner

As it helps to have the special case for monotonic g at hand before proving the main result, the proof will be given in two stages.

THEOREM. *Let $g(t)$ and $h(t)$ be real functions defined on the interval $[a, b]$, such that $h(t)$ is Riemann integrable on $[a, b]$ with*

$$g(t) = \int_a^t h(u)\,du + C,$$

where C is some real constant. Further, let f be another real function that is Riemann integrable on the set $\{x \mid x = g(t), t \in [a, b]\}$. Then, the function $f(g(t))h(t)$ is Riemann integrable on $[a, b]$ and

$$\int_{g(a)}^{g(b)} f(x)\,dx = \int_a^b f(g(t))h(t)\,dt.$$

Proof. In order to make writing of some intervals a bit more convenient, sets of the form $\{x \mid x = g(t), t \in [p, q]\}$ will be denoted by $g([p, q])$. Since $g(t)$ is continuous, it reaches its maximum and minimum on $[a, b]$ as well as all the values in between at least once. Therefore, $g([a, b])$ is itself a compact interval. Also, f, g, and h are, as Riemann integrable functions, bounded, and there is a constant M such that $|f| \leq M$, $|g| \leq M$, and $|h| \leq M$ on $[a, b]$. We start by proving the special case under the additional assumption that either $h(t) > 0$ for every $t \in [a, b]$ or alternatively $h(t) < 0$ for every $t \in [a, b]$. The continuity and monotonicity of g then imply $[g(a), g(b)] = g([a, b])$ or $[g(b), g(a)] = g([a, b])$ depending on whether $h > 0$ or $h < 0$, respectively. Let

$$a = t_0 < t_1 < t_2 < \ldots < t_{n-1} < t_n = b$$

be any partition of $[a, b]$, where $0 < t_i - t_{i-1} < \delta$, $1 \leq i \leq n$ and δ can be made less than any positive number by increasing n. Then, we have

$$g(a) = g(t_0) < g(t_1) < g(t_2) < \ldots < g(t_{n-1}) < g(t_n) = g(b)$$

if $h(t) > 0$ or

$$g(b) = g(t_n) < g(t_{n-1}) < g(t_{n-2}) < \ldots < g(t_1) < g(t_0) = g(a)$$

if $h(t) < 0$, as a partition of $g([a,b])$. Also, denote the division points by

$$\begin{aligned} x_0 &= g(t_0) = g(a) \\ x_1 &= g(t_1) \\ &\dots \\ x_n &= g(t_n) = g(b) \end{aligned}$$

if $h(t) > 0$ and

$$\begin{aligned} x_0 &= g(t_n) = g(b) \\ x_1 &= g(t_{n-1}) \\ &\dots \\ x_n &= g(t_0) = g(a) \end{aligned}$$

if $h(t) < 0$. Using this notation, we have $x_0 < x_1 < x_2 < \ldots < x_{n-1} < x_n$ in both cases. Let us consider the sum

$$\sum_{i=1}^{n} f(g(\xi_i))\,|\kappa_i|(t_i - t_{i-1})$$

with $\xi_i \in [t_{i-1}, t_i]$ and κ_i chosen so that

$$g(t_i) - g(t_{i-1}) = \kappa_i(t_i - t_{i-1})$$

where $\inf_{t\in[t_{i-1},t_i]} h(t) \le \kappa_i \le \sup_{t\in[t_{i-1},t_i]} h(t)$, which is always possible based on the mean value theorem for definite integrals. We obtain

$$(1)\qquad \begin{aligned} &\sum_{i=1}^{n} f(g(\xi_i))\,|\kappa_i|\,(t_i - t_{i-1}) \\ &= \begin{cases} \sum\limits_{i=1}^{n} f(g(\xi_i))\,(g(t_i) - g(t_{i-1})) & \text{if } h(t) > 0 \text{ on } [a,b] \\ \sum\limits_{i=1}^{n} f(g(\xi_i))\,(g(t_{i-1}) - g(t_i)) & \text{if } h(t) < 0 \text{ on } [a,b] \end{cases} \\ &= \sum_{i=1}^{n} f(\rho_i)(x_i - x_{i-1}) \end{aligned}$$

