Arc Length, Functions of Bounded Variation and Total Variation
by Ng Tze Beng

This article gives a formula for the length of a curve in $\mathbf{R}^n$. This is a general formula, which does not assume differentiability and in the derivation of this formula, some interesting results about function of bounded variation will be used. We shall give a more complete picture with the method and concepts used in its derivation than otherwise given in most elementary text.

Definition 1. A curve in $\mathbf{R}^n$ is a continuous function $f: [a, b] \to \mathbf{R}^n$, where the non-trivial interval $[a, b]$ is called the parameter interval and f is called a *parametrized* curve.

We can consider an estimate of the length by taking points on this curve and take the length of a inscribed polygonal curve passing through these points. To do this we take a partition Δ : $a = x_0 < x_1 < x_2 < \ldots < x_n = b$ for the interval $[a, b]$. Let P_0 be the point $f(x_0)$ and $P_i = f(x_i)$. Then the length of the polygonal curve $P_0P_1\ldots P_n$ is an approximation of the arc length P_0P_n.

We shall consider $\mathbf{R}^n$ as the usual space with Euclidean metric. That is, the distance between two points is the usual Euclidean distance. We have the usual Euclidean norm function,

$$\|\ \| : \mathbf{R}^n \to \mathbf{R},$$

defined by $\|V\| = \sqrt{(V_1^2 + V_2^2 + V_3^2 + \ldots + V_n^2)}$, where V_i is the i-th component of V. The distance between two points V and W in $\mathbf{R}^n$ is then given by $\| V - W\|$.

The length of the polygonal curve $P_0P_1\ldots P_n$ is given by

$$|P_0P_1| + |P_1P_2| + \cdots + |P_{n-1}P_n| \quad \text{or} \quad \sum_{i=1}^{n} |P_{i-1}P_i|.$$

Now, the length of each line segment $|P_{i-1}P_i|$ is the length of the line joining $f(x_{i-1})$ to $f(x_i)$. Thus, by the Euclidean distance,

$$|P_{i-1}P_i| = \|f(x_i) - f(x_{i-1})\|.$$

Therefore, the length of the polygonal curve $P_0P_1\ldots P_n$ is given by

$$\sum_{i=1}^{n} \|(f(x_i) - f(x_{i-1}))\| . \qquad \text{------------------} \qquad (1)$$

We define the arc length of the curve $f(x)$, for $a \le x \le b$ to be the least upper bound of all possible polygonal approximation as given by (1), if it exists. Therefore, the length of the curve f is just the total variation $T_f[a, b]$. If it exists, the curve is called a *rectifiable curve*, otherwise it is not rectifiable. Thus, by the definition of a function of bounded variation, f is rectifiable, if and only if, f is of bounded variation. Not all continuous curves on a closed and bounded domain are rectifiable.

Now we consider the component functions of f. First an easy observation.

Theorem 2. A curve $f: [a, b] \to \mathbf{R}^n$ is rectifiable if and only if each component function is continuous of bounded variation.

Theorem 2 is an easy consequence of the definition of bounded variation and the following inequality,

$$|V_j| \leq \|V\| = \sqrt{(V_1^2 + V_2^2 + V_3^2 + \ldots + V_n^2)} \leq |V_1| + |V_2| + |V_3| + \ldots + |V_n|.$$

Using the first part of this inequality, we can show that if f is of bounded variation, then each component function f_j is also of bounded variation. Using the second part of the above inequality, we can show that if each f_j is of bounded variation, then f is of bounded variation. Also note that f is continuous, if and only if, each component of f is continuous. One may use the above inequality to prove this or simply observe that a map into a product space is continuous, if and only if, each component is continuous.

The next result is to show that there is a relation between the derivative of the variation function of f and the norm of the derived function of f. Note that if f is of bounded variation, then it is differentiable almost everywhere. In particular, if g: $[a, b] \to \mathbf{R}$ is of bounded variation, then g' is measurable and summable or integrable. This is a consequence of the fact that if g is of bounded variation, then it is the difference of two increasing functions. (See Theorem 13 in *"Monotone Function, Function of Bounded Variation, Fundamental Theorem of Calculus"*.). Note that for any increasing function g on $[a, b]$, g is differentiable almost everywhere and g' is measurable and Lebesgue integrable. (For a reference, see Proposition 24 in "*Change of Variable or Substitution in Riemann and Lebesgue Integration*".)

Lemma 3. Suppose g: $[a, b] \to \mathbf{R}$ is of bounded variation. Then the variation function, $V_g : [a, b] \to \mathbf{R}$, defined by $V_g(x) = T_g[a, x]$ for $a < x \leq b$ and $V_g(a) = 0$, satisfies

$$(V_g)'(x) = |g'(x)| \text{ almost everywhere.}$$

This results carries over to curve in $\mathbf{R}^n$.

Lemma 4. Suppose f: $[a, b] \to \mathbf{R}^n$ is of bounded variation. Then the variation function, $T_f : [a, b] \to \mathbf{R}$, defined by $T_f(x) = T_f[a, x]$ for $a < x \leq b$ and $T_f(a) = 0$, satisfies

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

Lemma 3 is well known although not as well known as the weaker statement when g is also absolutely continuous.

To prove Lemma 4 we require several results.

Lemma 5. Suppose f: $[a, b] \to \mathbf{R}^n$ is such that each component is Lebesgue integrable. Then $\left\| \int_a^b f(x)dx \right\| \leq \int_a^b \|f(x)\| dx$.

