Change of Variable or Substitution in Riemann and Lebesgue Integration

By Ng Tze Beng

Because of the fact that not all derived functions are Riemann integrable (see Example 2.2.2.7 on page 139 of *Theorems and Counterexamples in Mathematics* by B.R. Gelbaum and JMH Olmsted), in applying the change of variable formula to Riemann integration we need to be a little careful.

Suppose F and g are differentiable functions, which can be composed to give the composite $F \circ g$. Then $F \circ g$ is differentiable and by the *Chain rule* for differentiation,

$$(F \circ g)'(x) = F'(g(x))\, g'(x) = f(g(x))\, g'(x),$$

where $F'(x) = f(x)$. Thus, $F \circ g$ is an antiderivative of $f(g(x))\, g'(x)$ and F is an antiderivative of f and so we can write

$$\int f(g(x))\, g'(x) dx = F(g(x)) + C.$$

This is a statement about antiderivative and not about Riemann integration. For instance, $f(g(x))\, g'(x)$ need not be Riemann integrable. Here is an example when this is the case.

Example 1. Let $F:[0, 1] \to \mathbf{R}$ be the function as given in Example 2.2.2.7 on page 139 of *Theorems and Counterexamples in Mathematics* by B.R. Gelbaum and JMH Olmsted. F is differentiable on [0, 1] but the derivative $F' = f$ is not continuous on a Cantor set of positive measure in [0, 1] and so f is not Riemann integrable. Now define $g:[-1,0] \to [0,1]$ by $g(x) = x + 1$ for x in $[-1, 0]$. Then g is differentiable on $[-1,0]$ and $g'(x) = 1$ for x in $[-1,0]$. We also have that $f(g(x))$ is not Riemann integrable since it is discontinuous on a set of positive measure, which is a translation of the Cantor set of positive measure in [0,1]. This set is of positive measure because Lebesgue measure is translation invariant. Therefore, $f(g(x))\, g'(x)$ has an antiderivative even though it is not Riemann integrable. The formula

$$\int_{-1}^{0} f(g(x))g'(x)dx = \int_{g(-1)}^{g(0)} f(x)dx$$

does not hold since the integrands on both sides are not Riemann integrable.

Example 2. Even if f is continuous, the formula may not hold because $f(g(x))\, g'(x)$ may not be Riemann integrable simply because $g'(x)$ is not Riemann integrable. We may take $f(x) = \exp(x)$ the exponential function and $g:[0, 1] \to \mathbf{R}$ be the function as given in Example 2.2.2.7 on page 139 of *Theorems and Counterexamples in Mathematics* by B.R. Gelbaum and JMH Olmsted. That is, $g'(x)$ is discontinuous on a Cantor set C of positive measure in [0,1]. Then it is easy to see that $f(g(x))\, g'(x)$ is also discontinuous at every point x in C using the fact that $f \circ g$ is continuous and non-zero and $g'(x) = 0$ on C. Once again, the formula

$$\int_{0}^{1} f(g(x))g'(x)dx = \int_{g(0)}^{g(1)} f(x)dx$$

does not hold simply because the left hand side does not exist.

Now suppose g:[a, b] $\to$ [c, d] is a differentiable function and f : [c, d] $\to$ **R** is a bounded function. A necessary condition for the change of variable formula,

$$\int_a^b f(g(x))g'(x)dx = \int_{g(a)}^{g(b)} f(x)dx, \quad \text{---------------------} \quad \text{(A)}$$

to hold is that

1. $f(g(x))\, g'(x)$ is Riemann integrable on [a, b] and

2. f is Riemann integrable over a domain containing the range of g.

Theorem 1. If g: [a, b] $\to$ [c, d] is a differentiable function and f : [c, d] $\to$ **R** is Riemann integrable and has an antiderivative and if $f(g(x))\, g'(x)$ is Riemann integrable on [a, b], then formula (A) holds.

Proof. Let F be an antiderivative of f. Then $F' = f$. Also we have that $F(g(x))$ is an antiderivative of $f(g(x))\, g'(x)$ by the Chain rule. Then, since $f(g(x))\, g'(x)$ is Riemann integrable, by Darboux Theorem,

$$\int_a^b f(g(x))g'(x)dx = F(g(b)) - F(g(a)) .$$

Since f is Riemann integrable with antiderivative F, again by Darboux Theorem,

$$\int_{g(a)}^{g(b)} f(x)dx = F(g(b)) - F(g(a)) .$$

Therefore, we have $\int_a^b f(g(x))g'(x)dx = \int_{g(a)}^{g(b)} f(x)dx$. This completes the proof.

Remark 1. 1. It is not necessary that the function g be differentiable at the end points a and b for the application of Darboux Theorem. We can replace the condition on g:[a, b] $\to$ **R** by requiring that g be continuous on [a, b] and differentiable on the open interval (a, b). We then further require that the range of g be contained in the domain of f so that the composite function $f \circ g$ can be defined. Then the conclusion of Theorem 1 follows.

2. If the function f is continuous on its domain, then f is Riemann integrable and has an antiderivative given by the Fundamental Theorem of Calculus. Thus, we often replace the condition on f by continuity as in Theorem 2 below.

The following is the usual version of change of variable formula or substitution.

Theorem 2. If g: [a, b] $\to$ [c, d] is a continuous function, differentiable on the open interval (a, b) so that g' : (a, b) $\to$ **R** is continuous and bounded and if f : [c, d] $\to$ **R** is continuous, then formula (A) holds.

Proof. By assumption $f \circ g$ is continuous on [a, b] and so is bounded on (a, b). Since g' : (a, b) $\to$ **R** is continuous and bounded, $f(g(x))\, g'(x)$ or $(f \circ g)\, g'$ is continuous and bounded on (a, b). Therefore, $(f \circ g)\, g'$ is bounded and continuous almost everywhere on [a, b] and so by Lebesgue Theorem, $(f \circ g)\, g'$ is Riemann integrable on [a, b]. Since f is continuous on [c, d], by the Fundamental theorem of Calculus , f

has an antiderivative on [c, d], given by $F(x)=\int_c^x f(t)dt$. Therefore, by Theorem 1 and the remark following Theorem 1, formula (A) holds.

Under what condition can we guarantee the Riemann integrability of $(f\circ g)\, g'$ on [a, b]? Formula (A) requires that f be Riemann integrable on a domain containing the range of g. If we impose sufficient condition on g', we can deduce that $(f\circ g)\, g'$ is Riemann integrable on [a, b]. It is not true in general that if f is Riemann integrable and g is continuous, then $f\circ g$ is Riemann integrable. For a counter example see Example 5 of *Composition and Riemann Integrability*.

Theorem 3. Suppose $f:[c, d]\to \mathbf{R}$ is Riemann integrable and $g: [a, b]\to [c, d]$ is a continuously differentiable strictly increasing function mapping [a, b] onto [c, d]. Then $(f\circ g)\, g'$ is Riemann integrable on [a, b] and formula (A) holds, that is,

$$\int_a^b f(g(x))g'(x)dx=\int_c^d f(x)dx .$$

Remark 2. In theorem 3 we do not require that f has an antiderivative. This includes simple step functions.

Proof of Theorem 3. Since f is Riemann integrable on [c, d], given any $\varepsilon > 0$, there exists a partition W for [c, d],

$$W: c = l_0 < l_1 < \ldots < l_n = d$$

such that

$$U(W, f) - L(W, f) < \varepsilon/2 . \quad \text{--------------------------- (1)}$$

(See Theorem 1 of *Riemann integral and bounded function.)*
Now, since $g: [a, b]\to [c, d]$ is a strictly monotonic increasing bijective map, its inverse is also a strictly monotonic increasing bijection and so

$$g^{-1}W: a = z_0 < z_1 < \ldots < z_n = b ,$$

where $z_i = g^{-1}\, l_i$, i =1,2, …, n, is a partition for [a, b].
Because f is Riemann integrable, f is bounded on [c, d]. That is, there exists a real number $M > 0$ such that $|f(x)| < M$ for all x in [c, d]. Since $g' :[a, b] \to \mathbf{R}$ is continuous on [a, b], g' is uniformly continuous. Therefore, given $\varepsilon > 0$, there exists $\delta > 0$ such that

$$|x - y| < \delta \Rightarrow |g'(x) - g'(y)| < \varepsilon/(2(3M+1)(b-a)) . \quad \text{------------ (2)}$$

Now refine the partition $g^{-1}W$, by adding points if need be, to get a partition,

$$Q: a = x_0 < x_1 < \ldots < x_k = b , \text{ where } k \geq n,$$

with $\| Q \| = \max \{x_i - x_{i-1} : i =1, 2,\ldots, k\} < \delta$. This means $g^{-1}W \subseteq Q$. Then $gQ = P$ is a refinement of W since $W \subseteq g(Q) = P$. Let

$$P: c = y_0 < y_1 < \ldots < y_k = d ,$$

then $y_i = g(x_i)$, i =1, 2,…, k. Therefore, by the definition of upper and lower Riemann sums, (see for example *Riemann integral and bounded function*),

$$L(W, f) \leq L(P, f) \leq U(P, f) \leq U(W, f)$$

and so we have, $U(P, f) - L(P, f) \leq U(W, f) - L(W, f) < \varepsilon/2$, that is,

$$U(P, f) - L(P, f) < \varepsilon/2. \quad \text{---------------------- (3)}$$

This means,

$$\sum_{i=1}^{k} \sup\{ f(x) - f(y) : x, y \in [y_{i-1}, y_i]\}\Delta y_i < \frac{\varepsilon}{2} . \quad \text{---------------------- (4)}$$

Since g is differentiable, by the Mean Value Theorem, there exists $\xi_i \in [x_{i-1}, x_i]$ such that

$\Delta y_i = y_i - y_{i-1} = g(x_i) - g(x_{i-1}) = g'(\xi_i)(x_i - x_{i-1}) = g'(\xi_i)\Delta x_i$ and
$g'(\xi_i) > 0$ for $i = 1, 2, \dots, k$. ---------------------------- (5)

We shall now show that $(f \circ g)\, g'$ satisfies Riemann's condition to conclude that it is Riemann integrable.

$$U(Q, (f\circ g)\, g') - L(Q, (f\circ g)\, g')$$
$$= \sum_{i=1}^{k} \sup\{ f\circ g(x)g'(x) - f\circ g(y)g'(y) : x, y \in [x_{i-1}, x_i]\}\,\Delta x_i . \quad \text{----------------(6)}$$

But for x, y in $[x_{i-1}, x_i]$,

$$|f\circ g(x)\, g'(x) - f\circ g(y)\, g'(y)|$$
$$= |[f\circ g(x) - f\circ g(y)]\,[g'(x) - g'(\xi_i)] + [f\circ g(x) - f\circ g(y)]\, g'(\xi_i) + f\circ g(y)\,[g'(x) - g'(y)]|$$
$$\leq |f(g(x)) - f(g(y))|\,|g'(x) - g'(\xi_i)| + |f(g(x)) - f(g(y))|\, g'(\xi_i) + |f(g(y))|\,|g'(x) - g'(y)|$$
$$\leq 2M\,\varepsilon/(2(3M+1)(b-a)) + |f(g(x)) - f(g(y))|\, g'(\xi_i) + M\,\varepsilon/(2(3M+1)(b-a))$$
$$< \varepsilon/(2(b-a)) + |f(g(x)) - f(g(y))|\, g'(\xi_i).$$

Therefore, for each $i = 1, 2, \dots, k$,

$$\sup\{|f\circ g(x)\, g'(x) - f\circ g(y)\, g'(y)| : x, y \in [x_{i-1}, x_i]\}$$
$$< \varepsilon/(2(b-a)) + \sup\{|f(g(x)) - f(g(y))| : x, y \in [x_{i-1}, x_i]\}\,|g'(\xi_i)|.$$
$$= \varepsilon/(2(b-a)) + \sup\{|f(x) - f(y)| : x, y \in [y_{i-1}, y_i]\}\, g'(\xi_i),$$

since g is a bijection and $g([x_{i-1}, x_i]) = [y_{i-1}, y_i]$, where $y_i = g(x_i)$.
Therefore,

$$U(Q, (f\circ g)\, g') - L(Q, (f\circ g)\, g')$$
$$= \sum_{i=1}^{k} \sup\{|f\circ g(x)g'(x) - f\circ g(y)g'(y)| : x, y \in [x_{i-1}, x_i]\}\,\Delta x_i$$
$$< \sum_{i=1}^{k} \frac{\varepsilon}{2(b-a)}\Delta x_i + \sum_{i=1}^{k} \sup\{|f(x) - f(y)| : x, y \in [x_{i-1}, x_i]\}\, g'(\xi_i)\Delta x_i$$
$$= \frac{\varepsilon}{2} + \sum_{i=1}^{k} \sup\{|f(x) - f(y)| : x, y \in [x_{i-1}, x_i]\}\,\Delta y_i \text{ , by (5) ,}$$
$$< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \quad \text{by (4).}$$

Therefore, $U(Q, (f\circ g)\, g') - L(Q, (f\circ g)\, g') < \varepsilon$ and so $(f\circ g)\, g'$ satisfies the Riemann's condition and so it is Riemann integrable on $[a, b]$. (See Theorem 1 of *Riemann integral and bounded function*.)

We can also show that the upper and lower Riemann integrals of $(f \circ g)\, g'$ are the same to conclude that $(f \circ g)\, g'$ is Riemann integrable. (See Theorem 1 of *Riemann integral and bounded function*.) This will turn out to be a more efficient way to prove the theorem.

We shall use what we have just proved. Note that the upper Riemann sum of $(f\circ g)\, g'$ with respect to the partition Q (as given before with $\|Q\| < \delta$) is defined to be

$$U(Q, (f\circ g)\, g') = \sum_{i=1}^{k} \sup\{ f\circ g(x) g'(x) : x\in[x_{i-1}, x_i]\}\Delta x_i \, .$$

But

$\sup\{ f\circ g(x)\, g'(x) : x \in [x_{i-1}, x_i]\}$

$= \sup\{ f\circ g(x)\, (g'(x) - g'(\xi_i)) + f\circ g(x)\, g'(\xi_i) : x \in [x_{i-1}, x_i]\}$

$\leq \sup\{ M\, \varepsilon/(2(3M+1)(b-a)) + f\circ g(x)\, g'(\xi_i) : x \in [x_{i-1}, x_i]\}$, by (2),

$< \varepsilon/(6(b-a)) + \sup\{ f\circ g(x) : x \in [x_{i-1}, x_i]\} g'(\xi_i)$, since $g'(\xi_i) > 0$,

$= \varepsilon/(6(b-a)) + \sup\{ f(x) : x \in [y_{i-1}, y_i]\} g'(\xi_i)$, since g is a bijection.

Therefore,

$$U(Q, (f\circ g)\, g') < \sum_{i=1}^{k} \frac{\varepsilon}{6(b-a)}\Delta x_i + \sum_{i=1}^{k} \sup\{ f(x) : x\in[y_{i-1}, y_i]\}\, g'(\xi_i)\Delta x_i$$

$$= \frac{\varepsilon}{6} + \sum_{i=1}^{k} \sup\{ f(x) : x\in[y_{i-1}, y_i]\}\,\Delta y_i = \frac{\varepsilon}{6} + U(P, f)$$

$$< \varepsilon/6 + L(P, f) + \varepsilon/2 \quad \text{by (3).}$$

Thus, using the fact that the upper Riemann integral of $(f\circ g)\, g' \leq U(Q, (f\circ g)\, g')$ and that $L(P, f) \leq$ lower Riemann intergral *of* f, we have,

$$U\int_a^b f\circ g(x) g'(x) dx \leq U(Q, (f\circ g) g') < L\int_c^d f(x) dx + \varepsilon ,$$

where $U\int_a^b f\circ g(x) g'(x) dx$ is the upper Riemann integral of $(f\circ g)\, g'$ and $L\int_c^d f(x) dx$ is the lower Riemann integral of f.

