AN APPROXIMATION OF JORDAN DECOMPOSABLE FUNCTIONS FOR A LIPSCHITZ FUNCTION

IBRAHEEM ALOLYAN

Received 21 March, 2007; accepted 14 November, 2007; published 30 November, 2007.
Communicated by: L. Leindler

MATHEMATICS DEPARTMENT, COLLEGE OF SCIENCE, KING SAUD UNIVERSITY P.O.Box: 2455, RIYADH 11451, SAUDI ARABIA.
ialolyan05@yahoo.com.

ABSTRACT. The well known Jordan decomposition theorem gives the useful characterization that any function of bounded variation can be written as the difference of two increasing functions. Functions which can be expressed in this way can be used to formulate an exclusion test for the recent Cellular Exclusion Algorithms for numerically computing all zero points or the global minima of functions in a given cellular domain [2, 8, 9]. In this paper we give an algorithm to approximate such increasing functions when only the values of the function of bounded variation can be computed. For this purpose, we are led to introduce the idea of \( \epsilon \)-increasing functions. It is shown that for any Lipschitz continuous function, we can find two \( \epsilon \)-increasing functions such that the Lipschitz function can be written as the difference of these functions.

Key words and phrases: Bounded variation, Jordan decomposition, \( \epsilon \)-increasing.

2000 Mathematics Subject Classification 26A45, 26A48.
1. INTRODUCTION

In the theorem of optimization, we have the problem of finding the global minimum of a function defined on a compact domain. In order to solve this problem, Cellular Exclusion Algorithms is used for numerically computing the global minima of functions in a given domain [2, 8, 9]. These algorithms formulate an exclusion test to discard cells that does not contain the global minimum. One of the tests, that are used in Exclusion Algorithms, can be applied to functions that can be written as the difference of two increasing functions [2].

The exclusion test in [2] was applied only to polynomials because they can be easily written as the difference of two increasing polynomials. However, if the function is not a polynomial, we find ourselves in some need of the concept of $\epsilon$-increasing condition which is weaker than increasing condition. As a consequence, we have got a solution to the optimization problem as shown in Theorem 3.1.

The well known Jordan decomposition theorem gives the useful characterization that any function of bounded variation can be written as the difference of two increasing functions. In practice, these functions can not be computed explicit. Therefore, we approximate these two functions and then prove that these functions are $\epsilon$-increasing in Theorem 3.1.

The notion of functions of bounded variation plays a very significant and important role in the theory of real functions [1, 5], numerical analysis [3, 4] and optimization [8]. In the literature, several properties of these functions have been discussed (see for example [1, 6, 7, 8, 10]); nevertheless, we focus our attention to one of these properties known as Jordan Decomposition Theorem JDT.

Decomposable functions, which result from JDT, plays an important role in optimization [2, 8]. For example, the Exclusion Algorithm uses decomposable functions as a test function for the minimization condition [2, 9].

Let $f : [a, b] \to \mathbb{R}$ be a function and $\mathcal{P} := \{x_i \in [a, b] : a = x_0 < x_1 < \ldots < x_{m+1} = b\}$ be a partition of $[a, b]$. We recall that the variation of $f$ over $\mathcal{P}$ is the nonnegative real number

$$V_{\mathcal{P}}[f; a, b] = \sum_{i=1}^{m+1} |f(x_i) - f(x_{i-1})|.$$

The function $f$ is a function of bounded variation if there exists a number $M$ such that for every partition $\mathcal{P}$ of $[a, b]$, we have

$$\sum_{i=1}^{m+1} |f(x_i) - f(x_{i-1})| \leq M.$$

The total variation of $f$ on $[a, b]$ is defined to be the number

$$V_f[a, b] := \sup_{\mathcal{P} | [a, b]} \sum_{i=1}^{m+1} |f(x_i) - f(x_{i-1})|,$$

where $\mathcal{P} | [a, b]$ means “$\mathcal{P}$ is a partition of $[a, b]$” [7]. For simplicity, we will write $\sup_{\mathcal{P}}$ instead of $\sup_{\mathcal{P} | [a, b]}$. The set of all functions of bounded variation on $[a, b]$ is denoted by $BV[a, b]$ and we have the following proposition which follows immediately from the definition of functions of bounded variation.