Proof. $\left\| \int_a^b f(x)dx \right\|^2 = \left\langle \int_a^b f(x)dx, \int_a^b f(x)dx \right\rangle$, where $<, >$ is the inner product for $\mathbf{R}^n$. Let $\alpha = \int_a^b f(x)dx$. Then

$$\|\alpha\|^2 = \left\langle \int_a^b f(x)dx, \int_a^b f(x)dx, \right\rangle = \left\langle \alpha, \int_a^b f(x)dx \right\rangle$$

$$= \int_a^b \langle a, f(x)\rangle dx = \sum_{i=1}^{n} \alpha_i \int_a^b f_i(x)dx$$, where α_i and f_i are respectively the i-th components of α and f,

$$= \sum_{i=1}^{n} \int_a^b \alpha_i f_i(x)dx = \int_a^b \left(\sum_{i=1}^{n} \alpha_i f_i(x) \right) dx$$

$$= \int_a^b \langle \alpha, f(x)\rangle dx \le \int_a^b \|\alpha\| \, \|f(x)\| dx,$$ by the Schwarz inequality,

$$= \|\alpha\| \int_a^b \|f(x)\| dx .$$

Hence $\|\alpha\| \le \int_a^b \|f(x)\| dx$. This proves the required inequality.

Now we state a weaker version of Lemma 4.

Lemma 6. Suppose $f : [a, b] \to \mathbf{R}^n$ is absolutely continuous. Then the variation function, $T_f : [a, b] \to \mathbf{R}$, defined by $T_f(x) = T_f[a, x]$ for $a < x \le b$ and $T_f(a) = 0$, satisfies

$$(T_f)'(x) = \|f'(x)\|$$

almost everywhere on $[a, b]$.

Proof. Recall that $f : [a, b] \to \mathbf{R}^n$ is absolutely continuous if for any ε > 0, there exists some δ > 0 such that for any finite disjoint open intervals, $(a_1, b_1), (a_2, b_2), \ldots, (a_n\, b_n)$ in $[a, b]$, such that $\sum_{i=1}^{n}(b_i - a_i) < \delta$, then we have $\sum_{i=1}^{n} \| f(b_i) - f(a_i) \| < \varepsilon$. Obviously, each component of f is also absolutely continuous.
Firstly, observe that if f is of bounded variation, then

$$\| f(x) - f(y) \| \le T_f(x) - T_f(y) \quad \text{for } x > y.$$

From this we easily deduce, that for almost all x in $[a, b]$,

$$\|f'(x)\| \le (T_f)'(x) . \quad \text{---------------------------------} \quad (2)$$

Now fixed a x in $[a, b]$. For any partition, $\Delta : a = x_0 < x_1 < x_2 < \ldots < x_k = x$, for the interval $[a, x]$,

$$\sum_{i=1}^{k} \|f(x_i) - f(x_{i-1})\| = \sum_{i=1}^{k} \left\| \int_{x_{i-1}}^{x_i} f'(t)dt \right\|,$$

because $f_j(x_i) - f_j(x_{i-1}) = \int_{x_{i-1}}^{x_i} f_j'(t)dt$ for each j as f_j is absolutely continuous,

$$\le \sum_{i=1}^{k} \int_{x_{i-1}}^{x_i} \|f'(t)\| dt$$, by Lemma 5,

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

Thus, as $T_f(x)$ is the supremum for such sum, for any partition, $\Delta : a = x_0 < x_1 < x_2 < \ldots < x_k = x$, we have

$$T_f(x) \le \int_a^x \|f'(t)\| dt.$$

Therefore, for any $y > x$ we get, by taking the interval$[x, y]$ in place of $[a, x]$ above,

$$T_f[x, y] \le \int_x^y \|f'(t)\| dt.$$

Thus, $\dfrac{T_f(y)-T_f(x)}{y-x}=\dfrac{T_f[x,y]}{y-x}\le\dfrac{\int_x^y\|f'(t)\|dt}{y-x}$. Taking limits as y tends to x from above we have, $\lim_{y\to x^+}\dfrac{T_f(y)-T_f(x)}{y-x}\le\lim_{y\to x^+}\dfrac{\int_x^y\|f'(t)\|dt}{y-x}$, whenever the limits exist.

Similarly, interchanging the role of x and y for $x > y$ we get

$$\lim_{y\to x^-}\frac{T_f(x)-T_f(y)}{x-y}\le\lim_{y\to x^-}\frac{\int_y^x\|f'(t)\|dt}{x-y},$$ whenever the limits exist.

Therefore, $(T_f)'(x)\le\|f'(x)\|$ almost everywhere. Thus, $(T_f)'(x)=\|f'(x)\|$ almost everywhere. This completes the proof of Lemma 6.

Note that since f is of bounded variation on $[a, b]$, f_i' is measurable and integrable and so $\|f'(x)\|$ is integrable since it is dominated by the integrable function $\sum_{i=1}^{n}|f_i'(x)|$ because of the inequality $\|f'(x)\|\le\sum_{i=1}^{n}|f_i'(t)|$.

Now we shall prove Lemma 4. Our strategy is to write f as the sum of an absolutely continuous function and a singular function. Indeed for any function $f:[a,b]\to\mathbf{R}^n$ of bounded variation, $f = g + h$, where g is absolutely continuous and h is a singular function such that $h'(x) = 0$ almost everywhere. We can let g be defined by $g(x)=\int_a^x f'(t)dt$ and $h(x) = f(x) - g(x)$.