Since this is true for arbitrary $\varepsilon > 0$, we have that

$$U\int_a^b f\circ g(x) g'(x) dx \leq L\int_c^d f(x) dx . \text{ --------------- (7)}$$

(More precisely, the upper Riemann integral of $(f\circ g)\, g'$ on $[a, b] \leq$ lower Riemann integral of f on $[c, d]$.)

Now the lower Riemann sum of $(f\circ g) g'$ with respect to the partition Q is defined to be

$$L(Q, (f\circ g) g') = \sum_{i=1}^{k} \inf\{ f\circ g(x) g'(x) : x\in[x_{i-1}, x_i]\}\Delta x_i \, .$$

But $\inf\{ f\circ g(x)\, g'(x) : x \in [x_{i-1}, x_i]\}$

$= \inf\{ f\circ g(x)\, (g'(x) - g'(\xi_i)) + f\circ g(x)\, g'(\xi_i) : x \in [x_{i-1}, x_i]\}$

$\geq \inf\{ - M\, \varepsilon/(2(3M+1)(b-a)) + f\circ g(x)\, g'(\xi_i) : x \in [x_{i-1}, x_i]\}$

$> -\varepsilon/(6(b-a)) + \inf\{ f\circ g(x) : x \in [x_{i-1}, x_i]\} g'(\xi_i)$, since $g'(\xi_i) > 0$,

$= -\varepsilon/(6(b-a)) + \inf\{ f(x) : x \in [y_{i-1}, y_i]\} g'(\xi_i)$, since g is a bijection.

Therefore,

$$L(Q, (f\circ g) g') > \sum_{i=1}^{k} -\frac{\varepsilon}{6(b-a)}\Delta x_i + \sum_{i=1}^{k} \inf\{ f(x) : x\in[y_{i-1}, y_i]\}\, g'(\xi_i)\Delta x_i$$

$$= -\frac{\varepsilon}{6} + \sum_{i=1}^{k} \inf\{ f(x) : x\in[y_{i-1}, y_i]\}\,\Delta y_i = -\frac{\varepsilon}{6} + L(P, f) \, . \text{ ------------- (8)}$$

Now,

$$U\int_c^d f(x) dx \leq U(P, f) < L(P, f) + \frac{\varepsilon}{2}, \text{ by (3),}$$

$$< L(Q,(f\circ g)g') + \frac{\varepsilon}{6} + \frac{\varepsilon}{2}, \text{ by (8) above,}$$

$$< L\int_a^b f\circ g(x)g'(x)dx + \varepsilon,$$

using the fact that the upper Riemann integral of $f \le U(P, f)$ and that $L(Q, (f\circ g)g')$ $\le$ the lower Riemann integral of $(f\circ g)\, g'$.

Since this is true for any ε > 0, $U\int_c^d f(x)dx \le L\int_a^b f\circ g(x)g'(x)dx$. (More precisely, the upper Riemann integral of f on $[c, d]$ ≤ lower Riemann integral of $(f\circ g)\, g'$ on $[a, b]$.) Therefore, because lower Riemann integral is always less than or equal to the upper Riemann integral, we have $L\int_a^b f\circ g(x)g'(x)dx \le U\int_a^b f\circ g(x)g'(x)dx$. This then with (7) and the above inequality gives us,

$$U\int_c^d f(x)dx \le L\int_a^b f\circ g(x)g'(x)dx \le U\int_a^b f\circ g(x)g'(x)dx \le L\int_c^d f(x)dx.$$

Since f is Riemann integrable on $[c, d]$, $U\int_c^d f(x)dx = L\int_c^d f(x)dx$ and so

$$L\int_a^b f\circ g(x)g'(x)dx = U\int_a^b f\circ g(x)g'(x)dx = \int_c^d f(x)dx.$$

Hence $(f\circ g)\, g'$ is Riemann integrable on $[a, b]$ and

$$\int_a^b f\circ g(x)g'(x)dx = U\int_a^b f\circ g(x)g'(x)dx = \int_c^d f(x)dx.$$

This completes the proof.

The case when g is strictly decreasing is given as follows.

Theorem 4. Suppose $f : [c, d] \to \mathbf{R}$ is Riemann integrable and $g: [a, b] \to [c, d]$ is a continuously differentiable strictly decreasing function mapping $[a, b]$ onto $[c, d]$. Then $(f\circ g)\, g'$ is Riemann integrable on $[a, b]$ and

$$\int_a^b f(g(x))g'(x)dx = \int_d^c f(x)dx = \int_c^d \left(-f(x)\right)dx.$$

The proof is exactly the same except that we use the function $-f$ instead of f *and* noting that this time round, $g'(\xi_i) < 0$ *and in the equation*

$$\Delta y_i = g'(\xi_i)\Delta x_i,$$

the points y_i *are oriented in the opposite direction, that is,*

$$d = y_0 > y_1 > \ldots > y_k = c,$$

where $y_i = g(x_i)$, $i = 1, 2, \ldots, k$. Care should be exercised when taking supremum or infimum, use $-g'(\xi_i) > 0$. For instance,

$\sup\{ f\circ g(x)\, g'(x): x \in [x_{i-1}, x_i]\}$

$= \sup\{ f\circ g(x)\,(g'(x) - g'(\xi_i)) + f\circ g(x)\, g'(\xi_i): x \in [x_{i-1}, x_i]\}$

$\le \sup\{ M\,\varepsilon/(2(3M+1)(b-a)) + f\circ g(x)\, g'(\xi_i): x \in [x_{i-1}, x_i]\}$

$< \varepsilon/(6(b-a)) + \sup\{-f\circ g(x) : x \in [x_{i-1}, x_i]\}(-g'(\xi_i))$, since $-g'(\xi_i) > 0$.

There is a version of Theorem 1 for Lebesgue integrals. We shall state the result as follows. Note that it is a consequence of the following theorem.

Theorem 5. Suppose $f : [a, b] \to \mathbf{R}$ is differentiable on $[a, b]$ and its derivative f' is bounded, that is, there exists a number $K \ge 0$ such that for all x in $[a, b]$, $|f'(x)| \le K$. Then (1) f is *absolutely continuous* and

(2) the derived function f': $[a, b] \to \mathbf{R}$ is Lebesgue integrable, and for any x in $[a, b]$, the Lebesgue integral,

$$\int_a^x f'(t)dt = f(x) - f(a).$$

This is a well known result from Lebesgue integration theory. Part of the theorem is a consequence of the characterisation of functions satisfying the conclusion of Darboux Theorem or "Fundamental Theorem of Calculus" with Riemann integral replaced by Lebesgue integral and Riemann integrability replaced by Lebesgue integrability in terms of *absolute continuity*. This is a concept, which says that $f : [a, b] \to \mathbf{R}$ is absolutely continuous if for any $\varepsilon > 0$, there exists some $\delta > 0$ such that for any finite number of disjoint open intervals, (a_1, b_1), (a_2, b_2),…, $(a_n\ b_n)$ in $[a, b]$, such that $\sum_{i=1}^{n}(b_i - a_i) < \delta$, then we have $\sum_{i=1}^{n}\left| f(b_i) - f(a_i)\right| < \varepsilon$. Obviously, absolute continuity implies continuity. If f is differentiable and has a bounded derivative, it follows from the Mean Value Theorem that f is absolutely continuous on $[a, b]$. Thus, Theorem 5 is just the characterisation of functions satisfying the conclusion of the "Fundamental Theorem of Calculus" for Lebesgue integrals, with bounded derivative guaranteeing the absolute continuity of f.

Theorem 6. If g: $[a, b] \to [c, d]$ is a differentiable function such that its derivative is bounded and if $f : [c, d] \to \mathbf{R}$ is a bounded function, which has an antiderivative, then we have the following equality for Lebesgue integrals.

$$\int_a^b f(g(x))g'(x)dx = \int_{g(a)}^{g(b)} f(x)dx.$$

Proof. Suppose F is an antiderivative of f. Then since its derivative f is bounded on $[c, d]$, F is absolutely continuous. By Theorem 5, $F' = f$ is Lebesgue integrable. Also the composite function $F \circ g$ is differentiable and its derivative, by the *Chain Rule*, is given by $(f \circ g)g'$, which is bounded since both $f \circ g$ and g' are bounded. Therefore, by Theorem 5,

$$\int_a^b f(g(x))g'(x)dx = \int_a^b \left(F \circ g\right)'(x)dx = F(g(b) - F(g(a)) = \int_{g(a)}^{g(b)} f(x)dx.$$

This proves the theorem.

Remark 3. For the functions f and g as given in Example 2, by Theorem 6, formula (A) for Lebesgue integrals holds. Though $(f \circ g)g'$ *there* is not *Riemann integrable,* it is *nevertheless Lebesgue integrable.*

There is a version of Theorem 3 in terms of Lebesgue integral. The difficulty with the proof lies mainly in showing that the *Chain Rule* is true almost everywhere unlike Theorem 6 when the Chain Rule does apply easily.

We state the theorem and then we shall describe the problem in detail.

Theorem 7. If g: $[a, b] \to [c, d]$ is a monotone increasing (not necessarily strictly increasing) absolutely continuous function mapping $[a, b]$ onto $[c, d]$ and $f : [c, d] \to$

R is a Lebesgue integrable function, then we have the following equality for Lebesgue integrals.

$$\int_a^b f(g(x))g'(x)dx = \int_c^d f(x)dx = \int_{g(a)}^{g(b)} f(x)dx .$$

Now we set up the functions involved in the Theorem. Since f is Lebesgue integrable on $[c, d]$, we can define the indefinite integral on $[c, d]$ by

$$F(x) = \int_c^x f(t)dt ,$$

for any x in $[c, d]$. Then F: $[c, d] \to \mathbf{R}$ is an absolutely continuous function, which is differentiable almost everywhere on $[c, d]$ and $F' = f$ almost everywhere on $[c, d]$. Since $g :[a, b] \to [c, d]$ is monotonic and absolutely continuous, the composite function $F \circ g :[a, b] \to \mathbf{R}$ is also absolutely continuous. We shall phrase this result in the following proposition.

Proposition 8. Suppose $g :[a, b] \to [c, d]$ is monotonic and absolutely continuous and F: $[c, d] \to \mathbf{R}$ is absolutely continuous. Then $F \circ g :[a, b] \to \mathbf{R}$ is also absolutely continuous.

Proof. We shall assume that g is monotonic increasing and absolutely continuous. Since F is absolutely continuous, given any $\varepsilon > 0$, there exists $\delta_1 > 0$ such that for any finite disjoint open intervals, (c_1, d_1), (c_2, d_2),…, (c_n, d_n) in $[c, d]$,

$$\sum_{i=1}^{n} (d_i - c_i) < \delta_1 \Rightarrow \sum_{i=1}^{n} \left| F(d_i) - F(c_i) \right| < \varepsilon . \quad \text{-----------------------} \quad (9)$$

Thus, for such a $\delta_1 > 0$, since g is absolutely continuous, there exists $\delta_2 > 0$ such that for any finite disjoint open intervals, (a_1, b_1), (a_2, b_2),…, (a_k , b_k) in $[a, b]$, with $\sum_{i=1}^{k} (b_i - a_i) < \delta_2$, we have $\sum_{i=1}^{k} \left| g(b_i) - g(a_i) \right| < \delta_1$. But then since g is monotonic increasing, the interval $(g(a_i), g(b_i))$ is either empty when $g(a_i) = g(b_i)$ or a nontrivial open interval and the intervals $(g(a_1), g(b_1))$, $(g(a_2), g(b_2))$,…, $(g(a_k), g(b_k))$ are then pairwise disjoint in $[c, d]$. Thus, it follows from (9) that

$$\sum_{i=1}^{k} (b_i - a_i) < \delta_2 \Rightarrow \sum_{i=1}^{k} \left| F \circ g(b_i) - F \circ g(a_i) \right| = \sum_{i=1}^{k} \left| F(g(b_i)) - F(g(a_i)) \right| < \varepsilon .$$

If g is monotonic decreasing, then the intervals, $(g(b_1), g(a_1))$, $(g(b_2), g(a_2))$,…, $(g(b_k), g(a_k))$, are pairwise disjoint open intervals in $[c, d]$. Therefore, it follows from (9) that we can also conclude that given any $\varepsilon > 0$, there exists $\delta_2 > 0$ such that

$$\sum_{i=1}^{k} (b_i - a_i) < \delta_2 \Rightarrow \sum_{i=1}^{k} \left| F \circ g(b_i) - F \circ g(a_i) \right| = \sum_{i=1}^{k} \left| F(g(b_i)) - F(g(a_i)) \right| < \varepsilon .$$

Therefore, $F \circ g$ is absolutely continuous. This completes the proof.

Before we embark on the proof of Theorem 7, we shall establish some results that we shall use in the proof. We shall be using the Vitali Covering Theorem, a very useful theorem for our purpose. The theorem is stated later for reference.

From now on we shall use the term “measure” interchangeably with “Lebesgue outer measure”.

Proposition 9. Suppose g: $[a, b] \to \mathbf{R}$ is an absolutely continuous function. If E is a subset of $[a, b]$ of measure zero, then its image $g(E)$ is also of measure zero.

Proof. Since g is absolutely continuous, given any $\varepsilon > 0$, there exists some $\delta > 0$ such that for any finite disjoint open intervals, (a_1, b_1), (a_2, b_2),…, (a_n, b_n), with

$$\sum_{i=1}^{n}(b_i - a_i) < \delta \text{, we have } \sum_{i=1}^{n}|g(b_i) - g(a_i)| < \varepsilon . \quad \text{----------------} \quad (10)$$

At this point, it is useful to note that when proving statement about set being of measure zero, it is equivalent to proving the same of the same set minus a set of countable number of points, since such a set is of measure zero. Very often for technical reason, we may also add countable or finite number of points to it. One of the use of this device involves the fact that any open set in **R** is at most the union of countable open intervals, which are pairwise disjoint. What we may actually need is the union of countable closed intervals, which are non-overlapping in the sense that any two sets in this collection can have at most one point in common. Since the measure of $\{g(a), g(b)\}$ is zero, we may assume that $E \subseteq (a, b)$. Because E is of measure zero, for this $\delta > 0$, there exists an open set I such that $E \subseteq I \subseteq (a, \mathrm{b})$ and the measure $m(I - E) < \delta$. Since $m(E) = 0$, $m(I) < \delta$. Now I , being an open set, is the union of countable number of disjoint open intervals, $\{ I_i ,\ i = 1, \dots\}$, each of finite length. The number of open intervals I_i may be finite. Let $I_i = (a_i , b_i)$ and its closure $J_i = [a_i , b_i]$. Then $E \subseteq J \subseteq [a, b]$, where $J = \cup J_i$. Hence, $g(E) \subseteq g(J) = \bigcup_i g(J_i)$. Since g is continuous and each J_i is closed and bounded, by the Extreme Value Theorem, each $g(J_i) = [c_i,\ d_i]$ is also a closed and bounded interval, where c_i is the absolute minimum of g on J_i and d_i is the absolute maximum of g on J_i and there exist x_i , y_i in J_i such that $g(x_i) = c_i$, $g(y_i) = d_i$, either $x_i \le y_i$ or $y_i \le x_i$. Denote the interval $[x_i, y_i]$ or $[y_i , x_i]$ by K_i . Then $K_i \subseteq J_i$ and $g(K_i) = g(J_i)$ for each i. Thus, for any integer n,

$$\sum_{i=1}^{n}|y_i - x_i| = m\left(\bigcup_{i=1}^{n}K_i\right) \le m\left(\bigcup_{i=1}^{n}J_i\right) = \sum_{i=1}^{n}|b_i - b_i| = \sum_{i=1}^{n}m(I_i) \le m(I) < \delta .$$

This implies by (10) that

$$m\left(\bigcup_{i=1}^{n}g(J_i)\right) \le \sum_{i=1}^{n}m\big(g(J_i)\big) = \sum_{i=1}^{n}m\big(g(K_i)\big) = \sum_{i=1}^{n}|g(y_i) - g(x_i)| < \varepsilon .$$

Since this is true for any integer n, $m\left(\bigcup_{i=1}^{\infty}g(J_i)\right) \le \varepsilon$. Hence,

$m\big(g(E)\big) \le m\left(\bigcup_{i=1}^{\infty}g(J_i)\right) \le \varepsilon$. Since this is true for any $\varepsilon > 0$, $m(g(E)) = 0$. This completes the proof.