**Proposition 1.1.** [7] If $f : [a, b] \to \mathbb{R}$ is a Lipschitz function on $[a, b]$ (i.e., there is a constant $C$ such that

$$|f(x) - f(y)| \leq C|x - y| \quad \forall x, y \in [a, b],$$

where $|x - y|$ denotes the distance between $x$ and $y$. This proposition is known as the Lipschitz condition.
then \( f \in BV[a, b] \).

If we view the sum in (1.1) as a sum of positive and negative parts of the differences \( f(x_i) - f(x_{i-1}) \), then we can define \( P_f[a, b] \) to be the summation of the positive parts of \( f(x_i) - f(x_{i-1}) \) and \( N_f[a, b] \) to be the summation of the negative parts, i.e.,

\[
P_f[a, b] := \sup_{i=1}^{m+1} \sum (f(x_i) - f(x_{i-1}))^+,
\]

(1.3)

\[
N_f[a, b] := \sup_{i=1}^{m+1} \sum (f(x_i) - f(x_{i-1}))^-,
\]

where \( x^+ := \max\{0, x\} \) and \( x^- := \max\{0, -x\} \), then we have

\[
V_f[a, b] = P_f[a, b] + N_f[a, b],
\]

\[
f(b) - f(a) = P_f[a, b] - N_f[a, b].
\]

Varying \( b \) in (1.3) we get two functions \( P_f[a, \cdot], N_f[a, \cdot] : [a, b] \rightarrow \mathbb{R}^+ \cup \{0\} \) defined by

\[
P_f[a, x] := \sup_{\mathcal{P}[a,x]} \sum_{i=1}^{m+1} (f(x_i) - f(x_{i-1}))^+,
\]

(1.4)

\[
N_f[a, x] := \sup_{\mathcal{P}[a,x]} \sum_{i=1}^{m+1} (f(x_i) - f(x_{i-1}))^-.
\]

It can be shown that these two functions are increasing on the interval \([a, b]\). If we take \( p_f(x) = P_f[a, x] + f(a) \) and \( n_f(x) = N_f[a, x] \) as Jordan functions then we have the following theorem.

**Theorem 1.1** (Jordan Decomposition). If \( f \) is a function of bounded variation on \([a, b]\) then \( f \) can be written as the difference of two increasing functions

\[
f(x) = p_f(x) - n_f(x).
\]

This theorem states that we can write \( f \) as the difference of two increasing functions where each function can be computed by finding the supremum among all partitions. However, the supremum sum over all partitions cannot be computed numerically. Therefore, we approximate the functions \( p_f \) and \( n_f \) by considering the uniform partition \( \mathcal{P} := \{a + i(x - a)/(m + 1)\}_{i=0}^{m+1} \) of the interval \([a, x]\), where \( m \in \mathbb{N} \) and then study the consequences of this approximation.

In section 2, we explain the need of defining \( \epsilon \)-increasing functions. In Section 3, we write our algorithm to approximate \( p_f \) and \( n_f \); furthermore, we state and prove Theorem 3.1 for the functions \( p_m \) and \( n_m \) resulting from Algorithm 3.1.

### 2. Need of \( \epsilon \)-Increasing Definition

In order to approximate the functions \( p_f \) and \( n_f \) in Theorem 1.1, we use the uniform partition \( \mathcal{P} := \{a + i(x - a)/(m + 1)\}_{i=0}^{m+1} \) of \([a, x]\), where \( m \in \mathbb{N} \). Then we define \( P_m(x) \) and \( N_m(x) \)
to be
\[ P_m(x) := \sum_{i=1}^{m+1} (f(x_i) - f(x_{i-1}))^+, \]
(2.1)
\[ N_m(x) := \sum_{i=1}^{m+1} (f(x_i) - f(x_{i-1}))^- . \]

These two functions approximate the functions \( P_f[a, x] \) and \( N_f[a, x] \), respectively. Moreover, \( P_m \) and \( N_m \) approach \( P_f[a, x] \) and \( N_f[a, x] \) as \( m \to \infty \). Unfortunately, \( P_m \) and \( N_m \) are not guaranteed to be increasing if the function \( f \) is not monotone. We prove this in the following proposition.