Proof of Lemma 4. We shall prove the lemma for the case when f is also singular, that is, $f'(x) = 0$ almost everywhere. That means for each i =1,2, …, n, $f_i'(x) = 0$ almost everywhere.
For any $y > x$, we have,

$$T_f(y)-T_f(x)=T_f[x,y]\le\sum_{i=1}^{n}T_{f_i}[x,y]=\sum_{i=1}^{n}T_{f_i}(y)-T_{f_i}(x),$$

where $f = (f_1, f_2, \ldots, f_n)$. Therefore, for $y > x$,

$$0\le\frac{T_f(y)-T_f(x)}{y-x}\le\sum_{i=1}^{n}\frac{T_{f_i}(y)-T_{f_i}(x)}{y-x}.$$

Hence, for almost all x in $[a, b]$,

$$0\le\lim_{y\to x^+}\frac{T_f(y)-T_f(x)}{y-x}\le\sum_{i=1}^{n}\lim_{y\to x^+}\frac{T_{f_i}(y)-T_{f_i}(x)}{y-x}.$$

Similarly, by interchanging the role of x and y, we have, for almost all x in $[a, b]$,

$$0\le\lim_{y\to x^-}\frac{T_f(y)-T_f(x)}{y-x}\le\sum_{i=1}^{n}\lim_{y\to x^-}\frac{T_{f_i}(y)-T_{f_i}(x)}{y-x}.$$

Thus, $0\le(T_f)'(x)\le\sum_{i=1}^{n}(T_{f_i})'(x)=\sum_{i=1}^{n}|f_i'(x)|=0$ almost everywhere by Lemma 3.

Hence, $(T_f)'(x)=\|f'(x)\|=0$ almost everywhere.

Now we proceed to consider the general case when f is of bounded variation. Write $f = g + h$, where g is absolutely continuous and h is a singular function. Then for $y > x$,

$$T_f(y)-T_f(x)=T_f[x,y]\le T_g[x,y]+T_h[x,y]=T_g(y)-T_g(x)+T_h(y)-T_h(x).$$

Therefore, for $y > x$,

$$\frac{T_f(y)-T_f(x)}{y-x} \le \frac{T_g(y)-T_g(x)}{y-x} + \frac{T_h(y)-T_h(x)}{y-x}.$$

Thus, as before, for almost all x,

$$\lim_{y\to x}\frac{T_f(y)-T_f(x)}{y-x} \le \lim_{y\to x}\frac{T_g(y)-T_g(x)}{y-x} + \lim_{y\to x}\frac{T_h(y)-T_h(x)}{y-x}.$$

This means, for almost all x,

$(T_f)'(x) \le (T_g)'(x) + (T_h)'(x) = (T_g)'(x) + 0$, since h is singular,

$= \|g'(x)\|$, by Lemma 6, since g is absolutely continuous,

$= \|f'(x)\|$, since $f'(x) = g'(x)$ almost everywhere.

On the other hand, by (2), for almost all x in $[a, b]$

$$\|f'(x)\| \le (T_f)'(x).$$

Therefore, $(T_f)'(x) = \|f'(x)\|$ almost everywhere.

Now it remains to look at the arc length of the curve $f : [a, b] \to \mathbf{R}^n$ itself. Note that this is just the total variation $T_f[a,b]$ of the function. By the definition of a curve, f is continuous and so uniformly continuous since the domain $[a, b]$ is compact. Thus, there is actually a practical way to obtain an approximation of the arc length. That is for each n, we consider a regular partition of $[a, b]$, $\Delta_n : a = x_0 < x_1 < x_2 < \ldots < x_n = b,$ such that the norm of the partition $\|\Delta_n\| = (b-a)/n$. Then the arc length L_n of the polygonal curve defined by the partition Δ_n tends to the arc length of f as n tends to infinity. We could not obtain totally an integral formula for the arc length of f in general. There will be an error term, which vanishes if f is absolutely continuous. The integral part of the formula will actually come from the absolutely continuous part of f. The next theorem will imply that the total variation of the function f is the sum of the total variation of its absolutely continuous part and that of the singular part.

Theorem 7. Suppose $g : [a, b] \to \mathbf{R}^n$ is absolutely continuous and $h : [a, b] \to \mathbf{R}^n$ is such that each component of h is of bounded variation and that $h'(x) = 0$ almost everywhere in $[a, b]$. Then $T_{g+h}[a, b] = T_g[a, b] + T_h[a, b]$.

Proof. First note that $g + h$ is of bounded variation. It then follows that for any $y > x$ in $[a, b]$,

$$T_{g+h}[x, y] \le T_g[x, y] + T_h[x, y].$$

Similarly, $T_h[x, y] = T_{g+h-g}[x, y] \le T_{g+h}[x, y] + T_{-h}[x, y] = T_{g+h}[x, y] + T_h[x, y]$.

Thus, we have

$$|T_{g+h}[x, y] - T_h[x, y]| \le T_g[x, y] \quad \text{---------------------------} \quad (3)$$

We shall show that $T_{g+h}(x) = T_g(x) + T_h(x)$ for all x in $[a, b]$. To do this, it is sufficient to show that $T_{g+h}(x) - T_h(x)$ is absolutely continuous.