The next proposition will be useful for redefining the function f.

Proposition 10. Suppose g :$[a, b] \to \mathbf{R}$ is absolutely continuous and monotonic increasing. If E is the set $\{x \in [a, b] : g'(x) = 0\}$, then its image $g(E)$ is of measure zero.

Proof. If E is of measure zero, then the measure of $g(E)$ is zero by Proposition 9. So we now consider the case when the measure of E is greater than 0. Since the measure of $\{g(a), g(b)\}$ is zero, we may assume that $E \subseteq (a, b)$. (If need be, just remove both points a, b from E.) Let $g([a, b]) = [c, d]$.

For each x in E, $g'(x) = 0$ and so given any $\varepsilon > 0$, there are arbitrary small intervals $[x, x + h]$ such that $g(x+h)-g(x)=\left|g(x+h)-g(x)\right|<\varepsilon h$, where $h > 0$. We may assume, without loss of generality, that each $[x, x + h]$ is contained in (a, b). Then these arbitrary small intervals form a Vitali covering of E. As g is absolutely continuous, choose $\delta > 0$ for the given $\varepsilon > 0$ satisfying (10) in the definition of absolute continuity. Then by the Vitali Covering Theorem, there is a finite disjoint collection of these intervals, $I_i = [x_i, y_i]$, $i = 1,2,\ldots,n$, in $[a, b]$, such that the measure of $E-\bigcup_{i=1}^{n} I_i$ is less than δ. That is to say, I_i , $i = 1,2,\ldots, n$ covers all of E except for the subset $E-\bigcup_{i=1}^{n} I_i$ of E, of measure less than δ. We order these intervals so that

$$a < x_1 < y_1 \le x_2 < y_2 \le x_3 < \ldots \le x_n < y_n < b.$$

Then the measure

$$m\left(g\left(E\cap\bigcup_{i=1}^{n} I_i\right)\right)\le m\left(\bigcup_{i=1}^{n} g\left(I_i\right)\right)=\sum_{i=1}^{n}\left(g(y_i)-g(x_i)\right)<\varepsilon\sum_{i=1}^{n}\left(y_i-x_i\right)\le\varepsilon(b-a).$$

Now, $(a,b)-\bigcup_{i=1}^{n} I_i$ is an open set containing $E'=E-\bigcup_{i=1}^{n} I_i$. Let $k = m(E')$. Then since $k=m\left(E-\bigcup_{i=1}^{n} I_i\right)<\delta$, there exists an open set J containing E' such that the measure $m(J-E')<\frac{1}{2}\left(\delta-k\right)$. Thus,

$$m(J)=m(J-E')+m(J\cap E')<\frac{1}{2}(\delta-k)+k=\frac{1}{2}(\delta+k)<\delta.$$

We may assume that $J\subseteq(a,b)-\bigcup_{i=1}^{n} I_i$. Since J is open, it is a union of countable pairwise disjoint intervals $J_i = (a_i, b_i)$ $i = 1,2,\ldots, N$. Here N may be infinity. Then $\sum_{i=1}^{N}\left|b_i-a\right|_i=\sum_{i=1}^{N} m\left(J_i\right)=m(J)<\delta$. Hence,

$$m\left(g\left(E-\bigcup_{i=1}^{n} I_i\right)\right)=m\big(g\left(E'\right)\big)\le m\big(g\left(J\right)\big)=m\left(g\left(\bigcup_{i=1}^{N} J_i\right)\right)$$

$$\le\sum_{i=1}^{N} m(g(J_i))=\sum_{i=1}^{N}\left(g(b_i)-g(a_i)\right)\le\varepsilon.$$

This gives us:

$$m(g(E))=m\left(g\left(E\cap\bigcup_{i=1}^{n} I_i\right)\right)+m\left(g\left(E-\bigcup_{i=1}^{n} I_i\right)\right)<\varepsilon(b-a+1).$$

Since ε is arbitrary, this implies that $m(g(E)) = 0$.

Remark.

Actually, a stronger version of Proposition 10 is true.

Proposition 10*
Suppose g is defined and finite on $[a, b]$. Suppose $E = \{x \in [a, b]: g$ is differentiable at x and $g'(x) = 0\}$. Then $m(g(E)) = 0$.
(See Theorem 3 of my article, *Functions Having Finite Derivatives, Bounded Variation, Absolute Continuity, the Banach Zarecki Theorem and de La Vallée Poussin's Theorem*.)

We state also a partial converse to Proposition 10*.

Theorem 11*. Suppose g has derivatives (finite or infinite) on a set E with $m(g(E)) = 0$. Then $g' = 0$ almost everywhere on E.

(This is Theorem 2, in my article *Change of Variables Theorems*)

We shall prove the assertion we make previously, namely the following proposition.

Proposition 11. Suppose $g :[a, b] \to \mathbf{R}$ is an absolutely continuous and monotonic increasing function. Suppose the range of g is $[c, d]$ and E is a subset of $[c, d]$ of measure zero. Let $H = \{x \in [a, b] : g'(x) \neq 0\}$. Then the measure of $g^{-1}(E) \cap H$ is zero. Thus, $g' = 0$ almost everywhere on $g^{-1}(E)$.

Proof. Let $H_n = \{ x \in [a, b] : |g'(x)| > 1/n\}$. Then $H = \bigcup_{n=1}^{\infty} H_n$. We shall show that the measure of $g^{-1}(E) \cap H_n = E_n$ is zero for each positive integer n.
Suppose on the contrary that the measure $m(g^{-1}(E) \cap H_n) = m(E_n) = k > 0$. Since $g(E_n) \subseteq E$ and E is of measure zero, the measure $m(g(E_n))$ is 0. Thus, given any $\varepsilon' > 0$, there exists an open set G' containing $g(E_n)$ such that $m(G') < \varepsilon'$. Then $G = g^{-1}(G')$ is an open set containing E_n. For each x in E_n, there exists arbitrary small interval $[x, x+h]$ such that

$$\left|\frac{g(x+h)-g(x)}{h}\right| > \frac{1}{n} \quad \text{or} \quad |g(x+h)-g(x)| > \frac{1}{n}h\,, \quad \text{------------} \quad (11)$$

where $h > 0$.
We may assume that these intervals are in G. Therefore, this collection forms a Vitali covering for E_n. Then by the Vitali Covering Theorem, given any $\varepsilon > 0$, there exists a finite collection of disjoint intervals, I_i, $i = 1,2,\ldots, N$, so that $I = I_1 \cup I_2 \cup \ldots \cup I_N$ covers a subset of E_n of outer measure $> k - \varepsilon$. *But if we write* $I_i = [x_i, x_i + h_i]$, then (11) implies that

$$\sum_{i=1}^{N} |g(x_i + h_i) - g(x_i)| > \frac{1}{n}\sum_{i=1}^{N} h_i = \frac{1}{n} m(I) > \frac{1}{n}(k - \varepsilon) \quad .$$

But since $g(I) \subseteq G'$ *and g is monotonic increasing ,*

$$\sum_{i=1}^{N} |g(x_i + h_i) - g(x_i)| = m(g(I)) < m(G') < \varepsilon' \,.$$

So, if we choose $\varepsilon' = \dfrac{k}{2n}$ and $\varepsilon = k/2$, we would obtain a contradiction. This means $m(E_n) = 0$. Since $g^{-1}(E) \cap H_n = E_n$, $g^{-1}(E)\cap H = \bigcup_{n=1}^{\infty} g^{-1}(E)\cap H_n = \bigcup_{n=1}^{\infty} E_n$.

As the measure $m\left(g^{-1}(E)\cap H\right) = m\left(\bigcup_{n=1}^{\infty} E_n\right) \le \sum_{i=1}^{\infty} m(E_n)$ and $m(E_n) = 0$ for each integer n, $m\left(g^{-1}(E)\cap H\right) = 0$. This completes the proof.

Similar result when g is a decreasing function also holds.

Proposition 12. Suppose g: $[a, b] \to \mathbf{R}$ is an absolutely continuous and monotonic decreasing function. If $E = \{x \in [a, b] : g'(x) = 0\}$, then its image $g(E)$ is of measure zero.

Proof. If g is decreasing, then $-g$ is increasing. Then Proposition 10 says that $-g(E)$ is of measure zero. Since measure is invariant under reflection, $g(E)$ is also of zero measure.

(Proposition 10* supersedes Proposition 12.)

Proposition 13. Suppose g:$[a, b] \to \mathbf{R}$ is an absolutely continuous and monotonic decreasing function. Suppose the range of g is $[c, d]$ and E is a subset of $[c, d]$ of measure zero. Let $H = \{x \in [a, b] : g'(x) \neq 0\}$. Then the measure of $g^{-1}(E) \cap H$ is zero.

Proof. If g maps onto $[c, d]$, then $-g$ maps onto $[-d, -c]$. Suppose E is a subset of $[c, d]$ of zero measure, then $-E$ is a subset of $[-d, -c]$ of zero measure. Then by Proposition 11, $(-g)^{-1}(-E)\cap H'$, where $H' = \{ x \in [a, b] : (-g)'(x) \neq 0\}$, is of measure zero. But $H' = H$ and $(-g)^{-1}(-E) = g^{-1}(E)$ and so the measure of $g^{-1}(E) \cap H$ is zero.

Proposition 13*. Suppose g :$[a, b] \to \mathbf{R}$ is finite on $[a, b]$ and the range of g is $[c, d]$ and E is a subset of $[c, d]$ of measure zero. Let $H = \{x \in [a, b] : g$ is differentiable (finite or infinitely) at x and $g'(x) \neq 0\}$. Then the measure of $g^{-1}(E) \cap H$ is zero. Thus, $g'(x) = 0$ almost everywhere on $g^{-1}(E)$.

Proof.
If $H = \varnothing$, then we have nothing to prove.
So we now assume that $H \neq \varnothing$.
If $g^{-1}(E)\cap H = \varnothing$, then we have nothing to prove.
Suppose now that $g^{-1}(E)\cap H \neq \varnothing$.
By hypothesis, g is differentiable (finite or infinitely) on $g^{-1}(E)\cap H$. Moreover, since $g\left(g^{-1}(E)\cap H\right) \subseteq E$ and E is of measure 0, $m\left(g\left(g^{-1}(E)\cap H\right)\right) = 0$. Then

by Theorem 11* (Theorem 2 of my article, *Change of Variables Theorems*), $g' = 0$ almost everywhere on $g^{-1}(E)\cap H$. But as there does not exists an x in $g^{-1}(E)\cap H$ such that $g'(x)=0$, $m\left(g^{-1}(E)\cap H\right)=0$.

Proposition 14. Suppose $g:[a, b] \to \mathbf{R}$ is absolutely continuous and monotonic. Suppose the range of g is $[c, d]$ and F: $[c, d] \to \mathbf{R}$ is an absolutely continuous function. Let $E = \{x \in [a, b] : (F \circ g)'(x) \neq 0 \text{ and } g'(x) = 0\}$. Then the measure of E is zero.

Proof. Let $E_n = \{ x \in [a, b] : |(F \circ g)'(x)| > 1/n \text{ and } g'(x) = 0 \}$. Then $E = \bigcup_{n=1}^{\infty} E_n$. We shall show that the measure of E_n is zero for each positive integer n. We may assume as usual that $E_n \subseteq (a, b)$.
Suppose on the contrary that the measure $m(E_n) = k > 0$.
Now we use the condition that the derivative of g is zero at every point of E_n . For any x in E_n , since $g'(x) = 0$, given any $K > 0$ there exists arbitrary small $h > 0$ such that

$$|g(x+h)-g(x)| < Kh \ . \quad \text{-------------------} \ (12)$$

We are going to choose a suitable K to give a contradiction. The choice of K will depend on k, n and the absolute continuity of F.
Given any $\varepsilon' > 0$, since $m(E_n) = k$, there exists an open set G containing E_n such that the measure of G, $m(G) < k + \varepsilon'$. We shall choose our ε' carefully.
For each x in E_n , since $(F \circ g)'(x) \neq 0$, there exists arbitrary small interval $[x, x+h]$ such that

$$\left|\frac{F(g(x+h))-F(g(x))}{h}\right| > \frac{1}{n} \text{ or}$$

$$|F(g(x+h))-F(g(x))| > \frac{1}{n}h, \quad \text{-----------------------------} \ (13)$$

where $h > 0$.
We may assume that these intervals are in G and also that (12) and (13) hold simultaneously for these arbitrary small intervals $[x, x+h]$. Therefore, this collection forms a Vitali covering for E_n . Then by the Vitali Covering Theorem, there exists a finite number of pairwise disjoint intervals, I_i , $i = 1,2,\ldots, N$, so that $I = I_1 \cup I_2 \cup \ldots \cup I_N$ covers a subset of E_n of outer measure > $k - \varepsilon'$. *But if we write* $I_i = [x_i, x_i + h_i]$, then (12) implies that

$$\sum_{i=1}^{N}|g(x_i+h_i)-g(x_i)| < K\sum_{i=1}^{N}h_i = Km(I) < Km(G) < K(k+\varepsilon') \quad \text{--------} \ (14)$$

and (13) gives

$$\sum_{i=1}^{N}\left|F\left(g(x_i+h_i)\right)-F\left(g(x_i)\right)\right| > \frac{1}{n}\sum_{i=1}^{N}h_i = \frac{1}{n}m(I) > \frac{1}{n}(k-\varepsilon') \ . \ \text{---------} \ (15)$$

Take $\varepsilon' = k/2$, Then $\frac{1}{n}\left(k-\varepsilon'\right)=\frac{k}{2n}$.
By definition of the absolute continuity of F, there exists $\delta > 0$ such that for any finite number of disjoint open intervals, (a_1, b_1), (a_2, b_2),…, (a_M , b_M) in $[c, d]$,

$$\sum_{i=1}^{M}(b_i - a_i) < \delta \Rightarrow \sum_{i=1}^{M}|F(b_i) - F(a_i)| < \frac{k}{2n} \quad . \quad \text{-----------------------------} \quad (16)$$

Now choose $K > 0$ so that $K(k + \varepsilon') = K(3k/2) < \delta$. Then (14) implies that $\sum_{i=1}^{N}|g(x_i + h_i) - g(x_i)| < K(k + \varepsilon') < \delta$. Therefore, since g is monotonic, if g is increasing $(g(x_i), g(x_i+h_i))$, $i = 1, 2, \ldots, n$ are pairwise disjoint or degenerate (i.e., $g(x_i) = g(x_i+h_i)$), and if g is decreasing, $(g(x_i+h_i), g(x_i))$, $i = 1, 2, \ldots, n$ are pairwise disjoint or degenerate. Therefore, by (16), we have

$$\sum_{i=1}^{N}\left|F\left(g(x_i + h_i)\right) - F\left(g(x_i)\right)\right| < \frac{k}{2n}.$$

But this contradicts (15). Hence, $m(E_n) = 0$ for each positive integer n. Consequently, as the measure $m(E) = m\left(\bigcup_{k=1}^{\infty} E_k\right) \le \sum_{k=1}^{\infty} m(E_k)$, $m(E) = 0$. This completes the proof.