with $\rho_i = g(\xi_i) \in [x_{i-1}, x_i]$ due to monotonicity of g. If $h(t) < 0$, then the intervals in Eq. (1) will be added up from right to left over the partition $x_0 < x_1 < x_2 < \ldots < x_{n-1} < x_n$. Furthermore, we have

$$\begin{aligned} &\left| \sum_{i=1}^{n} f(g(\xi_i))\,|h(\xi_i)|(t_i - t_{i-1}) - \sum_{i=1}^{n} f(\rho_i)(x_i - x_{i-1}) \right| \\ &\quad = \left| \sum_{i=1}^{n} f(g(\xi_i))\,|h(\xi_i)|(t_i - t_{i-1}) - \sum_{i=1}^{n} f(g(\xi_i))\,|\kappa_i|\,(t_i - t_{i-1}) \right| \end{aligned}$$

$$\leq \sum_{i=1}^{n} |f(g(\xi_i))| \, \big||h(\xi_i)| - |\kappa_i|\big| (t_i - t_{i-1}) \leq M \sum_{i=1}^{n} |h(\xi_i) - \kappa_i| \, (t_i - t_{i-1})$$

$$\leq M \sum_{i=1}^{n} (\sup_{\tau \in [t_{i-1}, t_i]} h(\tau) - \inf_{\tau \in [t_{i-1}, t_i]} h(\tau))(t_i - t_{i-1}) \to 0$$

for $\delta \to 0$ under refinement of the partition $t_0 < t_1 < \ldots < t_{n-1} < t_n$ due to integrability of h, see [4] or [7]. In view of $|x_i - x_{i-1}| \leq \max_{t \in [a,b]} |h(t)| \, |t_i - t_{i-1}| \leq M \, |t_i - t_{i-1}|$, the integrability of f on $g([a, b])$ implies

$$\sum_{i=1}^{n} f(g(\xi_i)) \, |h(\xi_i)| \, (t_i - t_{i-1}) \to \sum_{i=1}^{n} f(\rho_i)(x_i - x_{i-1}) \to \begin{cases} \int_{g(a)}^{g(b)} f(x) \, dx, \; h(t) > 0 \\ \int_{g(b)}^{g(a)} f(x) \, dx, \; h(t) < 0 \end{cases}$$

when the largest of the lengths of the intervals $t_i - t_{i-1}$ tends to 0. Since ξ_i are arbitrary values in $[t_{i-1}, t_i]$, it follows that

$$\int_a^b f(g(t)) \, |h(t)| \; dt = \begin{cases} \int_{g(a)}^{g(b)} f(x) \, dx \text{ if } h(t) > 0 \\ \int_{g(b)}^{g(a)} f(x) \, dx \text{ if } h(t) < 0 \end{cases}$$

and this is precisely our claim.

Next, we proceed to the general case when g is not monotonic. Obviously, $g(a)$, $g(b) \in g([a, b])$ and, since g takes all values between $g(a)$ and $g(b)$ due to continuity, we have $[g(a), g(b)] \subset g([a, b])$ or $[g(b), g(a)] \subset g([a, b])$ depending on which of the two values is greater. Again, let

$$a = t_0 < t_1 < t_2 < \ldots < t_{n-1} < t_n = b$$

be a partition with $\xi_i \in [t_{i-1}, t_i]$, $0 < t_i - t_{i-1} < \delta$ and $1 \leq i \leq n$, where δ can be made arbitrary small for large n. By applying the mean value theorem for definite integrals twice, we readily obtain

$$\int_{g(t_{i-1})}^{g(t_i)} f(x) \, dx = \lambda_i \, (g(t_i) - g(t_{i-1})) = \lambda_i \mu_i (t_i - t_{i-1}),$$

where $|\lambda_i| \leq M$ and $\inf_{t \in [t_{i-1}, t_i]} h(t) \leq \mu_i \leq \sup_{t \in [t_{i-1}, t_i]} h(t)$, as well as

$$(2) \quad \left| \int_{g(t_{i-1})}^{g(t_i)} f(x) \, dx - f(g(\xi_i)) h(\xi_i)(t_i - t_{i-1}) \right| \leq \left| \int_{g(t_{i-1})}^{g(t_i)} f(x) \, dx \right| + \\ + |f(g(\xi_i)) h(\xi_i)(t_i - t_{i-1})| \leq (|\lambda_i| \, |\mu_i| + |f(g(\xi_i))| \, |h(\xi_i)|) \, |t_i - t_{i-1}| \, .$$