Let $f(x) = T_{g+h}(x) - T_h(x)$. Then for any $y > x$,

$$\begin{aligned} |f(y) - f(x)| &= |T_{g+h}(y) - T_{g+h}(x) - (T_h(y) - T_h(x))| \\ &= |T_{g+h}[x, y] - T_h[x, y]| \\ &\le T_g[x, y] = T_g(y) - T_g(x), \text{ by inequality (3). ---------------} \quad (4) \end{aligned}$$

Since g is absolutely continuous, it follows that the variation function T_g is also absolutely continuous. Therefore, using (4) and the definition of absolute continuity, we can easily show that f is absolutely continuous.
Then, for almost all x in $[a, b]$,

$$\begin{aligned} f'(x) &= (T_{g+h})'(x) - (T_h)'(x) \\ &= \|(g+h)'(x)\| - \|h'(x)\|, \text{ by Lemma 4,} \\ &= \|g'(x)\|, \text{ since } h'(x) = 0 \text{ almost everywhere,} \\ &= (T_g)'(x), \text{ by Lemma 6.} \end{aligned}$$

Now since $f(x) - T_g(x)$ is absolutely continuous as both $f(x)$ and $T_g(x)$ are absolutely continuous and that its derivative $f'(x) - (T_g)'(x) = 0$ almost everywhere on $[a, b]$, $f(x) - T_g(x)$ is a constant. Because $f(a) = T_g(a) = 0$, we conclude that $f(x) = T_g(x)$ for all x in $[a, b]$. This means $T_{g+h}(x) - T_h(x) = T_g(x)$ and so $T_{g+h}(x) = T_g(x) + T_h(x)$ for all x in $[a, b]$. In particular, $T_{g+h}(b) = T_g(b) + T_h(b)$, i.e.,

$$T_{g+h}[a, b] = T_g[a, b] + T_h[a, b].$$

This completes the proof.

Theorem 8. Suppose $f: [a, b] \to \mathbf{R}^n$ is a rectifiable curve . Then the arc length of f, L_f is given by the following formula,

$$L_f = \int_a^b \|f'(t)\| dt + T_h[a, b],$$

for some singular function h such that $f = g + h$, where g is absolutely continuous and $h'(x) = 0$ for almost all x in $[a, b]$. We may take $h(x) = f(x) - \int_a^x f'(t)dt$. In particular, f is absolutely continuous, if and only if, $L_f = \int_a^b \|f'(t)\| dt$.

Proof. Let $g(x) = \int_a^x f'(t)dt$. Then g is absolutely continuous since it is an indefinite integral. Let $h(x) = f(x) - \int_a^x f'(t)dt$. Then $f = g + h$. Therefore, by Theorem 7,

$$\begin{aligned} L_f &= T_f[a, b] = T_{g+h}[a, b] = T_g[a, b] + T_h[a, b] \\ &= T_g(b) + T_h[a, b] \\ &= \int_a^b (T_g)'(t)dt + T_h[a, b], \text{ since } T_g \text{ is absolutely continuous,} \\ &= \int_a^b \|g'(t)\| dt + T_h[a, b], \text{ since } (T_g)'(t) = \|g'(t)\| \text{ almost everywhere,} \\ &= \int_a^b \|f'(t)\| dt + T_h[a, b], \text{ since } g'(t) = f'(t) \text{ almost everywhere.} \end{aligned}$$

If f is absolutely continuous, i.e., each component of f is absolutely continuous, then $h(x) = f(x) - \int_a^x f'(t)dt = f(x) - [f(x) - f(a)] = f(a)$ is a constant function and so $T_h[a, b] = 0$. Conversely suppose $T_h[a, b] = 0$, then $T_h[a, x] = 0$ for all x in $[a, b]$. Therefore, by Theorem 7, $T_f(x) = T_{g+h}(x) = T_{g+h}[a, x] = T_g[a, x] + T_h[a, x] = T_g[a, x] + 0 = T_g(x)$. Since g is absolutely continuous, $T_g(x)$ is also absolutely continuous and so $T_f(x) = T_g(x)$ is absolutely continuous. (See Theorem 23 of "*Change of Variable or Substitution in Riemann and Lebesgue Integration*" or Theorem 29.14 and its proof in "*Principles of Real Analysis*" by C.D. Aliprantis and Owen Burkinshaw.) It follows easily that f is absolutely continuous. This completes the proof.

Now we specialize to the graph of a function, the curve given by a continuous function $f: [a, b] \to \mathbf{R}$. Then the arc length of the curve in $\mathbf{R}^2$ is the arc length of the function $g: [a, b] \to \mathbf{R}^2$ given by $g(x) = (x, f(x))$. Thus g is rectifiable, if and only if, f is of bounded variation. The arc length of the curve is the total variation of g. Thus applying Theorem 8 to g we get:

Theorem 9. Suppose $f : [a, b] \to \mathbf{R}$ is continuous and the graph of f is rectifiable (hence f is of bounded variation). Then the arc length L of the curve is given by,

$$L = \int_a^b \sqrt{1 + (f'(t))^2}\, dt + T_h[a, b],$$

for some singular function h such that $h'(x) = 0$ almost everywhere and $h(x) = f(x) - \int_a^x f'(t)dt$. In particular, f is absolutely continuous, if and only if,

$$L = \int_a^b \sqrt{1 + (f'(t))^2}\, dt.$$

Proof. The graph of f is the curve in question. Let $g : [a, b] \to \mathbf{R}^2$ be defined by $g(x) = (x, f(x))$ for each x in $[a, b]$. Then g is of bounded variation. Then we can take a decomposition of g as follows,

$$g(x) = (x, F(x)) + (0, h(x)),$$

where $F(x) = \int_a^x f'(t)dt$ and $h(x) = f(x) - \int_a^x f'(t)dt$.
Then $G(x) = (x, F(x))$ is absolutely continuous and $H(x) = (0, h(x))$ is singular since $h'(x) = 0$ almost everywhere. Therefore, by Theorem 8, the arc length L is given by the arc length of g,

$$L = L_g = \int_a^b \| G'(t) \| dt + T_H[a, b] = \int_a^b \sqrt{1 + (F'(t))^2}\, dt + T_H[a, b]$$

$$= \int_a^b \sqrt{1 + (f'(t))^2}\, dt + T_h[a, b],$$

since $F'(x) = f'(x)$ almost everywhere and the total variation $T_H[a, b] = T_{(0,h)}[a, b] = T_h[a, b]$.
f is absolutely continuous, if and only if, $h(x)$ is a constant function, if and only if,

$$L = L_g = \int_a^b \sqrt{1 + (f'(t))^2}\, dt + T_h[a, b] = \int_a^b \sqrt{1 + (f'(t))^2}\, dt.$$

This completes the proof.

Now we proceed to show that a sequence of polygonal length tends to the arc length. This is a consequence of the following result.

Theorem 10. Suppose $f : [a, b] \to \mathbf{R}^n$ is a rectifiable curve. Then for any $\varepsilon > o$ there exists a $\delta > 0$ such that if P: $a = x_0 < x_1 < x_2 < \ldots < x_n = b$ is a partition for $[a, b]$ with $\|P\| < \delta$, then

$$T_f[a, b] - \varepsilon < \sum_{i=1}^{n} \|(f(x_i) - f(x_{i-1}))\| \le T_f[a, b].$$

Proof. Since $T_f[a, b] = \sup\{\sum_{i=1}^{n} \|(f(x_i) - f(x_{i-1}))\| : \Delta : a = x_0 < x_1 < \cdots < x_n = b$ is a partition for $[a, b]\}$,

given any $\varepsilon > 0$, there exists a partition, $Q : a = T_0 < T_1 < T_2 < \ldots < T_N = b$, *such that*

$$T_f[a, b] - \varepsilon < \sum_{i=1}^{N} \|(f(T_i) - f(T_{i-1}))\| \le T_f[a, b]. \text{ ----------------- } (5)$$

Let $L = T_f[a, b]$. Now f is continuous on $[a, b]$ implies that f is uniformly continuous on $[a, b]$. Therefore, there exists a $\delta > 0$ such that whenever $|x - y| < \delta$, we have

$$\|f(y) - f(x)\| < \frac{\varepsilon}{4N} \text{ . -- } (6)$$

Suppose now P: $a = t_0 < t_1 < t_2 < \ldots < t_m = b$ is any partition such that the norm of the partition, $\|P\| < \delta$. Then the length of the polygonal curve defined by this partition is given by

$$L_P = \sum_{i=1}^{m} \|(f(t_i) - f(t_{i-1}))\|.$$

Let $S = Q \cup P$ and write the partition as

S: $a = s_0 < s_1 < s_2 < \ldots < s_M = b$.

Then the length of the polygonal curve defined by S is given by

$$L_S = \sum_{i=1}^{M} \|(f(s_i) - f(s_{i-1}))\| .$$

Obviously, by the triangle inequality, $L_P \le L_S$ and $L_Q \le L_S$. We shall show that our partition P satisfies the conclusion of the theorem.

Note that S is formed by adding points in Q to P. Adding a point T_i to P will increase the polygonal length of P by at most $\varepsilon/(2N)$. This is seen as follows.

Suppose for some j, $t_{j-1} < T_i < t_j$,

$$L_{P\cup\{T_i\}} = \sum_{k=1}^{j-1} \|f(t_k) - f(t_{k-1})\| + \|f(T_i) - f(t_{j-1})\| + \|f(t_j) - f(T_i)\| + \sum_{k=j+1}^{m} \|f(t_k) - f(t_{k-1})\|$$

Then,

$$\begin{aligned} L_{P\cup\{T_i\}} - L_P &= \|f(T_i) - f(t_{j-1})\| + \|f(t_j) - f(T_i)\| - \|f(t_j) - f(t_{j-1})\| \\ &\le \|f(T_i) - f(t_{j-1})\| + \|f(t_j) - f(T_i)\| \\ &\le \frac{\varepsilon}{4N} + \frac{\varepsilon}{4N} = \frac{\varepsilon}{2N} \quad \text{by (6).} \end{aligned}$$

Therefore, since Q has at most $N-1$ points different from points in P, we conclude that

$$L_S - L_P = L_{P\cup Q} - L_P \le (N-1)\cdot \frac{\varepsilon}{2N} \le \frac{\varepsilon}{2}.$$

Hence, $L_P \ge L_S - \frac{\varepsilon}{2} \ge L_Q - \frac{\varepsilon}{2} > L - \frac{\varepsilon}{2} - \frac{\varepsilon}{2} = L - \varepsilon$ by (5). Thus, the partition P: $a = t_0 < t_1 < t_2 < \ldots < t_m = b$ satisfies $T_f[a, b] - \varepsilon < L_P = \sum_{i=1}^{m} \|(f(t_i) - f(t_{i-1}))\| \le T_f[a, b]$. This completes the proof.