A stronger result than Proposition 14 holds as follows:

Proposition 14*. Suppose $g : [a, b] \to \mathbf{R}$ is a function of bounded variation. Suppose the range of g is $[c, d]$ and $F: [c, d] \to \mathbf{R}$ is an absolutely continuous function. Let $E = \{x \in [a, b] : (F \circ g)'(x) \neq 0 \text{ and } g'(x) = 0\}$. Then the measure of E is zero.

Proof.
Since g is of bounded variation, g has finite derivatives almost everywhere on $[a, b]$. Let $A = \{x \in [a,b] : g \text{ is differentiable at } x \text{ and } g'(x) = 0\}$. By Proposition 10*, $m\left(g(A)\right) = 0$. Since F is absolutely continuous, F is an N-function, i.e., F maps sets of measure zero to sets of measure zero. Therefore, $m\left(F\left(g(A)\right)\right) = 0$. Let $B = F \circ g(A) = F\left(g(A)\right)$. Then B is of measure zero. Let

$$H = \left\{x \in [a,b] : F \circ g \text{ is differentiable at } x \text{ (finite or infinitely) and } \left(F \circ g\right)' \neq 0\right\}.$$

Then by Proposition 13*, the measure of $\left(F \circ g\right)^{-1}(B) \cap H$ is zero.
Since $E \subseteq \left(F \circ g\right)^{-1}(B) \cap H$, measure of E is zero,

Remark. The use of Proposition 10* does not require g to be of bounded variation, indeed it only requires g to be a finite function, that is, g has finite values. We can drop the condition in Proposition 14* that g be a function of bounded variation and replace it by just g is a function finite on $[a, b]$.

Before we proceed with the proof of Theorem 7, we state the Vitali Covering Theorem.

Definition 15. For a subset E of $\mathbf{R}$, a collection $\mathcal{C}$ of intervals is said to cover E in the sense of Vitali, if given x in E and any $\varepsilon > 0$, there is an interval J in $\mathcal{C}$ with $x \in J$ and $0 < \mu(J) < \varepsilon$, where $\mu(J) = m(J)$ is the length of J.

Theorem 16. Vitali Covering Theorem.
Suppose $E \subseteq \mathbf{R}$ has finite Lebesgue outer measure and is covered in the sense of Vitali by a class $\mathscr{C}$ of intervals. Then there is a countable disjoint subclass $J \subseteq \mathscr{C}$ such that the outer measure $m\left(E - \bigcup\{I : I \in J\}\right) = 0$.

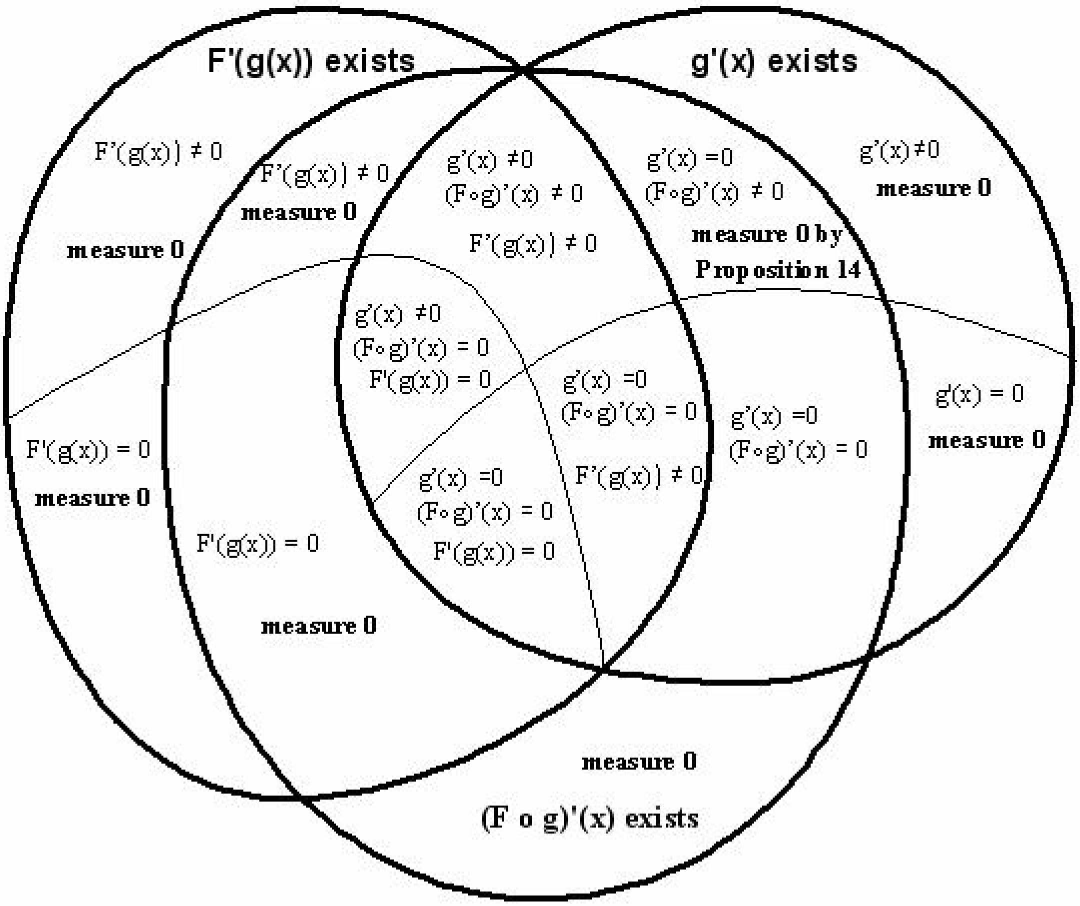
The following more useful form of the theorem provides a finite covering that covers enough of the set E for use.

Corollary 17. Suppose $E \subseteq \mathbf{R}$ has finite Lebesgue outer measure and is covered in the sense of Vitali by a class $\mathscr{C}$ of intervals. Then given any $\varepsilon > 0$, there is a finite set J_1, J_2, …, J_n of disjoint intervals of $\mathscr{C}$ such that

$$m\left(E - \bigcup_{i=1}^{n} J_i\right) < \varepsilon .$$

For the proof of Theorem 16 see page 225 of "Introduction to Measure and Integration" by S.J. Taylor or page 98 of "Real Analysis" by H.L. Royden.

We now summarize the idea of the proof of Theorem 7 in the following Venn diagram.



Proof of Theorem 7.

If F: $[c, d] \to \mathbf{R}$ is given by $F(x) = \int_c^x f(t)dt$ and g :$[a, b] \to [c, d]$ is monotonic increasing, surjective and absolutely continuous, then the composite $F \circ g$:$[a, b] \to \mathbf{R}$, by Proposition 8, is also absolutely continuous and consequently, the Lebesgue integral,

$$\int_a^b (F \circ g)'(x)dx = F \circ g(b) - F \circ g(a)$$

$$= F(g(b)) - F(g(a)) = F(d) - F(c) = \int_{g(a)}^{g(b)} f(x)dx ,$$

since g(a) = c and g(b) = d.
Then we may ask the question:
When is $(F\circ g)'(x) = F'(g(x))\, g'(x) = f(g(x))g'(x)$ almost everywhere on $[a, b]$?
It turns out that this is not an easy question. To prove the theorem, we need to answer this question in the affirmative and the Venn diagram above illustrates the situation as we proceed.

Because $F\circ g$ is absolutely continuous and hence of bounded variation on $[a, b]$, $F\circ g$ is differentiable almost everywhere on $[a, b]$. Since g is monotonic increasing, g is also differentiable almost everywhere on $[a, b]$. Thus, it is enough to consider subset K of $[a, b]$, where both $(F\circ g)'(x)$ and $g'(x)$ exist for every x in K, because the complement of K in $[a, b]$ is of measure zero.
If x is in K, either $g'(x) = 0$ or $g'(x) \neq 0$.
Now suppose x is in K and $g'(x) \neq 0$.
As $(F\circ g)'(x)$ exists, $F'(g(x))$ exists. This is seen as follows. For $h \neq 0$, if $g(x + h) - g(x) \neq 0$, then

$$\frac{F(g(x+h)) - F(g(x))}{h} = \frac{F(g(x+h)) - F(g(x))}{g(x+h) - g(x)} \cdot \frac{g(x+h) - g(x)}{h} .$$

Consider the function $G(h) = \dfrac{g(x+h) - g(x)}{h}$ for $h \neq 0$.

Since $\lim_{h\to 0} G(h) = g'(x) \neq 0$, there exists $\delta > 0$ such that $0 < |h| < \delta$ implies $G(h) \neq 0$ and so $k(h) = g(x + h) - g(x) \neq 0$. Thus, for $0 < |h| < \delta$, we have

$$\frac{\frac{F(g(x+h))-F(g(x))}{h}}{\frac{g(x+h)-g(x)}{h}} = \frac{F(g(x+h)) - F(g(x))}{g(x+h) - g(x)} .$$

Hence, $$\lim_{h\to 0} \frac{F(g(x+h)) - F(g(x))}{g(x+h) - g(x)} = \frac{\lim_{h\to 0} \frac{F(g(x+h))-F(g(x))}{h}}{\lim_{h\to 0} \frac{g(x+h)-g(x)}{h}} = \frac{(F\circ g)'(x)}{g'(x)} = L .$$

That means given $\varepsilon > 0$, there exists $\delta > 0$ such that $0 < |h| < \delta \Rightarrow k(h) \neq 0$ and

$$\left| \frac{F(g(x+h)) - F(g(x))}{g(x+h) - g(x)} - L \right| < \varepsilon .$$

Since k is continuous on the interval $(-\delta, \delta)$, $k((-\delta, \delta))$ is an interval containing 0.

On the interval $(-\delta, \delta)$, k is monotonic increasing and since $k(h) = 0$, if and only if , $h = 0$, $k((-\delta, \delta))$ contains an open interval containing 0. Thus, there exists $\delta' > 0$ such that $(-\delta', \delta') \subseteq k((-\delta, \delta))$. Therefore, this implies that for every p in $(-\delta', \delta')$, there exists a h in $(-\delta, \delta)$ such that $k(h) = p$. Therefore, $0 < |p| < \delta'$ implies that

$$\left|\frac{F(g(x)+p)-F(g(x))}{p}-L\right| = \left|\frac{F(g(x)+k(h))-F(g(x))}{k(h)}-L\right|$$
$$= \left|\frac{F(g(x+h))-F(g(x))}{g(x+h)-g(x)}-L\right| < \varepsilon \ .$$

This shows that $\lim\limits_{p\to 0}\dfrac{F(g(x)+p)-F(g(x))}{p}$ exists and so F is differentiable at $g(x)$.
This means that if x is in K and $g'(x) \neq 0$, then the chain rule $(F\circ g)'(x) = F'(g(x))g'(x)$ holds.
Now consider the case x in K *and* $g'(x) = 0$
If $F'(g(x))$ exists, then we have the chain rule and so $(F\circ g)'(x) = F'(g(x))\, g'(x) = 0$.
On the other hand, if $g'(x) = 0$ and $(F\circ g)'(x) = 0$, the equality still holds and there is no contribution to the Lebesgue integrals on both sides.
It remains to check the case when $g'(x) = 0$ and $(F\circ g)'(x) \neq 0$. Of course, this means $g'(x) = 0$ and $(F\circ g)'(x) \neq 0$ and $F'(g(x))$ does not exist for its existence would imply that $(F\circ g)'(x) = 0$. By Proposition 14, the set $\{x \in [a, b]: g'(x) = 0, (F\circ g)'(x) \neq 0$ (therefore $F'(g(x))$ does not exist)$\}$ has measure zero. We shall also show that the set $\{x \in [a, b]: g'(x) \neq 0$, and $F'(g(x)) \neq f(g(x))\}$ has measure zero. If $F'(g(x)) \neq f(g(x))$, then $g(x)$ belongs to a set of measure zero, since $F' = f$ almost everywhere on $[c, d]$. Therefore, the set $\{x \in [a, b]: g'(x) \neq 0$, and $F'(g(x)) \neq f(g(x))\} \subseteq g^{-1}(E) \cap H$, where E is a set of measure zero and $H=\{x \in [a, b]: g'(x) \neq 0\}$. Thus, by Proposition 11, the measure of $g^{-1}(E) \cap H$ is zero and so $\{x \in [a, b]: g'(x) \neq 0$, and $F'(g(x)) \neq f(g(x))\}$ has measure zero. Thus, we have,

$$\left(F\circ g\right)'(x) = \begin{cases} \left(F\circ g\right)'(x), & \text{when } g'(x)\neq 0, \\ 0, & \text{when } g'(x)=0 \end{cases} \quad \text{almost everywhere on } [a, b],$$

$$= \begin{cases} F'(g(x))g'(x), & \text{when } g'(x)\neq 0, \\ 0, & \text{when } g'(x)=0 \end{cases} \quad \text{almost everywhere on } [a, b],$$

$$= \begin{cases} F'(g(x))g'(x), & \text{when } g'(x)\neq 0, \\ F'(g(x))g'(x), & \text{when } g'(x)=0 \text{ and } F'(g(x)) \text{ exists}, \\ 0, & \text{when } g'(x)=0 \text{ and } F'(g(x)) \text{ does not exist} \end{cases} \quad \text{almost everywhere on } [a, b],$$

$$= \begin{cases} f(g(x))g'(x), & \text{when } g'(x)\neq 0, \\ 0, & \text{when } g'(x)=0 \end{cases} \quad \text{almost everywhere on } [a, b],$$

$$= f(g(x))g'(x) \ \text{ almost everywhere on } [a, b].$$

Therefore, $\int_c^d f(x)dx = F(d)-F(c) = \int_a^b \left(F\circ g\right)'(x)dx = \int_a^b f(g(x))g'(x)dx$.

From the above proof itself, we can easily construct a similar proof for the case when g is an absolutely continuous decreasing function. Hence, we state the following theorem for record.