**Proposition 2.1.** If the function \( f : [a, b] \to \mathbb{R} \) is not monotone on \([a, b]\), then no \( m \) can be chosen so that the functions \( P_m \) and \( N_m \) in (2.1) are increasing with respect to the uniform partition \( \mathcal{P} := \{a + i(x-a)/(m+1)\}_{i=0}^{m+1} \).

**Proof.** Suppose first that \( f \) is increasing on the interval \([a, c]\) and decreasing on the interval \([c, d]\) for \( c, d \in (a, b) \) and \( c < d \). In this case, the function \( P_m \) is increasing on \([a, c]\); nevertheless, if we let \( e = \min\{c + (c-a)/m, d\} \) and consider the interval \( \mathcal{I} = (c, e) \), then \( P_m(x) \) is less than \( P_m(c) \) for all \( x \in \mathcal{I} \). In order to prove that, let \( x \) be any point in \((c, e)\), and \( \mathcal{P} \) be the uniform partition of \([a, x]\), then \( f(x) < f(c) \) and \( x_m = a + m(x-a)/(m+1) \in (a+m(c-a)/(m+1), c) \) and we have

\[
\begin{align*}
P_m(x) &= \sum_{i=1}^{m+1} (f(x_i) - f(x_{i-1}))^+ \\
&= f(x_m) - f(x_0) + (f(x) - f(x_m))^+ \\
&< f(x_m) - f(x_0) + f(c) - f(x_m) \\
&= f(e) - f(x_0) = P_m(c).
\end{align*}
\]

Therefore, the function \( P_m \) is not increasing on \((c, e)\). If the function \( f \) is decreasing on the interval \([a, j]\) and increasing on the interval \([j, k]\) for \( j, k \in (a, b) \) and \( j < k \) then the function \( N_m \) is not increasing on the interval \((j, k)\) by the same argument that was done for \( P_m \) in the first case.

We define next the concept of oscillation of a function.

**Definition 2.1 (Oscillation).** We say that a continuous function \( f : [a, b] \to \mathbb{R} \) oscillates \( k \) times on the interval \([a, b]\) if there are exactly \( k \) points \( s_1, s_2, \ldots, s_k \in (a, b) \) such that for all \( i = 1, \ldots, k \), the value \( f(s_i) \) is either a strict local maximum or a strict local minimum. If the function \( f \) has an infinite number of maximum and minimum points on \((a, b)\), we say that \( f \) oscillates infinitely often.

For example, the function \( f : [-4, 4] \to \mathbb{R} \), defined by \( f(x) = x(x-1)(x-2)(x-3) \), oscillates 3 times and the function \( g : (0, 1) \to \mathbb{R} \), defined by \( g(x) = x \sin(1/x) \), oscillates infinitely often. In the following example, the function \( f \) is smooth and oscillates only one time; nevertheless, the function \( P_m \) in this case is not increasing for any choice of \( m \).

**Example 2.1.** Let \( f : [0, 2] \to \mathbb{R} \) be defined by \( f(x) := 1 - (1 - x)^2 \). In theory, this function can be written as the difference of two increasing functions, e.g., \( p_f(x) = 2x \) and \( n_f(x) = x^2 \). In order to compute \( p_m \) and \( n_m \) numerically, we notice that this function is increasing on the interval \([0, 1]\); therefore, the value of \( P_m(x) \) will increase on this interval for any choice of \( m \).
Nevertheless, on the interval \((1, (m + 1)/m)\); \(P_m\) will take values less than \(P_m(1) = 1\). To explain this, let \(x\) be any point in \((1, (m + 1)/m)\), and \(\mathcal{P} := \{ix/(m + 1)\}_{i=0}^{m+1}\) be the uniform partition of \([0, x]\), then \(f(x) < f(1)\) and \(mx/(m + 1) \in (m/(m + 1), 1)\) and we have

\[
P_m(x) = \sum_{i=1}^{m+1} (f(x_i) - f(x_{i-1}))^+
\]

\[
= f(mx/(m + 1)) - f(0) + (f(x) - f(mx/(m + 1)))^+
\]

\[
< f(mx/(m + 1)) - f(0) + (f(1) - f(mx/(m + 1)))
\]

\[
= f(1) - f(0) = 1 = P_m(1).
\]