Now it remains to investigate how Lemma 3 can be proven. We shall use the Lebesgue decomposition of an increasing function to prove the assertion of the Lemma. We shall recall some interesting properties of increasing function. We shall define the saltus function of an increasing function. This is made possible because the set of points of discontinuity of an increasing function is at most countable.

Theorem 11. Suppose f: $[a, b] \to \mathbf{R}$ is a monotone function. Then the set of discontinuity of f is countable.

This is Theorem 3 in my article "*Monotone Function, Function of Bounded Variation and Fundamental Theorem of Calculus*".

We shall decompose an increasing function as a sum of a continuous function and a 'jump' function.

Definition 12. Suppose f: $[a, b] \to \mathbf{R}$ is an increasing function. Then the discontinuity of f can only be jump discontinuity. By Theorem 11, the set of discontinuity of f is countable. So let the set of discontinuity be $\{x_1, x_2, \ldots, x_n, \ldots\}$ where $x_1 < x_2 < \ldots < x_n < \ldots$, i.e., the x_i 's is ordered in an increasing order. We define the saltus function (or jump function) $s : [a, b] \to \mathbf{R}$ of f by

$s(a) = 0$,

$$s(x) = f(a+) - f(a) + \sum_{x_k < x} [f(x_k+) - f(x_k-)] + f(x) - f(x-), \text{ for } a < x \le b,$$

where $f(x^+) = \lim_{k \to x^+} f(k)$ and $f(x^-) = \lim_{k \to x^-} f(k)$ are the respective right and left limits at x. Plainly, $s(x)$ is an increasing function.

Let $\phi(x) = f(x) - s(x)$ for x in $[a, b]$. Then we shall show that $\phi : [a, b] \to \mathbf{R}$ is an increasing, continuous function.

Theorem 13. Suppose $f : [a, b] \to \mathbf{R}$ is an increasing function. Let s be the saltus function of f. Then the function $\phi : [a, b] \to \mathbf{R}$ defined by $\phi(x) = f(x) - s(x)$ for x in $[a, b]$ is increasing and continuous on $[a, b]$.

Proof. For $a \le x < y \le b$,

$$s(y) - s(x) = \sum_{x \le x_k < y} [f(x_k+) - f(x_k-)] + f(y) - f(y-) - [f(x) - f(x-)]$$

$$\le \begin{cases} \sum_{x < x_k < y} [f(x_k+) - f(x_k-)] + f(y) - f(y-), & \text{if } x \text{ is not a point of discontinuity} \\ f(x+) - f(x) + \sum_{x < x_k < y} [f(x_k+) - f(x_k-)] + f(y) - f(y-), & \text{if } x \text{ is a point of discontinuity} \end{cases}$$

$\le f(y) - f(x)$ by Theorem 2 of "*Monotone Function, Function of Bounded Variation and Fundamental Theorem of Calculus*".

Therefore, $\phi(x) = f(x) - s(x) \le f(y) - s(y) = \phi(y)$ and so ϕ is an increasing function.

Now, if x is a point of continuity of f, then obviously, for $x < y$,

$$f(x+) - f(x) \le s(y) - s(x)$$

and if x is a point of discontinuity of f, i.e., $x = x_j$ for some j, then

$$s(y) - s(x) = \sum_{x \le x_k < y} [f(x_k+) - f(x_k-)] + f(y) - f(y-) - [f(x) - f(x-)]$$
$$= f(x+) - f(x) + \sum_{x < x_k < y} [f(x_k+) - f(x_k-)] + f(y) - f(y-)$$
$$\ge f(x+) - f(x).$$

Therefore, $f(x+) - f(x) \le s(y) - s(x)$ for $x < y$.
Hence, $f(x+) - f(x) \le \lim_{y \to x^+} s(y) - s(x) = s(x+) - s(x)$. Therefore,

$$\phi(x+) = f(x+) - s(x+) \le f(x) - s(x) = \phi(x).$$

Since ϕ is increasing, $\phi(x) \le \phi(x+)$. It follows that $\phi(x) = \phi(x+)$.
Now, for $x < y$, if x is a point of discontinuity of f,

$$s(y) - s(x) = \sum_{x \le x_k < y} [f(x_k+) - f(x_k-)] + f(y) - f(y-) - [f(x) - f(x-)]$$
$$= f(x+) - f(x) + \sum_{x < x_k < y} [f(x_k+) - f(x_k-)] + f(y) - f(y-)$$
$$\ge f(y) - f(y-).$$

Plainly, if x is a point of continuity of f, then $f(x) - f(x-) = 0$ and so

$$s(y) - s(x) = \sum_{x \le x_k < y} [f(x_k+) - f(x_k-)] + f(y) - f(y-) - [f(x) - f(x-)]$$

$$= \sum_{x \le x_k < y} [f(x_k+) - f(x_k-)] + f(y) - f(y-)$$
$$\ge f(y) - f(y-).$$

Therefore, $s(y) - \lim_{x \to y^-} s(x) \ge f(y) - f(y-)$. I.e., $s(y) - s(y-) \ge f(y) - f(y-)$.

Thus, $\phi(y-) = f(y-) - s(y-) \geq f(y) - s(y) = \phi(y)$. Since ϕ is increasing, $\phi(y-) \leq \phi(y)$. Thus, $\phi(y-) = \phi(y)$ for any y in $[a, b]$. It then follows that $\phi(x) = \phi(x+) = \phi(x-)$. Thus, ϕ is continuous at x for any x in $[a, b]$. This completes the proof.