Theorem 18. If g: $[a, b] \to [c, d]$ is a monotone decreasing (not necessarily strictly decreasing) absolutely continuous function mapping $[a, b]$ onto $[c, d]$ and $f : [c, d] \to \mathbf{R}$ is a Lebesgue integrable function, then we have the following equality for Lebesgue integrals.

$$\int_a^b f(g(x))g'(x)dx = -\int_c^d f(x)dx = \int_{g(a)}^{g(b)} f(x)dx .$$

Remark 4. It is clear that, since any Riemann integrable function is also Lebesgue integrable and the two integrals are the same, Theorem 3 is a special case of Theorem 7 because any g: $[a, b] \to [c, d]$ which is monotonic and continuosly differentiable on $[a, b]$ has bounded derivative and so is absolutely continuous. However, we may not deduce from Theorem 7 that with the condition of Theorem 3, $f(g(x))g'(x)$ is Riemann integrable. We may conclude that $f(g(x))g'(x)$ is Lebesgue integrable but need an extra effort to show that it is Riemann integrable by showing that it is indeed continuous except on a set of measure zero in $[a, b]$. We may replace the requirement that g be strictly increasing by just increasing, courtesy of Theorem 7.
For simplicity and for the reader who may not be familiar with measure theory, we choose to present the two proofs. It is of course possible to modify the proof of Theorem 3 by not requiring the strict monotonicity on g. Likewise, Theorem 4 is a special case of Theorem 18.

Remark 5. It is also possible to replace the function f by another function equal to f almost everywhere. We assume the hypothesis of Theorem 7, that is, g: $[a, b] \to [c, d]$ is a monotone increasing (not necessarily strictly increasing) absolutely continuous function mapping $[a, b]$ onto $[c, d]$ and $f : [c, d] \to \mathbf{R}$ is a Lebesgue integrable function. By Proposition 12, the measure $m(g(E))$ is zero, where $E = \{x \in [a, b] : g'(x) = 0\}$. Now we can define $f_E : [c, d] \to \mathbf{R}$ by $f_E(y) = f(y)$, if y is not in $g(E)$ and $f_E(y) = 0$, if y is in $g(E)$. Then $f_E = f$ almost everywhere on $[c, d]$ and we have
$\int_c^d f_E(x)dx = \int_a^b f_E(g(x))g'(x)dx = \int_a^b f(g(x))g'(x)dx$.

The conclusion of Theorem 7 need not be true if we drop the monotonicity condition on g as illustrated in the following example.

Example 3. Let g:$[0, 1] \to \mathbf{R}$ be defined by $g(x) = \begin{cases} x^2 \sin^2\left(\frac{\pi}{2x}\right), & x > 0 \\ 0, & x = 0 \end{cases}$ and let f :$[0, 1] \to \mathbf{R}$ be defined by $f(x) = \begin{cases} \frac{1}{2\sqrt{x}}, & x > 0 \\ 0, & x = 0 \end{cases}$. Then f is not bounded on [0, 1].
The function f is Lebesgue integrable and the integral is given by the improper Riemann integral F :$[0, 1] \to \mathbf{R}$ defined by $F(x) = \int_0^x f(t)dt = \sqrt{x}$. Then the

composite function $F \circ g$ is not absolutely continuous even though both F and g are.

This is because $F \circ g(x) = \begin{cases} \left| x\sin\left(\frac{\pi}{2x}\right)\right|, & x > 0 \\ 0, & x = 0 \end{cases}$ and is obviously not of bounded variation on [0, 1] and so cannot be absolutely continuous on [0, 1].

Note that F is differentiable at x for $x > 0$ and $F'(x) = f(x) = \dfrac{1}{2\sqrt{x}}$ for $x > 0$ but not differentiable at $x = 0$.

Also note that g is differentiable on $[0, \infty)$ and

$$g'(x) = \begin{cases} 2x\sin^2\left(\frac{\pi}{2x}\right) - \pi\sin\left(\frac{\pi}{2x}\right)\cos\left(\frac{\pi}{2x}\right), & x > 0 \\ 0, & x = 0 \end{cases}.$$

Since F is differentiable on (0, 1] and g is differentiable on [0, 1], $(F \circ g)'(x) = f(g(x))g'(x)$ almost everywhere on [0, 1]. Note that

$$f(g(x)) = \begin{cases} \dfrac{1}{2x\left|\sin\left(\frac{\pi}{2x}\right)\right|}, & x > 0 \text{ and } x \neq \dfrac{1}{2k}\text{ , integer } k \geq 1 \\ 0, & x = 0 \text{ or } x = \dfrac{1}{2k}\text{ , integer } k \geq 1 \end{cases}$$ so that

$$f(g(x))g'(x) = \begin{cases} \left|\sin\left(\frac{\pi}{2x}\right)\right| - \dfrac{\pi}{2}\cdot\dfrac{\sin\left(\frac{\pi}{2x}\right)}{\left|\sin\left(\frac{\pi}{2x}\right)\right|}\cdot\dfrac{\cos\left(\frac{\pi}{2x}\right)}{x}, x > 0 \text{ and } x \neq \dfrac{1}{2k}\text{ , integer } k \geq 1 \\ 0, x = 0 \text{ or } x = \dfrac{1}{2k}, \text{ integer } k \geq 1 \end{cases}.$$

For $x = \dfrac{1}{2k}$, integer $k \geq 1$,

$$\lim_{h\to 0}\frac{F\circ g(\frac{1}{2k}+h) - F\circ g(\frac{1}{2k})}{h} = \lim_{h\to 0}\frac{\left|(\frac{1}{2k}+h)\sin\left(\dfrac{\pi}{2(\frac{1}{2k}+h)}\right)\right|}{h}.$$

Observe that $\dfrac{\pi}{2(\frac{1}{2k}+h)} = \dfrac{k\pi}{1+kh}$ and for sufficiently small $h > 0$,

$(k-1)\pi\ \dfrac{\pi}{2(\frac{1}{2k}+h)} = \dfrac{k\pi}{1+kh} < k\pi$ and for $h < 0$ and sufficiently small $|h|$, i.,e, when

$0 < 1+kh < 1,\ (k+1)\pi > \dfrac{\pi}{2(\frac{1}{2k}+h)} = \dfrac{k\pi}{1+kh} > k\pi$.

Thus, for k even,

$$\lim_{h\to 0^+}\frac{\left|(\frac{1}{2k}+h)\sin\left(\dfrac{\pi}{2(\frac{1}{2k}+h)}\right)\right|}{h} = -\lim_{h\to 0^+}\frac{(\frac{1}{2k}+h)\sin\left(\dfrac{\pi}{2(\frac{1}{2k}+h)}\right)}{h}$$

$$= -\lim_{h\to 0^+} \frac{\sin\left(\frac{\pi}{2(\frac{1}{2k}+h)}\right) - (\frac{1}{2k}+h)\cos\left(\frac{\pi}{2(\frac{1}{2k}+h)}\right)\cdot\frac{\pi}{2}\frac{1}{(\frac{1}{2k}+h)^2}}{1} = \frac{\pi}{k},$$

and

$$\lim_{h\to 0^-} \frac{\left|(\frac{1}{2k}+h)\sin\left(\frac{\pi}{2(\frac{1}{2k}+h)}\right)\right|}{h} = \lim_{h\to 0^-} \frac{(\frac{1}{2k}+h)\sin\left(\frac{\pi}{2(\frac{1}{2k}+h)}\right)}{h}$$

$$= \lim_{h\to 0^-} \frac{\sin\left(\frac{\pi}{2(\frac{1}{2k}+h)}\right) - (\frac{1}{2k}+h)\cos\left(\frac{\pi}{2(\frac{1}{2k}+h)}\right)\cdot\frac{\pi}{2}\frac{1}{(\frac{1}{2k}+h)^2}}{1} = -\frac{\pi}{k},$$

Thus, as the left and right limits are not the same, $F \circ g$ is not differentiable at x for $x = \frac{1}{4i}$, $i \geq 1$.

For k odd,

$$\lim_{h\to 0^+} \frac{\left|(\frac{1}{2k}+h)\sin\left(\frac{\pi}{2(\frac{1}{2k}+h)}\right)\right|}{h} = \lim_{h\to 0^+} \frac{(\frac{1}{2k}+h)\sin\left(\frac{\pi}{2(\frac{1}{2k}+h)}\right)}{h}$$

$$= \lim_{h\to 0^+} \frac{\sin\left(\frac{\pi}{2(\frac{1}{2k}+h)}\right) - (\frac{1}{2k}+h)\cos\left(\frac{\pi}{2(\frac{1}{2k}+h)}\right)\cdot\frac{\pi}{2}\frac{1}{(\frac{1}{2k}+h)^2}}{1} = -\frac{\pi}{k}$$

and

$$\lim_{h\to 0^-} \frac{\left|(\frac{1}{2k}+h)\sin\left(\frac{\pi}{2(\frac{1}{2k}+h)}\right)\right|}{h} = -\lim_{h\to 0^-} \frac{(\frac{1}{2k}+h)\sin\left(\frac{\pi}{2(\frac{1}{2k}+h)}\right)}{h}$$

$$= -\lim_{h\to 0^+} \frac{\sin\left(\frac{\pi}{2(\frac{1}{2k}+h)}\right) - (\frac{1}{2k}+h)\cos\left(\frac{\pi}{2(\frac{1}{2k}+h)}\right)\cdot\frac{\pi}{2}\frac{1}{(\frac{1}{2k}+h)^2}}{1} = \frac{\pi}{k}.$$

It follows that $F \circ g$ is not differentiable at x for $x = \frac{1}{4i+2}$, $i \geq 0$. Hence, $F \circ g$ is not differentiable at x for $x = \frac{1}{2k}$, $k \geq 1$. Thus, $F \circ g$ is differentiable on [0, 1] except for a denumerable subset in [0, 1] and $(F \circ g)'(x) = f(g(x))g'(x)$ for $x \neq \frac{1}{2k}$, $k \geq 1$ or $x \neq 0$.

Note that $F \circ g$ is continuous on [0, 1]. Therefore, since $F \circ g$ is not absolutely continuous on [0, 1], by Corollary 2 of my article, *"When is a continuous functions on a closed and bounded interval be of bounded variation, absolutely continuous? The answer and application to generalized change of variable for Lebesgue integral"*, $\left(F \circ g\right)'$ cannot be Lebesgue integrable on [0,1]. We may verify this fact directly as follows.

$$\left| f(g(x))g'(x)\right| \geq \frac{\pi}{2} \cdot \frac{\left|\cos\left(\frac{\pi}{2x}\right)\right|}{x} - \left|\sin\left(\frac{\pi}{2x}\right)\right| \quad \text{for } x > 0 \text{ and } x \neq \frac{1}{2k}, \text{ integer } k \geq 1.$$

Now $\frac{\pi}{2} \cdot \frac{\left|\cos\left(\frac{\pi}{2x}\right)\right|}{x}$ is not Lebesgue integrable on [0, 1]. Observe that

$$\int_t^1 \frac{\left|\cos\left(\frac{\pi}{2x}\right)\right|}{x} dx = \int_{\frac{\pi}{2}}^{\frac{\pi}{2}\frac{1}{t}} \frac{\left|\cos\left(y\right)\right|}{y} dx \text{ and so}$$

$$\int_0^1 \frac{\left|\cos\left(\frac{\pi}{2x}\right)\right|}{x} dx = \lim_{t \to 0^+} \int_t^1 \frac{\left|\cos\left(\frac{\pi}{2x}\right)\right|}{x} dx = \lim_{t \to \infty} \int_{\frac{\pi}{2}}^{t} \frac{\left|\cos\left(y\right)\right|}{y} dy .$$

Note that

$$\int_{\frac{\pi}{2}k}^{\frac{\pi}{2}(k+1)} \frac{\left|\cos\left(y\right)\right|}{y} dy \geq \frac{2}{\pi(k+1)} \int_{\frac{\pi}{2}k}^{\frac{\pi}{2}(k+1)} \left|\cos\left(y\right)\right| dy = \frac{2}{\pi(k+1)}$$

and so $\int_{\frac{\pi}{2}}^{\frac{\pi}{2}(N+1)} \frac{\left|\cos\left(y\right)\right|}{y} dy \geq \sum_{k=1}^{N} \frac{2}{\pi(k+1)}$.

Therefore, since $\sum_{k=1}^{\infty} \frac{2}{\pi(k+1)} = \infty$, $\int_0^1 \frac{\left|\cos\left(\frac{\pi}{2x}\right)\right|}{x} dx = \infty$. Since $\int_0^1 \left|\sin\left(\frac{\pi}{2x}\right)\right| dx \leq 1$, it follows that $\int_0^1 \left| f(g(x))g'(x)\right| dx = \infty$. This means $\left(F \circ g\right)'$ is not Lebesgue integrable on [0, 1].

However, $\left(F \circ g\right)'$ is improperly Riemann integrable. This is because

$G(x) = \frac{\pi}{2} \cdot \frac{\sin\left(\frac{\pi}{2x}\right)}{\left|\sin\left(\frac{\pi}{2x}\right)\right|} \cdot \frac{\cos\left(\frac{\pi}{2x}\right)}{x}, x > 0 \text{ and } x \neq \frac{1}{2k}$, integer $k \geq 1$ is improperly Riemann integrable on [0, 1]. We deduce this as follows. It is easy to observe that

$G(x)$ changes sign in the intervals,

$$\left(\frac{1}{4k+1}, \frac{1}{4k}\right), \left(\frac{1}{4k+2}, \frac{1}{4k+1}\right), \left(\frac{1}{4k+3}, \frac{1}{4k+2}\right), \left(\frac{1}{4k+4}, \frac{1}{4k+3}\right) .$$

Now $\left|\int_{\frac{1}{k+1}}^{\frac{1}{k}} G(x) dx\right| = \left|\frac{\pi}{2} \int_{\frac{1}{k+1}}^{\frac{1}{k}} \frac{\sin\left(\frac{\pi}{2x}\right)}{\left|\sin\left(\frac{\pi}{2x}\right)\right|} \cdot \frac{\cos\left(\frac{\pi}{2x}\right)}{x} dx\right| = \left|\frac{\pi}{2} \int_{\frac{1}{k+1}}^{\frac{1}{k}} \frac{\cos\left(\frac{\pi}{2x}\right)}{x} dx\right|$

$$= \left|\frac{\pi}{2} \int_k^{k+1} \frac{\cos\left(\frac{\pi}{2} y\right)}{y} dy\right| \leq \frac{\pi}{2k} \int_k^{k+1} \left|\cos\left(\frac{\pi}{2} y\right)\right| dy = \frac{1}{k} .$$

Note also that $\left|\frac{\pi}{2}\int_k^{k+1}\frac{\cos\left(\frac{\pi}{2}y\right)}{y}dy\right| \geq \frac{1}{k+1}$. This means

$$\frac{1}{k+1} \leq \left|\int_{\frac{1}{k+1}}^{\frac{1}{k}} G(x)dx\right| \leq \frac{1}{k}.$$

Hence $\left|\int_{\frac{1}{k+1}}^{\frac{1}{k}} G(x)dx\right| \to 0$ as $k \to \infty$. Moreover, the sequence $\left|\int_{\frac{1}{n+1}}^{\frac{1}{n}} G(x)dx\right|$ is a decreasing sequence. Therefore, by the Alternating Series Test, $\int_0^1 G(x)dx$ is convergent. Consequently,

$$f(g(x))g'(x) = \left|\sin\left(\frac{\pi}{2x}\right)\right| - G(x), x > 0 \text{ and } x \neq \frac{1}{2k} \text{ , integer } k \geq 1$$

is improperly Riemann integrable on [0, 1].