Let ε be any positive real number. In view of the inequality

$$\sum_{i=1}^{n} \sup_{\tau', \tau'' \in [t_{i-1}, t_i]} |h(\tau') - h(\tau'')| \, (t_i - t_{i-1}) < \varepsilon^2$$

that holds for partitions with a sufficiently small δ, based on integrability of h, the combined length of $[t_{i-1}, t_i]$ on which $\sup_{\tau', \tau'' \in [t_{i-1}, t_i]} |h(\tau') - h(\tau'')| \geq \varepsilon$ must be less than ε. From (2), it trivially follows for every such interval that

$$(3) \qquad \left| \int_{g(t_{i-1})}^{g(t_i)} f(x)\,dx - f(g(\xi_i))h(\xi_i)(t_i - t_{i-1}) \right| \leq 2M^2 \, |t_i - t_{i-1}| \, .$$

We also apply the estimate (2) to $[t_{i-1}, t_i]$ on which $|h(t)| < \varepsilon$ for all t. This yields

$$(4) \qquad \left| \int_{g(t_{i-1})}^{g(t_i)} f(x)\,dx - f(g(\xi_i))h(\xi_i)(t_i - t_{i-1}) \right| \leq 2M\varepsilon \, |t_i - t_{i-1}| \, .$$

Let the set Q be comprised of the remaining $[t_{i-1}, t_i]$, i.e. those intervals on which it holds that $\sup_{\tau', \tau'' \in [t_{i-1}, t_i]} |h(\tau') - h(\tau'')| < \varepsilon$ and at the same time there is some $t \in [t_{i-1}, t_i]$ for which $|h(t)| \geq \varepsilon$. In that case, if $[t_{i-1}, t_i]$ contains a point p such that $h(p) \geq \varepsilon$, then, because of $h(q) \geq h(p) - |h(p) - h(q)| > h(p) - \varepsilon \geq 0$, the value $h(q)$ must be positive in every point $q \in [t_{i-1}, t_i]$. Likewise, if $[t_{i-1}, t_i]$ contains a point p such that $h(p) \leq -\varepsilon$, then in every point $q \in [t_{i-1}, t_i]$ the value $h(q)$ must be negative, since $h(q) \leq h(p) + |h(q) - h(p)| < h(p) + \varepsilon \leq 0$. As $h(t)$ does not change the sign on the individual intervals $[t_{i-1}, t_i]$, the special case of the theorem for monotonic g guarantees that $f(g(t))h(t)$ is integrable on every $[t_{i-1}, t_i] \in Q$ and

$$\int_{g(t_{i-1})}^{g(t_i)} f(x)\,dx = \int_{t_{i-1}}^{t_i} f(g(t))h(t)\,dt.$$

Using the estimates (3) and (4), we can sum over the intervals to obtain

$$(5) \quad \left| \int_{g(a)}^{g(b)} f(x)\,dx - \sum_{i=1}^{n} f(g(\xi_i))h(\xi_i)(t_i - t_{i-1}) \right| \leq 2M^2\varepsilon + 2M\varepsilon(b-a) + \\ + \left| \sum_{\substack{i=1 \\ [t_{i-1}, t_i] \in Q}}^{n} \left(\int_{t_{i-1}}^{t_i} f(g(t))h(t)\,dt - f(g(\xi_i))h(\xi_i)(t_i - t_{i-1}) \right) \right|$$

where the first two terms come from the intervals $\notin Q$ and the last one from those $\in Q$. Now, $f(g(t))h(t)$ is already known to be integrable on the subset of $[a, b]$ that is the union of intervals $\in Q$. Therefore, and by virtue of the additivity of domain for definite integrals, we can repeatedly halve the intervals $\in Q$ until their maximum length decreases below some suitable threshold and the last sum in (5) becomes no greater than, say, ε. This results in