Now we summarize the above in the following:

Theorem 14. Suppose $f: [a, b] \to \mathbf{R}$ is an increasing function. Then f can be decomposed as a sum of a continuous function and a saltus type function as follows.

$$f(x) = \phi(x) + s(x) \text{ for all } x \text{ in } [a, b],$$

where $\phi(x) = f(x) - s(x)$ is increasing and continuous and $s(x)$ is the saltus function of f. Obviously $s'(x) = 0$ almost everywhere on $[a, b]$ and $f'(x) = \phi'(x)$ almost everywhere.

We shall seek for a better decomposition of f. We would like to decompose the function into a sum of an absolutely continuous function and a singular function. This is the Lebesgue decomposition of a monotone function.

Theorem 15. Suppose $f: [a, b] \to \mathbf{R}$ is an increasing function. Then

$$f(x) = F(x) + g(x) + s(x) \text{ for all } x \text{ in } [a, b], \quad \text{-------------------} \quad (7)$$

where $F(x)$ is an increasing, absolutely continuous function, $g(x)$ is a continuous increasing function, which is singular, i.e., $g'(x) = 0$ almost everywhere on $[a, b]$ and $s(x)$ is a saltus function of f.

Proof. Take the decomposition $f(x) = \phi(x) + s(x)$ given by Theorem 14.
Let $F(x) = \int_a^x f'(t)dt$. Then $F(x)$ is absolutely continuous being the indefinite integral of an integrable function. Then $F'(x) = f'(x) = \phi'(x)$ almost everywhere on $[a, b]$. Let $g(x) = \phi(x) - F(x)$. Then $g(x)$ is continuous since both F and g are continuous on $[a, b]$. Also $g'(x) = \phi'(x) - F'(x) = f'(x) - f'(x) = 0$ almost everywhere on $[a, b]$. Note that g is also increasing. For $a \leq x < y \leq b$,

$$\begin{aligned} g(x) - g(y) &= \phi(x) - \phi(y) - [F(x) - F(y)] \\ &= \phi(x) - \phi(y) + [F(y) - F(x)] = \phi(x) - \phi(y) + \int_x^y \phi'(t)dt \\ &\leq \phi(x) - \phi(y) + \phi(y) - \phi(x) = 0, \end{aligned}$$

because $\int_x^y \phi'(t)dt \leq \phi(y) - \phi(x)$ for increasing function ϕ
(see for example Proposition 24 in "*Change of Variable or substitution in Riemann and Lebesgue Integration*").

Therefore, $g(x) \leq g(y)$ for $x < y$. Hence, g is increasing.
Thus, $f(x) = \phi(x) + s(x) = F(x) + g(x) + s(x)$ is the desired decomposition.

Next we shall consider for a function of bounded variation $f: [a, b] \to \mathbf{R}$ the most efficient way of decomposing $f(x) - f(a)$ as the difference of two increasing functions.

Let Q: $a = x_0 < x_1 < x_2 < \ldots < x_n = b$ be a partition for $[a, b]$. Define

$$p(Q) = \sum_{i=1}^{n} \max[0, (f(x_i) - f(x_{i-1})],$$

$$n(Q) = -\sum_{i=1}^{n} \min[0, (f(x_i) - f(x_{i-1})],$$

$$t(Q) = p(Q) + n(Q) = \sum_{i=1}^{n} |f(x_i) - f(x_{i-1}|.$$

Note that $p(Q)$ is just the sum over the terms for which $f(x_i) - f(x_{i-1}) \geq 0$, $-n(Q)$ is the sum over the terms for which $f(x_i) - f(x_{i-1}) \leq 0$ and so

$$p(Q) - n(Q) = \sum_{i=1}^{n}(f(x_i) - f(x_{i-1})) = f(b) - f(a) .$$

Define for a function $f : [a, b] \to \mathbf{R}$ of bounded variation,
the positive variation,

$P_f[a,b] = \sup \{ p(Q): Q \text{ a partition of } [a, b] \}$,

the negative variation,

$N_f[a,b] = \sup \{ n(Q): Q \text{ a partition of } [a, b] \}$

and the total variation,

$T_f[a,b] = \sup \{ t(Q): Q \text{ a partition of } [a, b] \}$.

Plainly, all three variations exist since f is of bounded variation.
Since $p(Q) + n(Q) = t(Q)$,

$$P_f[a,b] + N_f[a,b] = T_f[a,b]. \quad \text{-------------------------} \quad (8)$$

We deduce this as follows. Plainly, $p(Q) + n(Q) \leq T_f[a,b]$. It follows that $P_f[a,b] + N_f[a,b] \leq T_f[a,b]$. On the other hand, $t(Q)= p(Q) + n(Q) \leq P_f[a,b] + N_f[a,b]$, consequently $T_f[a,b] \leq P_f[a,b] + N_f[a,b]$. Thus, we have $P_f[a,b] + N_f[a,b] = T_f[a,b]$.
Similarly, since $p(Q) - n(Q) = f(b) - f(a)$,

$$P_f[a,b] - N_f[a,b] = f(b) - f(a). \quad \text{----------------------} \ (9)$$

Now we define the positive variation function $P_f : [a, b] \to \mathbf{R}$ of f by

$P_f(a)= 0$, $P_f(x) = P_f[a,x]$ for $a < x \leq b$.