Now, for any 0< t < 1, $F \circ g$ is continuous on [t, 1] and differentiable on [0, 1] except on a denumerable subset. Moreover, on [t, 1],

$$\left(F \circ g\right)'(x) = f(g(x))g'(x) = \left|\sin\left(\frac{\pi}{2x}\right)\right| - G(x), x \geq t \text{ and } x \neq \frac{1}{2k} \text{ , integer } k \geq 1.$$

Now $G(x)$ is Lebesgue integrable on [t, 1], because there exists an integer N such that $0 < \frac{1}{2N} < t$ and $\int_{\frac{1}{2N}}^1 |G(x)|dx = \sum_{k=1}^{2N-1}\int_{\frac{1}{k+1}}^{\frac{1}{k}} |G(x)|dx \leq \sum_{k=1}^{2N-1}\frac{1}{k} < \infty$. Therefore, $\left(F \circ g\right)'$ is Lebesgue integrable on [t, 1] as $\left|\sin\left(\frac{\pi}{2x}\right)\right|$ is Lebesgue integrable on [t, 1]. Therefore, by Corollary 2 of *"When is a continuous functions on a closed and bounded interval be of bounded variation, absolutely continuous? The answer and application to generalized change of variable for Lebesgue integral",* $F \circ g$ is absolutely continuous on [t, 1]. Therefore,

$$\int_t^1 \left(F \circ g\right)'(x)dx = F(g(1)) - F(g(t)) = 1 - \sqrt{t^2\sin^2\left(\frac{\pi}{2t}\right)} = 1 - t\left|\sin\left(\frac{\pi}{2t}\right)\right|$$

$$= \int_{g(t)}^{g(1)} f(x)dx. \qquad \text{------------------------ (1)}$$

Note that we may deduce that $\left(F \circ g\right)'$ is Lebesgue integrable on [t, 1] by simply noting that $\left(F \circ g\right)'$ is Riemann integrable on [t, 1] since it is continuous except on a denumerable set in [t, 1] and that it is bounded on [t, 1].

Therefore, the integral on both sides of (1) are Riemann integrals.
Hence the improper Riemann integral,

$$\int_0^1 \left(F \circ g\right)'(x)dx = \lim_{t \to 0^+}\int_t^1 \left(F \circ g\right)'(x)dx = \lim_{t \to 0^+}\int_{g(t)}^1 f(x)dx$$

$$= \lim_{y \to 0^+}\int_y^1 f(x)dx = \int_0^1 f(x)dx.$$

But $\int_t^1 (F\circ g)'(x)dx = \int_t^1 f(g(x))\cdot g'(x)dx$ and so we have the improper Riemann integral,

$$\int_0^1 f(g(x))\cdot g'(x)dx = \int_0^1 f(x)dx = \int_{g(0)}^{g(1)} f(x)dx .$$

The change of variable formula for Lebesgue integral in this case fails simply because $(F\circ g)'$ is not Lebesgue integrable on [0, 1]. However, $(F\circ g)'$ is improperly Riemann integrable on [0, 1] and the formula holds with the integrals on both sides taken to be the improper Riemann integral.

Thus, in view of the fact that *f* is unbounded in the above example, we consider function *f* , which is bounded and Lebesgue integrable on the interval [*a*, *b*]. We can now obtain the change of variable formula even if *g* is not monotonic but still absolutely continuous. This will involve some subtle technical results in measure theory, functions of bounded variation or absolute continuity. We shall state a weaker version of the theorem, which is mostly what we need.

Theorem 19. Suppose *g*: [*a*, *b*] → **R** is an absolutely continuous function and *f* : [*c*, *d*] → **R** is a continuous function such that the range of *g* is contained in [*c*, *d*]. Then we have the following equality for Lebesgue integrals.

$$\int_a^b f(g(x))g'(x)dx = \int_{g(a)}^{g(b)} f(x)dx .$$

We shall need the following technical result.

Proposition 20. Suppose *f* : [*c*, *d*] → **R** is a bounded and Lebesgue integrable function. Define *F* : [*c*, *d*] → **R** by $F(x) = \int_c^x f(t)dt$. Then *F* satisfies a Lipschitz condition, that is, there is a constant *K* such that $|F(y) - F(x)| \le K| y - x|$ for all *x* and *y* in [*c*, *d*]. Equivalently, if *F* is absolutely continuous and *F′* is bounded, then *F* satisfies a Lipschitz condition.

Proof. If *F* is absolutely continuous on [*c*, *d*], then $F(x) = \int_c^x F'(t)dt + F(c)$ for all *x* in [*c*, *d*]. Since *F′* is bounded almost everywhere in [*c*, *d*], there exists a *K* such that $|F'(x)| < K$ almost everywhere on [*c*, *d*]. Then for any *y* > *x* in [*c*, *d*] $\int_x^y |F'(t)|dt \le \int_x^k K = K|y-x|$. Thus, for any *y* > *x*,

$$|F(y)-F(x)| = \left|\int_c^y F'(t)dt - \int_c^x F'(t)dt\right| = \left|\int_x^y F'(t)dt\right| \le K|y-x| .$$

Therefore, *F* satisfies a Lipschitz condition with constant *K*.

A crucial step in proving Theorem 19 is to show that the composite function $F\circ g$ is absolutely continuous when *g* is absolutely continuous and *F* is also absolutely continuous with bounded derivative almost everywhere. Indeed, this is a consequence of the following proposition.

Proposition 21. Suppose g: $[a, b] \to \mathbf{R}$ is an absolutely continuous function and F : $[c, d] \to \mathbf{R}$ is an absolutely continuous function satisfying a Lipschitz condition with constant K, such that the range of g is contained in $[c, d]$. Then $F\circ g$ is absolutely continuous.

Proof. Since g is absolutely continuous, given any $\varepsilon > 0$, there exists $\delta > 0$ such that for any collection of disjoint open intervals (a_1, b_1), (a_2, b_2),…, $(a_n\ b_n)$ in $[a, b]$,

$$\sum_{i=1}^{n}(b_i - a_i) < \delta \Rightarrow \sum_{i=1}^{n}|g(b_i) - g(a_i)| < \frac{\varepsilon}{K}.$$

Then for each i =1, …, n, since F satisfies a Lipschitz condition with constant K,

$$|F(g(b_i)) - F(g(a_i))| \le K|g(b_i) - g(a_i)|.$$

Therefore, taking summation,

$$\sum_{i=1}^{n}|F(g(b_i)) - F(g(a_i))| \le \sum_{i=1}^{n}K|g(b_i) - g(a_i)| = K\sum_{i=1}^{n}|g(b_i) - g(a_i)| < K\cdot\frac{\varepsilon}{K} = \varepsilon.$$

Hence, $\sum_{i=1}^{n}(b_i - a_i) < \delta \Rightarrow \sum_{i=1}^{n}|F\circ g(b_i) - F\circ g(a_i)| < \varepsilon$. Therefore, $F\circ g$ is absolutely continuous.

Proof of Theorem 19. Since $f : [c, d] \to \mathbf{R}$ is continuous, it is bounded and Lebesgue integrable. Therefore, the function $F : [c, d] \to \mathbf{R}$ defined by $F(x) = \int_c^x f(t)dt$ is absolutely continuous and satisfies a Lipschitz condition. It follows by Proposition 21, that $F\circ g$ is absolutely continuous. Hence, we have

$$\int_a^b (F\circ g)'(x)dx = F\circ g(b) - F\circ g(a) = \int_c^{g(b)} f(x)dx - \int_c^{g(a)} f(x)dx = \int_{g(a)}^{g(b)} f(x)dx.$$

By the Fundamental Theorem of calculus, since f is continuous on $[c, d]$, F is differentiable everywhere on $[c, d]$. Note that g is differentiable almost everywhere on $[a, b]$, since g is absolutely continuous. Therefore, $F\circ g$ is differentiable almost everywhere on $[a, b]$ because if g is differentiable at x and since F is differentiable at g(x), $F\circ g$ is differentiable at x. In particular, $(F\circ g)'(x) = F'(g(x))\, g'(x) = f(g(x))\, g'(x)$ almost everywhere on $[a, b]$. Therefore,

$$\int_a^b f(g(x))g'(x)dx = \int_a^b (F\circ g)'(x)dx = \int_{g(a)}^{g(b)} f(x)dx.$$

This completes the proof.

Suppose $f : [c, d] \to \mathbf{R}$ is not necessarily continuous everywhere but just bounded and Lebesgue integrable. Then the question remains if $(F\circ g)'(x) = f(g(x))\, g'(x)$ almost everywhere.

We shall need the analogue of Proposition 10, Proposition 11 and Proposition 14.

Firstly, we recall the definition of the total variation function of a function of bounded variation. A function g is said to be of bounded variation on $[a, b]$ if the total variation

$$V_g[a,b] = \sup\left\{\sum_{i=1}^{n}|g(x_i) - g(x_{i-1})| : a = x_0 < x_1 < \cdots < x_n = b \text{ is a partition for}[a,b]\right\}$$

exists or equivalently, there exists a constant $K > 0$ such that for any partition $a = x_0 < x_1 < \cdots < x_n = b$, $\sum_{i=1}^{n} |g(x_i) - g(x_{i-1})| < K$. If g: $[a, b] \to \mathbf{R}$ is an absolutely continuous function, then the total variation function $V_g : [a,b] \to \mathbb{R}$ is defined by $V_g(x) = 0$ and $V_g(x) = V_g[a, x]$, the total variation of the function g on $[a, x]$ for x in $(a, b]$. This is well defined because, for any partition $a = x_0 < x_1 < \ldots < x_n = x$,

$$\sum_{i=1}^{n} |g(x_i) - g(x_{i-1})| \le V_g[a,b],$$

so that $\sup\left\{ \sum_{i=1}^{n} |g(x_i) - g(x_{i-1})| : a = x_0 < x_1 < \cdots < x_n = x \text{ is a partition for} [a, x] \right\}$

exists.

The following results concerning the properties of function of bounded variation and of absolutely continuous functions can be found on page 267 to page 269 in "*Principles of Real Analysis*" by C.D. Aliprantis and Owen Burkinshaw or in my article "*Monotone Function, Function of Bounded Variation, Fundamental Theorem of Calculus*".

Theorem 22.
1. g: $[a, b] \to \mathbf{R}$ is a function of bounded variation, if and only if, g is the difference of two increasing functions.
2. Any function g: $[a, b] \to \mathbf{R}$ of bounded variation is differentiable almost everywhere on $[a, b]$.
3. If g: $[a, b] \to \mathbf{R}$ is of bounded variation, then the variation function $V_g : [a, b] \to \mathbf{R}$ and the function $V_g - g$ are both increasing functions.
4. If g: $[a, b] \to \mathbf{R}$ is of bounded variation, then for any x and y in $[a, b]$ *with* $a \le x \le y \le b$ we have $|g(y) - g(x)| \le V_g(y) - V_g(x)$. In particular, $|g(x) - g(a)| \le V_g(x)$.

The next result concerns absolutely continuous functions.

Theorem 23. Suppose g: $[a, b] \to \mathbf{R}$ is absolutely continuous on $[a, b]$. Then the following statements hold.
1. g is of bounded variation.
2. The variation function $V_g : [a, b] \to \mathbf{R}$ is also absolutely continuous and so g is the difference of two increasing absolutely continuous functions.

We shall need the following technical result about increasing function.

Proposition 24. Suppose g: $[a, b] \to \mathbf{R}$ is an increasing function.
Then g is differentiable almost everywhere and the derivative g' is Lebesgue integrable (therefore measurable) and $\int_a^x g'(t)dt \le g(x) - g(a)$ for any x in $[a, b]$.

The proof of Proposition 24 can be found on page 100 of Royden's "*Real Analysis*".

We shall need the following consequence of absolute continuity.

Proposition 25. Suppose g: $[a, b] \to \mathbf{R}$ is absolutely continuous on $[a, b]$. Then for any x in $[a, b]$, $V_g(x) = \int_a^x |g'(t)|\,dt$.

Proof. By Theorem 22 part 4, for any $y \neq x$ in $[a, b]$, $(V_g(y) - V_g(x))/(y - x) \geq (g(y) - g(x))/(y - x)$. Consequently, $V_g'(x) \geq g'(x)$ almost everywhere on $[a, b]$. We also have , for any $y \neq x$ in $[a, b]$, $(V_g(y) - V_g(x))/(y - x) \geq -(g(y) - g(x))/(y - x)$ and so we have $V_g'(x) \geq -g'(x)$ almost everywhere on $[a, b]$. Therefore, $V_g'(x) \geq |g'(x)|$ almost everywhere on $[a, b]$.

Then $\int_a^x |g'(t)|\,dt \leq \int_a^x V_g'(x)dx \leq V_g(x) - V_g(a) = V_g(x)$.

The last inequality follows from Proposition 24 since the function V_g is an increasing function.

On the other hand, for any x in $(a, b]$ and for any partition of $[a, x]$, $a = x_0 < x_1 < \ldots < x_n = x$, since g is absolutely continuous (and therefore the fundamental theorem for Lebesgue integral holds),

$$\sum_{i=1}^{n} |g(x_i) - g(x_{i-1})| = \sum_{i=1}^{n} \left| \int_{x_{i-1}}^{x_i} g'(t)dt \right| \leq \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} |g'(t)|\,dt = \int_a^x |g'(t)|\,dt.$$

Therefore,

$$\sup\left\{ \sum_{i=1}^{n} |g(x_i) - g(x_{i-1})| : a = x_0 < x_1 < \cdots < x_n = x \text{ is a partition for}[a, x] \right\} \leq \int_a^x |g'(t)|\,dt.$$

It follows that $V_g(x) \leq \int_a^x |g'(t)|\,dt$. Therefore, $V_g(x) = \int_a^x |g'(t)|\,dt$.

Corollary 26. Suppose g: $[a, b] \to \mathbf{R}$ is absolutely continuous on $[a, b]$. Then $V_g'(x) = |g'(x)|$ almost everywhere on $[a, b]$.

Proof, Since $V_g(x) = \int_a^x |g'(t)|\,dt$, $V_g'(x) = |g'(x)|$ almost everywhere on $[a, b]$. This is because g is absolutely continuous and so $|g'(x)|$ is Lebesgue integrable. It follows that the differentiation of the indefinite integral $\int_a^x |g'(t)|\,dt$ yields $|g'(x)|$ almost everywhere. See for example Theorem 10 on page 107 in Royden's "*Real Analysis*" or see the reference to Theorem 2 of my article "*Integration by Parts*".

Remark 6. The conclusion of Corollary 26 is actually true just for function of bounded variation defined on $[a, b]$. But the proof is much more delicate, it requires a careful handling of how the finite difference $g(x_i)$ - $g(x_{i-1})$ changes sign. One proof of this is to choose a dissection $a = x_0 < x_1 < \cdots < x_k = b$ of $[a, b]$ so that the summation $\sum_{i=1}^{k} |g(x_i) - g(x_{i-1})|$ differs from the total variation by say $1/2^n$. The choice is such that the series $\sum_{n=1}^{\infty} \frac{1}{2^n}$ converges. Define a function g_n so as to give the same absolute value $g_n(x_i) - g_n(x_{i-1}) = |g(x_i) - g(x_{i-1})|$ for each i. The main difficulty is to show that $g_n'(x) \to V_g'(x)$ almost everywhere on $[a, b]$. Note that $g_n'(x) = \pm\, g'(x)$ almost everywhere on $[a, b]$. By showing that the difference $V_g(x) - g_n(x)$ is increasing for each n and by

using the series function $\sum_{n=1}^{\infty}\left(V_g(x)-g_n(x)\right)$, which is convergent, so that if it is possible to show that this series function can be differentiated almost everywhere term by term, we can then conclude that $V_g{}'(x)-g_n{}'(x) \to 0$ almost everywhere and the result then follows. We can indeed use Fubini's theorem on differentiation of a series function to do this.
(For a proof, of this more general result see *Arc Length, Functions of Bounded Variation and Total Variation*, Lemma 3 and its proof,)

We shall need the following results to use when proving the nullity of a set or of its image under a function.