![Figure 1: The function \(P_m(x)\) in Example 2.1](image)

Therefore, the function \(P_m\) is not increasing on \((1, (m + 1)/m)\). In fact, \(P_m(x)\) equals the value \(1 = P_m(1)\) if there is \(i, 1 \leq i \leq m + 1\), such that \(i/(m + 1) = 1\), i.e., the point 1 is one of mesh points. Solving for \(x\), we get \(x = (m + 1)/i\); therefore, \(P_m(x)\) equals one if \(x = (m + 1)/(m + 1) = 1, x = (m + 1)/m, x = (m + 1)/(m - 1), \ldots \) or \(x = m + 1\). In Figure 1 we plot the function \(p_m\) on the interval \([1, 2]\) for \(m = 10\) and 100. We can see from the plots that for \(m = 10\), we have \(P_m(x) = 1\) on the interval \([1, 2]\) for the values \(x = 1, 10/9, 10/8, 10/7, 10/6, \) and 10/5 ; however, for all other values; \(P_m\) is strictly less than one. For \(m = 100\), the function \(P_m\) oscillates more because it equals one at \(1, 100/99, 100/98, \ldots , 100/50\) and it is strictly less than one for all other values on \([1, 2]\). Although the functions in Figure 1 increase in oscillations, the amplitude decreases as \(m\) increases.

Proposition 2.1 and Example 2.1 show that \(m\) cannot be chosen in such a way that \(P_m\) and \(N_m\) are always increasing. In fact, the functions \(P_m\) and \(N_m\) will oscillate in a small \(\epsilon\)-neighborhood of \(f(c)\). As \(m\) increases, the number \(\epsilon\) becomes smaller and smaller. This idea led us to introduce a new definition that we call \(\epsilon\)-increasing and we will use it in the Decomposition Algorithm.

**Definition 2.2** (\(\epsilon\)-Increasing). We say that the function \(f : [a, b] \rightarrow \mathbb{R}\) is \(\epsilon\)-increasing, where \(\epsilon\) is a nonnegative number, if \(f(x) \leq f(y) + \epsilon\) for all \(x < y\) and \(x, y \in [a, b]\).

If a function \(f\) is increasing, then \(f(x) \leq f(y)\) for all \(x \leq y\). This implies that \(f(x) \leq f(y) + \epsilon\) for all \(\epsilon \geq 0\); therefore, the function \(f\) is \(\epsilon\)-increasing.
If a function $f$ is $\epsilon_1$-increasing, then $f$ is $\epsilon$-increasing for all $\epsilon \geq \epsilon_1$. For example, the function $f : [-10^{-2}, \infty) \rightarrow \mathbb{R}$, defined by $f(x) = x^2$ is $10^{-4}$-increasing because $f(x) \leq f(y) + 10^{-4}$ for all $x \leq y$. In fact, $f$ is $\epsilon$-increasing for all $\epsilon \geq 10^{-4}$. However, this function is not increasing because $f(-10^{-2}) > f(0)$.

Similarly, we can define the concept of what we call $\epsilon$-decreasing for decreasing function of some perturbation.

3. Decomposition Algorithm

In this section, we introduce the Decomposition Algorithm to compute the functions $p_m$ and $n_m$ that approximate $p_J$ and $n_J$, respectively. These functions are computed with respect to the uniform partition $\pi := \{a + i(x - a)/(m + 1)\}_{i=0}^{m+1}$ of the interval $[a, x]$ where $m$ is the number of points between $a$ and $x$. Consequently, the functions $p_m$ and $n_m$ will not be always increasing. We showed in Example 2.1 that $p_m$ and $n_m$ cannot be increasing on $[a, b]$ even for smooth functions and for any value of $m$. However, we prove in Theorem 3.1 that the resulting functions, $p_m$ and $n_m$, are $\epsilon$-increasing under some assumptions. In this algorithm we define $p_m$ and $n_m$ to be

$$
\begin{align*}
p_m(x) : &= f(a) + \sum_{i=1}^{m+1} (f(x_i) - f(x_{i-1}))^+,
\end{align*}
$$