$$(6) \quad \left| \int_{g(a)}^{g(b)} f(x)\,dx - \sum_{i=1}^{n} f(g(\xi_i))h(\xi_i)(t_i - t_{i-1}) \right| \leq (2M(M + b - a) + 1)\varepsilon$$

for some partition P of $[a, b]$ and choice of division points t_i. Note that t_i have been added as needed and n increased accordingly. Also, the contents of Q change in agreement with its definition as the intervals that are being split are replaced by new intervals generated by each split, whereas the intervals $\notin Q$ are not altered in any way. Since ξ_i can be any value $\in [t_{i-1}, t_i]$, and because there exist sequences of values $f(g(\underline{\xi}_{i_k})h(\underline{\xi}_{i_k})$ and $f(g(\overline{\xi}_{i_k})h(\overline{\xi}_{i_k})$, where $\underline{\xi}_{i_k}, \overline{\xi}_{i_k} \in [t_{i-1}, t_i]$, which with increasing k converge to the infimum L_i and supremum U_i of $f(g(t))h(t)$ on each of the individual intervals $[t_{i-1}, t_i]$ that make up P, it follows that the lower and upper Darboux sum of $f(g(t))h(t)$ with respect to P must also both lie within $(2M(M+b-a)+1)\varepsilon$ from $\int_{g(a)}^{g(b)} f(x)dx$ – all we need to do is take one interval $[t_{i-1}, t_i]$ at a time to form $\sum_{i=1}^{n} L_i(t_i - t_{i-1})$ and $\sum_{i=1}^{n} U_i(t_i - t_{i-1})$ while simply relying on the fact that the absolute value is a continuous function. Therefore, as there is such a P for every ε, the lower and upper Darboux integrals of $f(g(t))h(t)$ on $[a, b]$ are equal to $\int_{g(a)}^{g(b)} f(x)\,dx$. ■

Corollary. *Let g be a real differentiable function on the interval $[a, b]$, its derivative g' Riemann integrable on $[a, b]$, and let f be a real function that is Riemann integrable on $\{x \mid x = g(t), t \in [a, b]\}$. Then $f(g(t))g'(t)$ is Riemann integrable on $[a, b]$ and*

$$\int_{g(a)}^{g(b)} f(x)\,dx = \int_a^b f(g(t))g'(t)\,dt.$$

Proof. In view of the fundamental theorem of calculus, the stated conditions clearly imply that $g(t) = \int_a^t g'(u)\,du + g(a)$. ■

3. Conclusion

We have used only the equivalence of Riemann's and Darboux's definitions and the fact that the sum $\sum_{i=1}^{n} \sup_{\tau', \tau'' \in [t_{i-1}, t_i]} |h(\tau') - h(\tau'')|\,(t_i - t_{i-1})$ necessarily goes to zero under refinement of the partitions when h is integrable. The theorem can, therefore, be included in introductory texts on analysis while keeping the exposition well within the grasp of undergraduates, and by doing so one would cover integration more thoroughly. This is arguably the best that we can do when it comes to the Riemann integration.

Acknowledgement. The author would like to thank the University of North Carolina at Charlotte for the continuing support over the years while he was pursuing his graduate studies there.

REFERENCES

[1] T. M. Apostol, *Mathematical Analysis*, Addison-Wesley, 1964.

[2] R. O. Davies, *An elementary proof of the theorem on change of variable in Riemann integration*, The Mathematical Gazette, **45** (1961), 23–25.

[3] H. Heuser, *Lehrbuch der Analysis I*, Teubner, 1990.

[4] H. Kestelman, *Change of variable in Riemann integration*, The Mathematical Gazette, **45** (1961), 17–23.

[5] J. Lu, *Is the composite function integrable?*, The American Mathematical Monthly, **106** (1999), 763–766.

[6] G. E. Shilov, *Elementary Real and Complex Analysis*, Dover, 1996.

[7] V. A. Zorich, *Mathematical Analysis I*. Springer, 2004.

ORCID: 0009-0001-6322-5018

E-mail: asaric@gmail.com

Received: 04.08.2025, *in revised form* 12.09.2025

Accepted: 25.09.2025