Proposition 27. Suppose $f\,:[a, b] \to \mathbf{R}$ is a monotone increasing absolutely continuous function and E is a measurable subset of $[a, b]$. Then

$$\int_E f'(x)dx = m(f(E)).$$

Proof. We begin by proving the theorem for the special case when E is an open subset of $[a, b]$. Since E is open, E = a countable (finite or denumerable) union of *disjoint* open intervals, say $\{U_1, U_2, \dots\}$. Thus,

$$m(f(E)) = m\left(f\left(\bigcup_n U_n\right)\right) = m\left(\bigcup_n f(U_n)\right) = \sum_n m\left(f(U_n)\right),$$

since $\{f(U_1), f(U_2), \dots\}$ is a collection of non-overlapping intervals,

$$= \sum_n \left(f(b_n) - f(a_n)\right), \text{ where } U_n = (a_n, b_n),$$

$$= \sum_n \int_{a_n}^{b_n} f'(t)dt, \text{ because } f \text{ is absolutely continuous},$$

$$= \int_E f'(t)dt.$$

Note that since E is measurable and f is absolutely continuous, $f(E)$ is measurable so that $m*(f(E)) = m(f(E))$. Also for any open U, $f(U)$ is measurable and so $m*(f(U)) = m(f(U))$.
For the general case, suppose E is a measurable subset in $[a, b]$. Then for each positive integer n, there exists an open set G_n such that $E \subseteq G_n$ and $m(G_n) < m(E) + 1/n$ and an open set H_n such that $f(E) \subseteq H_n$ and $m(H_n) < m(f(E)) + 1/n$. Thus, $\lim_{n\to\infty} m(G_n) = m(E)$ and $\lim_{n\to\infty} m(H_n) = m(f(E))$.
For each positive integer n, $f^{-1}(H_n)$ is open by continuity of f. Therefore, $C_n = f^{-1}(H_n) \cap G_n$ is also open and contains E. Then note that

$$m(E) \le \lim_{n\to\infty} m(C_n) \le \lim_{n\to\infty} m(G_n) = m(E)$$

and so $\lim_{n\to\infty} m(C_n) = m(E)$.
Similarly, since $f(E) \subseteq f(C_n) = f(f^{-1}(H_n) \cap G_n) \subseteq H_n$,

$$m(f(E)) \le \lim_{n\to\infty} m(f(C_n)) \le \lim_{n\to\infty} m(H_n) = m(f(E)).$$

It follows that $\lim_{n\to\infty} m(f(C_n)) = m(f(E))$.

Therefore,

$$m(f(E)) = \lim_{n\to\infty} m(f(C_n)) = \lim_{n\to\infty}\int_{C_n} f'(t)dt \text{, since } C_n \text{ is also open,}$$
$$= \int_{\bigcap_{n=1}^{\infty} C_n} f'(x)dx \text{, by Lebesgue Dominated Convergence Theorem,}$$
$$= \int_E f'(x)dx + \int_{\bigcap_{n=1}^{\infty} C_n - E} f'(x)dx = \int_E f'(x)dx,$$

because $m\left(\bigcap_{n=1}^{\infty} C_n - E\right) = 0$ as $E \subseteq \bigcap_{n=1}^{\infty} C_n$ and $m\left(\bigcap_{n=1}^{\infty} C_n\right) = m(E)$ for

$m(E) \le m\left(\bigcap_{n=1}^{\infty} C_n\right) = \lim_{n\to\infty} m(C_n) \le m(E)$.

This completes the proof.

An easy consequence of the above proposition is the following:

Proposition 28. Suppose g: $[a, b] \to \mathbf{R}$ is absolutely continuous on $[a, b]$. Then for any measurable subset E of $[a, b]$, $m(g(E)) \le \int_E |g'(t)|\,dt$. Hence, the measure of $g(\{x \in [a, b]: g'(x) = 0\})$ is zero.

Proof. We may assume that $E \subseteq (a, b)$. Take any open set G covering $V_g(E)$ such that $m(G) < m(V_g(E)) + \varepsilon$. Then since V_g is continuous, $O = V_g^{-1}(G)$ is an open set covering E. We may assume that $O \subseteq (a, b)$. O is at most a countable union of pairwise disjoint open intervals in $[a, b]$, $O = \bigcup_k I_k$. For each I_k, $m(g(I_k)) \le m(V_g(I_k))$ since for any x, y in I_k, $|g(x)-g(y)| \le V_g(b_k)-V_g(a_k) = m(V_g(I_k))$, where $I_k = (a_k, b_k)$. Thus, since $E \subseteq O$,

$$m(g(E)) \le m(g(O)) \le \sum_k m(g(I_k)) \le \sum_k m(V_g(I_k)) = m(V_g(O)) \le m(G) < m(V_g(E)) + \varepsilon .$$

Since ε is arbitrary, it follows that $m(g(E)) \le m(V_g(E))$. Therefore, since V_g is increasing and absolutely continuous, we have by Proposition 27,

$$m(g(E)) \le m(V_g(E)) = \int_E V_g'(t)dt = \int_E |g'(t)|\,dt .$$

The last equality is because $V_g'(x) = |g'(x)|$ almost everywhere on $[a, b]$ by Corollary 26.

The last assertion is now obvious.

Remark 7. The last assertion of Proposition 28 slightly generalises Proposition 10. The assertion is actually true without the assumption of absolute continuity. See Proposition 10*.

Our next result is the following proposition, guaranteeing that for an absolutely continuous function g, if it maps a set onto a set of measure zero, then its total variation function maps the same set onto a set of measure zero too.

Proposition 29. Suppose g: $[a, b] \to \mathbf{R}$ is an absolutely continuous function. Then for any subset E such that the measure of its image under g , $m(g(E))$, is zero, we have that $m(V_g(E)) = 0$.

Before we embark on the proof, consider a sequence of function g_n defined by a sequence of partitions of $[a, b]$ whose norm tends to zero and for which g_n on each of the subinterval of the partition is equal to $\pm\, g(x)$ + *constant.* Then each $g_n(E)$ has measure zero because on each of these subintervals reflection and translation preserve measure. Note that we would also want g_n to converge to the total variation function V_g almost everywhere. Put in geometrical terms, we would want to stretch g by a sequence of moves to V_g while preserving the total variation and each move carries the set E onto a set of measure zero. The proof below uses the same idea.

Proof of Proposition 29.

If measure of E is zero, we have nothing to prove, since both g and V_g are absolutely continuous. For each positive integer n, cover the set E by an open set O_n such that $m(O_n) \le m(E) + 1/n$. We may assume that $O_n \supseteq O_{n+1} \supseteq O_{n+2} \ldots$. Then since O_n is open, it is the disjoint union of at most countable number of open intervals, $I_1^n, I_2^n, \cdots, I_{s(n)}^n$. ($s(n)$ here may be infinite). Let $I_k^n = \left(a_k^n, b_k^n\right)$. Then $m(O_n) = \sum_{k=1}^{s(n)} m\left(I_k^n\right)$ and $m(O_n)$ tends to $m(E)$ as n tends to infinity. Hence, $m\left(\bigcap_{n=1}^{\infty} O_n\right) = m(E)$. Absolute continuity makes the proof easier.

Since $V_g{}'(x) = |g'(x)|$ almost everywhere on $[a, b]$, we may assume that *for every x in E* , $V_g{}'(x) = |g'(x)|$.

This is because g is absolutely continuous and so V_g is also absolutely continuous by Theorem 23 and so both maps set of zero measure to set of measure zero by Proposition 9. So we can replace E by $E - H$, where $H = \{x\in [a, b]: V_g{}'(x) \neq |g'(x)|\}$. Note that $m(H) = 0$ by Corollary 26. Thus, by Proposition 9, $m(V_g(H)) = 0$ because V_g is absolutely continuous. If we can show that $m(V_g(E - H)) = 0$, then since $m(V_g(H)) = 0$, $m(V_g(E)) \le m(V_g(E - H)) + m(V_g(H))$ implies that $m(V_g(E)) = 0$.

For each x in E, given any $\varepsilon > 0$, there exists arbitrary small $h > 0$ such that

$$V_g(x+h) - V_g(x) < \left|g(x+h) - g(x)\right| + \varepsilon h \,. \quad \text{------------------} \quad (17)$$

Thus, for each x in E there are arbitrary small intervals $[x, x+h]$ in O_n , hence in some I_k^n such that (17) holds. We shall define a series of family of intervals $\{J^n\}$ such that each family has the measure of its intersection with E the same as the measure of E as follows. For each integer n, by the Vitali Covering Theorem, there exists in O_n , countable pairwise disjoint intervals, $J_i^n = [x_i^n, x_i^n + h_i^n)$, $i = 1, \cdots$, such that (17) holds and $m\left(E - \bigcup_{i=1}^{\infty} J_i^n\right) = 0$ with $\bigcup_{i=1}^{\infty} J_i^n \subseteq O_n$.

Thus, for each i,

$m(V_g(J_i^n)) = V_g(x_i^n + h_i^n) - V_g(x_i^n) < \left|g(x_i^n + h_i^n) - g(x_i^n)\right| + \varepsilon h_i^n \le m(g(J_i^n)) + \varepsilon h_i^n$.

Therefore,

$$\sum_{i=1}^{\infty} m(V_g(J_i^n)) \le \sum_{i=1}^{\infty} m(g(J_i^n)) + \varepsilon \sum_{i=1}^{\infty} h_i^n \le \sum_{i=1}^{\infty} m(g(J_i^n)) + \varepsilon(b-a) . \quad \text{---------} \ (18)$$

Now $m\left(E \cap J^n\right) = m(E)$, where $J^n = \bigcup_{i=1}^{\infty} J_i^n$. In this way we define the family $\{J^n\}$.

Define the following characteristic functions,

$$L_i^n(y) = \begin{cases} 1, & \text{if } y \in g(J_i^n), \\ 0, & \text{otherwise} \end{cases} .$$

Then $\int_{-\infty}^{\infty} L_i^n(y)dy = m(g(J_i^n))$.

And so from (18) we get,

$$m(V_g(E)) = m(V_g(E \cap J^n)) \le \sum_{i=1}^{\infty} m(V_g(J_i^n)) \le \sum_{i=1}^{\infty} \int_{-\infty}^{\infty} L_i^n(y)dy + \varepsilon(b-a) . \text{--------} \ (19)$$

Now define $L^n(y) = \sum_{i=1}^{\infty} L_i^n(y)$. Note that for each n, $L^n(y) \le N(y)$ for all y, where $N(y)$ is the Banach Indicatrix function associated with the function g. That is $N(y)$ is the number of solutions of the equation $g(x) = y$ in $[a, b]$ if it is finite, infinity otherwise. $N(y)$ is integrable and the integral, $\int_{-\infty}^{\infty} N(y)dy = V_g(b)$ = the total variation of the function g on $[a, b]$. (For a reference see page 225 of the classic text, I.P Natanson's "Theory of Functions of a Real Variable".) Since $N(y)$ is integrable, the set on which $L^n(y) = \infty$ is of measure zero. Therefore, by the Lebesgue Monotone Convergence Theorem, $\int_{-\infty}^{\infty} L^n(y)dy = \sum_{i=1}^{\infty} \int_{-\infty}^{\infty} L_i^n(y)dy$. Thus, from (19), we obtain, for each integer $n > 0$,

$$m(V_g(E)) \le \int_{-\infty}^{\infty} L^n(y)dy + \varepsilon(b-a) . \qquad \text{-----------------------} \ (20)$$

Then by the Lebesgue Dominated Convergence Theorem, as $L^n(y)$ is dominated by $N(y)$, if L^n converges almost everywhere to a function L, then $\int_{-\infty}^{\infty} L(y)dy = \lim_{n\to\infty} \int_{-\infty}^{\infty} L^n(y)dy$. Now we want to show that the integral $\int_{-\infty}^{\infty} L(y)dy$ is zero. Note that, since the Banach indicatrix function $N(y)$ is summable, the set $D = \{y : N(y) = \infty\}$ is of measure zero. We now proceed to analyse the function $L(y)$. We shall show that $L(y) = 0$ almost everywhere, i.e., $\lim_{n\to\infty} L^n(y) = 0$ for almost all y in the closed interval $g([a, b]) = [c, d]$, where $c = \min\{g(x): x \in [a, b]\}$ and $d = \max\{g(x): x \in [a, b]\}$. Let $B = \bigcap_{n=1}^{\infty} \bigcup_{i=n}^{\infty} J^i$. We shall show next that if $L(y)$ is finite, then y is in $g(B)$. Firstly, we claim that $m(g(B)) = 0$. Note that $B \subseteq C = \bigcap_{n=1}^{\infty} O_n$ and $E \subseteq C$. $m(g(C)) \le m(g(E)) + m(g(C-E))$. Since $m(C-E) = 0$, we have that $m(g(C-E)) = 0$ as g is absolutely continuous. As $m(g(E)) = 0$, we have that $m(g(C)) \le 0$ and so $m(g(C)) = 0$. It follows that $m(g(B)) = 0$. Let K be the set in $\mathbb{R}$ for which $\lim_{n\to\infty} L^n(y) \ne 0$. Consider K

$-D$. We shall show that $K - D \subseteq g(B)$. Since each function, $L^n(y)$, is a non-negative integral function for each y in $K - D$, there exists a sequence $\{n_r\}$ such that $L^{n_r}(y) \geq 1$ because $\lim_{n\to\infty} L^n(y) \neq 0$. Thus, this means there exists at least one point x_{n_r} in $J_{n_r} = \bigcup_{i=1}^{\infty} J_i^{n_r}$ such that $g(x_{n_r}) = y$. Since $N(y) < \infty$, this means there can only be a finite number of solutions to $g(x_{n_r}) = y$. Therefore, the sequence $\left\{x_{n_r}\right\}_{r=1}^{\infty}$ contains a convergent constant subsequence. That is to say, some of the solution to $g(x) = y$ must occur infinitely often in the sequence $\left\{x_{n_r}\right\}_{r=1}^{\infty}$. Let such a solution be x_0. Then x_0 belongs to an infinite number of J^i and so in B, in particular $g(x_0) = y$ and hence $y \in g(B)$. Therefore, $K - D \subseteq g(B)$ and so $L(y) = 0$ almost everywhere and $\int_{-\infty}^{\infty} L(y)dy = 0$. Thus, by (20) we get

$$m(V_g(E)) \leq \lim_{n\to\infty} \int_{-\infty}^{\infty} L^n(y)dy + \varepsilon(b-a) = \int_{-\infty}^{\infty} L(y)dy + \varepsilon(b-a) = \varepsilon(b-a).$$

Since ε is arbitrary, we conclude that $m(V_g(E)) = 0$. This completes the proof.

Remark.
The conclusion of Proposition 29 is true for any function of bounded variation g.