(3.1)

$$
\begin{align*}
n_m(x) : &= \sum_{i=1}^{m+1} (f(x_i) - f(x_{i-1}))^-.
\end{align*}
$$

Algorithm 3.1. INPUT: $a$, $f$ and $x$. OUTPUT: $p$ and $n$ evaluated at $x$.
choose $m$ (number of points between $a$ and $x$).
\begin{align*}
d &:= (x - a)/(m + 1).
\end{align*}
for $i = 0 : m + 1$
\begin{align*}
x_i &:= a + id.
p &= f(a).
n &= 0.
\end{align*}
for $i = 1 : m + 1$
\begin{align*}
\text{if } s &= f(x_i) - f(x_{i-1}) \geq 0 
\text{ then } p &= p + s.
\text{else } n &= n + s.
\end{align*}

Note that we can choose $m + 1 = 2^k$ to decrease the number of computations when we increase $m$ (i.e., if we choose $m + 1 = 2^{k+1}$ then we already have computed the the values $f(x_i)$ for $i$ is even). The main question about this algorithm is how to compute $m$ such that $p$ and $n$ are increasing. Unfortunately, $m$ cannot always be computed even for smooth functions.

We have discussed in Proposition 1.1 that each Lipschitz function on $[a, b]$ is a function of bounded variation in this interval. In our next discussion, we choose the space of Lipschitz functions because in this space, the value $|f(x_i) - f(x_{i-1})|$ can be controlled by $|x_i - x_{i-1}|$. We come now to the main theorem in this paper which shows that $p_m$ and $n_m$ in Algorithm 3.1 have, in fact, a special property that can be exploited numerically.

**Theorem 3.1.** If $f : [a, b] \rightarrow \mathbb{R}$ is a Lipschitz function with a Lipschitz constant $C$, then for all $\epsilon > 0$ there exists an $m := m_\epsilon \in \mathbb{N}$ such that $p_m$ and $n_m$ in Algorithm 3.1 are $\epsilon$-increasing.
In order to prove this theorem, we need to prove the following propositions. We will always use the partitions \( \{x_i\}_{i=0}^{m+1} \) and \( \{y_i\}_{i=0}^{m+1} \) to be the uniform partitions of \([a, x]\) and \([a, y]\) respectively.

**Lemma 3.1.**

(1) If \( f \) is increasing on \([a, b]\) then \( p_m \) is increasing and we have \( p_m(x) = f(x) \) and \( n_m(x) = 0 \) for all \( x \in [a, b] \).

(2) If \( f \) is decreasing on \([a, b]\) then \( n_m \) is increasing and we have \( p_m(x) = f(a) \) and \( n_m(x) = f(a) - f(x) \) for all \( x \in [a, b] \).

**Proof.** This follows immediately from the way \( p_m \) and \( n_m \) are constructed in Algorithm 3.1.

In the following discussion, we will consider the function \( p_m \) and prove that this function is \( \epsilon \)-increasing. A similar argument can be made to prove that the function \( n_m \) is also \( \epsilon \)-increasing.

**Proposition 3.1.** For fixed \( x \in [a, b] \), the sequences \( (p_m(x)) \), and \( (n_m(x)) \) converge to \( p_J(x) \) and \( n_J(x) \), respectively. In other words, the sequences \( (p_m) \), and \( (n_m) \) converge pointwise to \( p_J \) and \( n_J \) respectively.