The negative variation function $N_f : [a, b] \to \mathbf{R}$ of f is defined similarly by

$N_f(a)= 0$, $N_f(x) = N_f[a,x]$ for $a < x \leq b$.

Finally the total variation function of f is of course defined by

$T_f(a)= 0$, $T_f(x) = T_f[a,x]$ for $a < x \leq b$.

Thus, it follows from (8), that

$$T_f(x) = P_f(x) + N_f(x) \quad \text{---------------------} \quad (10)$$

and from (9),

$$f(x) - f(a) = P_f(x) - N_f(x). \quad \text{-------------------} \ (11).$$

Plainly, these three variation functions are increasing functions.

We state the above conclusion as follows:

Theorem 16. Suppose $f : [a, b] \to \mathbf{R}$ is a function of bounded variation. Let $T_f(x)$, $P_f(x)$ and $N_f(x)$ be respectively the total, positive and negative variation functions of f. Then $f(x) - f(a) = P_f(x) - N_f(x)$ and $T_f(x) = P_f(x) + N_f(x)$.

The decomposition (11) is the most efficient in the sense of the following theorem.

Theorem 17. Suppose $f : [a, b] \to \mathbf{R}$ is a function of bounded variation. Suppose we have another decomposition of $f(x) - f(a)$,

$$f(x) - f(a) = g(x) - h(x)$$

where g and h are increasing functions with $g(a)=h(a) =0$. Then $P_f(x) \leq g(x)$ and $N_f(x) \leq h(x)$ on $[a, b]$.

Proof. Let Q: $a = x_0 < x_1 < x_2 < \ldots < x_n = b$ be a partition for $[a, b]$. Then

$$f(x_i) - f(x_{i-1}) = g(x_i) - g(x_{i-1}) - [h(x_i) - h(x_{i-1})] \leq g(x_i) - g(x_{i-1}).$$

Therefore, as $p(Q)$ is the sum over the terms for which $f(x_i) - f(x_{i-1}) \geq 0$,

$$p(Q) \leq \sum_{i=1}^{n}(g(x_i) - g(x_{i-1})) = g(b) - g(a) = g(b).$$

Thus $P_f[a,b] \leq g(b)$. It follows from this that $P_f[a,x] \leq g(x)$, for any $a < x \leq b$. This means $P_f(x) \leq g(x)$ for all x in $[a, b]$. Therefore, $P_f(x) - N_f(x) = f(x) - f(a) = g(x) - h(x) \leq g(x) - N_f(x)$ and so $N_f(x) \leq h(x)$.

Proof of Lemma 3. If $f : [a, b] \to \mathbf{R}$ is a function of bounded variation, then $(T_f)'(x) = |f'(x)|$ almost everywhere on $[a, b]$.

Let $P(x) = P_f(x)$ and $N(x) = N_f(x)$ be the positive and negative variation functions of f. Take their respective Lebesgue decomposition as given by Theorem 15,

$$P(x) = \int_a^x P'(t)dt + r_1(x) \quad \text{and}$$

$$N(x) = \int_a^x N'(t)dt + r_2(x)$$

where $r_1'(x) = r_2'(x) = 0$ almost everywhere on $[a, b]$ and both $r_1(x)$ and $r_2(x)$ are increasing functions satisfying $r_1(a) = r_2(a) = 0$.
Let $h(x) = \min(P'(x), N'(x))$. This is defined almost everywhere on $[a, b]$. Now define

$$w_1(x) = \int_a^x (P'(t) - h(t))dt + r_1(x) \quad \text{and}$$

$$w_2(x) = \int_a^x (N'(t) - h(t))dt + r_2(x).$$

Then obviously both $w_1(x)$ and $w_2(x)$ are increasing functions since both $P'(x) - h(x)$ and $N'(x) - h(x)$ are greater or equal to 0 almost everywhere on $[a, b]$. Also we have $w_1(a) = w_2(a) = 0$.
In particular,

$$w_1(x) - w_2(x) = P(x) - N(x) = f(x) - f(a). \quad \text{--------------------} \quad (12)$$

Therefore, by Theorem 17, $P(x) \leq w_1(x)$.

Thus, $w_1(x) = \int_a^x (P'(t) - h(t))dt + r_1(x) \geq P(x) = \int_a^x P'(t)dt + r_1(x)$ and so $\int_a^x -h(t))dt \geq 0$. Hence, $\int_a^x h(t))dt \leq 0$. But $\int_a^x h(t))dt \geq 0$, since $h(t) = \min(P'(t), N'(t)) \geq 0$ almost everywhere. Therefore, $\int_a^x h(t))dt = 0$ for all x and consequently, $h(x) = 0$ almost everywhere on $[a, b]$. That is to say, $\min(P'(x), N'(x)) = 0$ almost everywhere on $[a, b]$.
Now by (12), $f'(x) = P'(x) - N'(x)$ almost everywhere on $[a, b]$ and so

$$[P'(x) - h(x)] + [N'(x) - h(x)] = |P'(x) - N'(x)| = |f'(x)|$$

almost everywhere.
Hence, $|f'(x)| = P'(x) + N'(x) - 2h(x) = P'(x) + N'(x)$ almost everywhere since $h(x) = 0$ almost everywhere. But because $T_f(x) = P(x) + N(x)$, $(T_f)'(x) = P'(x) + N'(x)$ almost everywhere and so $(T_f)'(x) = |f'(x)|$ almost everywhere on $[a, b]$. This completes the proof of Lemma 3.

Ng Tze Beng
May 2006