Proposition 29*. Suppose g: $[a, b] \to \mathbf{R}$ is a function of bounded variation. Then for any subset E such that the measure of its image under g, $m(g(E))$, is zero, we have that $m(V_g(E)) = 0$. More precisely, $m(g(E)) = 0$ if and only if $m(V_g(E)) = 0$.

(For the proof see my article, *Functions of Bounded Variation and Johnson's Indicatrix*, Theorem 1.)

The next proposition is a slight generalisation of Proposition 11.

Proposition 30. Suppose g :$[a, b] \to \mathbf{R}$ is an absolutely continuous function. Suppose the range of g is $[c, d]$ and E is a subset of $[c, d]$ of measure zero. Let $H = \{ x \in [a, b] : g'(x) \neq 0\}$. Then the measure of $g^{-1}(E) \cap H$ is zero. Thus, $g'(x) = 0$ almost everywhere on $g^{-1}(E)$.

Proof. Since g is absolutely continuous, by Corollary 26, $V_g'(x) = |g'(x)|$ almost everywhere on $[a, b]$. Let $H' = \{ x \in [a, b] : V_g'(x) \neq 0\}$. Then H' is the same as the set $\{ x \in [a, b] : |g'(x)| \neq 0\} = H$ except perhaps on a subset of measure zero. That is $H' - C = H - C$ where $C = \{ x \in [a, b] : V_g'(x) \neq |g'(x)|\}$. Thus, we may assume that $H' = H$. This assumption will not affect the conclusion of the proposition. Therefore, $g^{-1}(E) \cap H = g^{-1}(E) \cap H'$. Then $g(g^{-1}(E) \cap H') \subseteq E$. Since E is of measure zero, $g(g^{-1}(E) \cap H')$ too is of measure zero. Then by Proposition 29, $m(V_g(g^{-1}(E) \cap H')) = 0$. Since V_g is monotonic increasing, by Proposition 11, $V_g^{-1}(V_g(g^{-1}(E) \cap H')) \cap H'$ is of measure zero. But $g^{-1}(E) \cap H' \subseteq V_g^{-1}(V_g(g^{-1}(E) \cap H'))$ and so $g^{-1}(E) \cap H' \cap H' \subseteq$

$V_g^{-1}(V_g(\, g^{-1}(E) \cap H'\,)\,) \cap H'$. Therefore, $g^{-1}(E) \cap H' = g^{-1}(E) \cap H$ is of measure zero. This completes the proof of Proposition 30.

Remark.

Proposition 30 is superseded by Proposition 13* without any condition on *g*.

We shall now state and prove the following generalisation of Theorem 19.

Theorem 31. Suppose g: $[a, b] \to \mathbf{R}$ is an absolutely continuous function and let $f : [c, d] \to \mathbf{R}$ be a bounded integrable function such that the range of g is contained in $[c, d]$. Then we have the following equality for Lebesgue integrals.

$$\int_a^b f(g(x))g'(x)dx = \int_{g(a)}^{g(b)} f(x)dx .$$

Proof. Since $f : [c, d] \to \mathbf{R}$ is Lebesgue integrable, we can define the function $F : [c, d] \to \mathbf{R}$ by $F(x) = \int_c^x f(t)dt$. By Proposition 20, F is absolutely continuous and satisfies a Lipschitz condition. It then follows by Proposition 21, that $F\circ g$ is absolutely continuous. Hence, we have

$$\int_a^b \left(F \circ g\right)'(x)dx = F \circ g(b) - F \circ g(a)$$
$$= \int_c^{g(b)} f(t)dt - \int_c^{g(a)} f(t)dt = \int_{g(a)}^{g(b)} f(t)dt .$$

It now remains to show that $(F\circ g)'(x) = F'(g(x))\, g'(x) = f(g(x))\, g'(x)$ almost everywhere on $[a, b]$. Since both $F\circ g$ and g are absolutely continuous, they are differentiable almost everywhere on $[a, b]$. Therefore, it is enough to consider subset K of $[a, b]$ on which both $(F\circ g)'(x)$ and $g'(x)$ exist finitely. If x is K, then either $g'(x) = 0$ or $g'(x) \neq 0$.

Suppose now that x is in K and $g'(x) \neq 0$. As F is differentiable almost everywhere on $[c, d]$, there exists a subset E in $[c, d]$ of measure zero such that F is differentiable everywhere in $[c, d] - E$. Let

$$H = \{x \in [a,b] : g \text{ is differentiable at } x \text{ finitely and } g'(x) \neq 0\}.$$

Then by Proposition 13*, measure of $g^{-1}(E) \cap H$ is zero. That is to say, $g'(x) = 0$ almost everywhere on $g^{-1}(E) \cap K$. Let $L = g^{-1}(E)$. Then for any x in $K - L = K - K \cap L$, $(F\circ g)'(x)$, $g'(x)$ and $F'(g(x))$ exist and so the Chain Rule holds on $K - L$. Since F is absolutely continuous, $C = F(E)$ is a set of measure zero and so by Proposition 13*, the measure of

$$\left(F \circ g\right)^{-1}(C) \cap \left\{x \in [a,b] : F \circ g \text{ is differentiable finitely at } x \text{ and } \left(F \circ g\right)'(x) \neq 0\right\}$$

is zero. Let $G = \left\{x \in [a,b] : F \circ g \text{ is differentiable finitely at } x \text{ and } \left(F \circ g\right)'(x) \neq 0\right\}$.

Note that $D = \left(F \circ g\right)^{-1}(C) = g^{-1}(F^{-1}(F(E))) \supseteq L$ and so $\left(F \circ g\right)'$ and g' are zero almost everywhere on $g^{-1}(E) \cap K = L \cap K$. Thus $\left(F \circ g\right)'$ and g' are zero on $L \cap K - L \cap H - L \cap G$. Note that both $L \cap H$ and $L \cap G$ are of measure zero and so

$L \cap (H \cup G)$ is of measure zero. We now let $K' = K - L \cap (H \cup G)$. Then on $K' - (L \cap K - L \cap (H \cup G))$ the Chain Rule holds and on $L \cap K - L \cap H - L \cap G$, $(F \circ g)'$ and g' are zero. Let $D'' = L \cap K - L \cap H - L \cap G$.

Hence, $(F \circ g)'(x) = F'(g(x)) \cdot g'(x)$ for x in $K'' = K' - D''$. Observe that

$D'' = \left\{ x : (F \circ g)'(x) = 0, g'(x) = 0 \text{ but } F'(g(x)) \text{ does not exist} \right\}$.

Next note that $F'(y) = f(y)$ almost everywhere on $[c, d]$. So there exists a set E' of measure zero in in $[c, d]$ such that $F'(y) = f(y)$ for all y in $[c, d] - E'$. Thus the set $\{x \in [a, b]: g'(x) \neq 0, \text{ and } F'(g(x)) \neq f(g(x))\} \subseteq g^{-1}(E') \cap H$, where $H = \{x \in [a, b]: g'(x) \neq 0\}$. Thus, by Proposition 13*, $\{x \in [a, b]: g'(x) \neq 0 \text{ and } F'(g(x)) \neq f(g(x))\}$ has measure zero. Consequently, $\left\{ x \in K'' : F'(g(x)) \neq f(g(x)) \text{ and } g'(x) \neq 0 \right\}$ has measure zero.

Thus, on $K'' - \left\{ x \in K'' : F'(g(x)) \neq f(g(x)) \text{ and } g'(x) \neq 0 \right\}$, either $F'(g(x)) = f(g(x))$ or $g'(x) = 0$.

Consequently, either $F'(g(x)) = f(g(x))$ or $g'(x) = (F \circ g)'(x) = 0$.

Schematically,

$$(F \circ g)'(x) = \begin{cases} (F \circ g)'(x), & \text{when } g'(x) \neq 0, \\ 0, & \text{when } g'(x) = 0 \end{cases} \quad \text{almost everywhere on } [a, b]$$

$$= \begin{cases} F'(g(x))g'(x), & \text{when } g'(x) \neq 0, \\ F'(g(x))g'(x), & \text{when } g'(x) = 0 \text{ and } F'(g(x)) \text{ exists,} \\ 0, & \text{when } g'(x) = 0 \text{ and } F'(g(x)) \text{ does not exist} \end{cases} \quad \text{almost everywhere on } [a, b]$$

$$= \begin{cases} f(g(x))g'(x), & \text{when } g'(x) \neq 0, \\ 0, & \text{when } g'(x) = 0 \end{cases} \quad \text{almost everywhere on } [a, b]$$

$= f(g(x))g'(x)$ almost everywhere on $[a, b]$.

Therefore, $\int_a^b f(g(x))g'(x)dx = \int_a^b (F \circ g)'(x)dx = \int_{g(a)}^{g(b)} f(x)dx$.

This completes the proof of Theorem 31.

We shall close the article with a proof of the Banach Indicatrix Theorem for continuous function of bounded variation.

Theorem 32 (Banach Indicatrix Theorem). Suppose $g: [a, b] \to \mathbf{R}$ is a continuous function of bounded variation. Then $\int_{-\infty}^{\infty} N(y)dy = V_g(b)$, where $N(y)$ is the Banach Indicatrix function defined by $N(y)$ is the number of solutions of the equation $g(x) = y$ in $[a, b]$ if it is finite, infinity otherwise. Here N is a function into the extended real number.

Proof. We shall make use of the fact that the function g is of bounded variation. That is to say the supremum of the sum $\sum_{i=1}^{n}\left|g(x_i)-g(x_{i-1})\right|$ over all possible partitions, $P: a = x_0 < x_1 < \cdots < x_n = b$ exists. That is,

$$\sup\left\{\sum_{i=1}^{n}\left|g(x_i)-g(x_{i-1})\right| : a = x_0 < x_1 < \cdots < x_n = b \text{ is a partition for} [a,b]\right\}$$

exists. Hence, we can use the oscillation with respect to a partition instead. We shall seek a sequence of partitions such that the oscillations with respect to the partitions converges to the total variation of g on $[a, b]$. If the total variation is zero, then g is a constant function and we have nothing to prove for $\int_{-\infty}^{\infty} N(y)dy = 0$. We now assume that the total variation $V_g(b) > 0$. Thus, there exists an integer K such that $1/K < V_g(b)$. Therefore, by the definition of the supremum, given any positive integer $r > K$, there exists a partition $N_r : a = x_0^r < x_1^r < \cdots < x_{n_r}^r = b$ such that

$$V_g(b) \geq \sum_{i=1}^{n_r}\left|g(x_i^r)-g(x_{i-1}^r)\right| > V_g(b) - \frac{1}{r} > 0.$$

Since g is continuous we shall consider using the measure of the image of each of the closed interval $[x^r_{i-1}, x^r_i] = \Delta^r_i$. By inserting more points if need be we may assume that the norm of the partition $\|N_r\| = \max\{ x^r_i - x^r_{i-1}, i = 1,\ldots,n_r\} < 1/r$. Thus, $\lim_{r\to\infty}\|N_r\| = 0$. By the Extreme Value Theorem and the Intermediate Value Theorem,

$$g(\Delta^r_i) = [g(a^r_i), g(b^r_i)],$$

where $a^r_i,\ b^r_i \in [x^r_{i-1}, x^r_i] = \Delta^r_i$.
Hence, for each i,

$$m(g(\Delta^r_i)) = g(b^r_i) - g(a^r_i) \geq |g(x^r_i) - g(x^r_{i-1})|.$$

Therefore, we have

$$V_g(b) \geq \sum_{i=1}^{n_r} m\big(g(\Delta_i^r)\big) \geq \sum_{i=1}^{n_r}\left|g(x_i^r)-g(x_{i-1}^r)\right| > V_g(b) - \frac{1}{r} > 0.$$

Now the oscillation of g with respect to the partition N_r is defined to be $Os_r = \sum_{i=1}^{n_r} m\big(g(\Delta_i^r)\big)$. We then have for each $r > K$,

$$Os_r = \sum_{i=1}^{n_r} m\big(g(\Delta_i^r)\big) > V_g(b) - \frac{1}{r}.$$

We choose the partitions N_r such that N_{r+1} contains all points in N_r. Define now

$$L_i^r(y) = \begin{cases} 1, & \text{if } y \in g([x_{i-1}^r, x_i^r]), \\ 0, & \text{otherwise} \end{cases}.$$

Then we have $\int_{-\infty}^{\infty} L_i^r(y)dy = m(E_i^r)$, where $E^r_i = g(\Delta^r_i) = [g(a^r_i), g(b^r_i)]$. It is easy to see by continuity that E^r_i and $g([x^r_{i-1}, x^r_i])$ are intervals, hence measurable and that $m(E^r_i) = m(g([x^r_{i-1}, x^r_i]))$. Define $L^r(y) = \sum_{i=1}^{n_r} L_i^r(y)$. Then $L^r(y)$ is a measurable function. Thus,

$$Os_r = \sum_{i=1}^{n_r} m\left(E_i^r\right) = \sum_{i=1}^{n_r} \int_{-\infty}^{\infty} L_i^r(y)dy = \int_{-\infty}^{\infty} L^r(y)dy > V_g(b) - \frac{1}{r} \ .$$

Because every subinterval of the partition N_{r+1} is part of a subinterval of N_r, $0 \le L^r(y) \le L^{r+1}(y)$. Let $L(y) = \lim_{r\to\infty} L^r(y)$. Then by the Lebesgue Monotone Convergence Theorem, $V_g(b) \ge \int_{-\infty}^{\infty} L(y)dy = \lim_{r\to\infty}\int_{-\infty}^{\infty} L^r(y)dy = \lim_{r\to\infty} Os_r \ge V_g(b)$. Therefore, $V_g(b) = \int_{-\infty}^{\infty} L(y)dy$. We shall now show that $L(y) = N(y)$ almost everywhere. Suppose that $N(y) = 0$, then this implies that y is not in $g([x^r_{i-1}, x^r_i))$ for all $r > K$ and all i. Therefore, $L_i^r(y) = L^r(y) = L(y) = 0$. Thus, $N(y) = L(y)$ if $N(y) = 0$. Now if $N(y) = t < \infty$, then since $\lim_{r\to\infty}\|N_r\| = 0$, each solution of the equation $y = g(x)$ will lie in a different subinterval $[x^r_{i-1}, x^r_i)$ for all large r, provided y is not equal to $g(b)$. This can be seen by taking all r such that $1/r < \min\{|p - q|: p$ and q are distinct solutions to $y = g(x)\}$. Thus, for all large r, $L_i^r(y) = 1$ for exactly t values of i and $L_i^r(y) = 0$ for all the other values of i. Hence $L^r(y) = t$ for all large r. Therefore, $L(y) = t$. Now we come to the case when $N(y) = \infty$. This means that there exists distinct t solutions of $g(x) = y$ no matter how large t is. Therefore, for all large r, following the argument above $L^r(y) \ge t$. This implies that $L(y) = \infty = N(y)$. Now for $y = g(b)$. If $N(y) = \infty$, then as above we see that $L(y) = \infty$. If $N(y) = t < \infty$, then $L(y) = t-1$. We have thus shown that $N(y) = L(y)$ almost everywhere on $\mathbb{R}$. Therefore, $\int_{-\infty}^{\infty} N(y)dy = \int_{-\infty}^{\infty} L(y)dy = V_g(b)$. This completes the proof of Theorem 32.

A/Prof. Ng Tze Beng
Email: tbengng@gmail.com