**Proof.** The proof of this proposition is straightforward. It suffices to note that

\[
\lim_{m \to \infty} \sum_{i=1}^{m+1} (f(x_i) - f(x_{i-1}))^+ = \sup_{\mathcal{P}[a,b]} \sum_{i=1}^{m+1} (f(x_i) - f(x_{i-1}))^+, \tag{3.2}
\]

and

\[
\lim_{m \to \infty} \sum_{i=1}^{m+1} (f(x_i) - f(x_{i-1}))^- = \sup_{\mathcal{P}[a,b]} \sum_{i=1}^{m+1} (f(x_i) - f(x_{i-1}))^-.
\]

**Proposition 3.2.** If \( x < y \), then for any \( \epsilon > 0 \), there exists an \( m \in \mathbb{N} \) such that

\[
|p_m(x) - p_m(y)| < 4C(m + 2)(y - x), \tag{3.2}
\]

\[
|n_m(x) - n_m(y)| < 4C(m + 2)(y - x). \tag{3.3}
\]

**Proof.** We prove (3.2) and notice that a similar argument holds for (3.3).

\[
|p_m(x) - p_m(y)| = \sum_{i=1}^{m+1} \left[ (f(x_i) - f(x_{i-1}))^+ - (f(y_i) - f(y_{i-1}))^+ \right]
\]

\[
\leq \sum_{i=1}^{m+1} |f(x_i) - f(x_{i-1})| + |f(y_i) - f(y_{i-1})|
\]

\[
\leq C \sum_{i=1}^{m+1} (x_i - x_{i-1}) + (y_i - y_{i-1})
\]

\[
\leq C \sum_{i=1}^{m+1} (|y_i - x_{i-1}| + |y_{i-1} - x_i|)
\]

\[
\leq 4C \sum_{i=1}^{m+1} (y_i - x_i)
\]

\[
\leq 4C(m + 2)(y - x).
\]
Proposition 3.3. For any $\epsilon > 0$, and $x \in [a, b]$, there exists an $m \in \mathbb{N}$ and an open neighborhood $I = (x - \delta, x + \delta)$ of $x$, where $\delta > 0$ such that

\begin{align}
|p_m(y) - p_J(y)| &\leq \epsilon \quad \forall y \in I, \\
|n_m(y) - n_J(y)| &\leq \epsilon \quad \forall y \in I.
\end{align}

Proof. Since the sequence of functions $(p_m)$ converges pointwise to $p_J$, we choose an $m$ such that

\begin{equation}
|p_m(x) - p_J(x)| < \epsilon/3.
\end{equation}

Choose $\delta$ such that $\delta < \epsilon/12C(m + 2)$, then we have from Proposition 3.2 that

\begin{equation}
|p_m(x) - p_m(y)| < \epsilon/3.
\end{equation}

Furthermore, we have

\begin{equation}
|p_J(x) - p_J(y)| < C\delta < \epsilon/3.
\end{equation}

Now we have

\begin{align*}
|p_m(y) - p_J(y)| &\leq |p_m(y) - p_m(x)| + |p_m(x) - p_J(x)| \\
&\quad + |p_J(x) - p_J(y)| \\
&< \epsilon.
\end{align*}

The inequality (3.5) can be proved similarly.

Theorem 3.2. The sequences $(p_m)$, and $(n_m)$ converge uniformly to $p_J$ and $n_J$, respectively.

Proof. Since the interval $[a, b]$ is compact, it follows immediately from Proposition 3.3 that $(p_m)$, and $(n_m)$ converge uniformly to $p_J$ and $n_J$, respectively.

After proving the previous propositions, and Theorem 3.2, we now have the tools to prove Theorem 3.1.

Proof of Theorem 3.1. Both functions $p_J$ and $n_J$ are increasing as we mentioned. For any $\epsilon > 0$ choose an $m$ such that $|p_m(x) - p_J(x)| \leq \epsilon/2$ for any $x \in [a, b]$. Thus, for any $x < y$,

\begin{equation*}
p_m(x) \leq p_J(x) + \epsilon/2 \leq p_J(y) + \epsilon/2 \leq p_m(y) + \epsilon.
\end{equation*}

This proves that the function $p_m$ is $\epsilon$-increasing.

If we choose $\epsilon$ to be the machine epsilon $\epsilon_{\text{mach}}$ of some computer, then the functions $p_m$ and $n_m$ will be increasing in this computer because if $x < y$ then $p_m(x) \leq p_m(y) + \epsilon_{\text{mach}} = p_m(y)$ and $n_m(x) \leq n_m(y) + \epsilon_{\text{mach}} = n_m(y)$.

REFERENCES


