Journal of Convex Analysis
Volume 11 (2004), No. 1, 131–146

Finite Convex Integration

D. Lambert
Unité d'Optimisation, Département de Mathématique,
Facultés Universitaires Notre Dame de la Paix, Namur, Belgium
delphine.lambert@math.fundp.ac.be

J.-P. Crouzeix
LIMOS, Université Blaise Pascal, Clermont-Ferrand, France

V. H. Nguyen
Unité d'Optimisation, Département de Mathématique,
Facultés Universitaires Notre Dame de la Paix, Namur, Belgium

J.-J. Strodiot
Unité d'Optimisation, Département de Mathématique,
Facultés Universitaires Notre Dame de la Paix, Namur, Belgium

Received January 21, 2003
Revised manuscript received September 24, 2003

This paper is concerned with the finite convex integration problem which for a given finite family of pairs of points $\{(x_i, x_i^*)\}$, consists in finding a convex function f such that $x_i^* \in \partial f(x_i)$ for all i.

1. Introduction

Assume that we are given a finite family $\{(x_i, x_i^*)\}_{i\in I} \subset X \times X^*$, where $X = X^* = \mathbb{R}^n$ (but the analysis works in a more general setting, for instance in Hilbert or Banach spaces). We are concerned with the following problem :

$$\text{Find a convex function } f : X \to \mathbb{R} \text{ such that } x_i^* \in \partial f(x_i) \ \forall\, i \in I \qquad (Int)$$

where $\partial f(x)$ denotes the Fenchel-subdifferential of f at x. Clearly, this problem is a finite version of the continuous integrability problem (for more details, see Rockafellar [1]): given a convex subset C of X and a point-to-set map $\Gamma : C \to X^*$:

$$\text{Find a convex function } f : C \to \mathbb{R} \text{ such that } \Gamma(x) \subset \partial f(x) \ \forall\, x \in C \qquad (Intc)$$

If f is a solution of one of these problems, its lsc-regularization (that is to say the greatest lower semi-continuous (lsc in short) function bounded from above by f) is also a solution. For this reason, we shall only retain the solutions which are lower semi-continuous.

Given a solution f of (Int), then any function of type $f + K$ where K is constant is a solution as well. Thus, we consider an additional initial-type condition: fix a couple (x_0, x_0^*) in the family and $\lambda_0 \in \mathbb{R}$, consider only the lower semi-continuous convex functions that satisfy

$$f(x_0) = \lambda_0 \text{ and } x_i^* \in \partial f(x_i) \ \ \forall\, i \in I.$$

We denote by $\mathcal{F}$ the set of all such functions. In both cases (continuous or finite), this solution set $\mathcal{F}$ is empty unless a cyclic-monotonicity type condition, that we call (CM),

ISSN 0944-6532 / $ 2.50

holds. However, when $\mathcal{F}$ is not empty, $\mathcal{F}$ is not necessarily a singleton as shown by the following example: $X = X^* = \mathbb{R}$, the family is $\{(0,0),(1,2),(-1,-2)\}$ and the initial condition is $f(0) = 0$. Then, the two following functions $f_1(x) = x^2$ and $f_2(x) = 2|x|$ belong to $\mathcal{F}$.

We shall show, in Sections 3 and 4, that there are two functions f^- and f^+ in $\mathcal{F}$ such that $f^- \leq f \leq f^+$ for all $f \in \mathcal{F}$. In Section 5, we describe algorithms for computing these two functions f^- and f^+. The gap between f^- and f^+ is studied in Section 6. In Section 7, we consider the case where the initial condition $f(x_0) = \lambda_0$ is replaced by the condition $f(\bar{x}) = \bar{\lambda}$ where $\bar{x}$ is an arbitrary point in X.

Finally, we conclude this paper with the particular case where $X = \mathbb{R}$.

The notation is the classical notation of convex analysis used in the tutorial book of Rockafellar [1].

2. The cyclic monotonicity property

A necessary condition for the existence of a solution of the continuous integrability problem ($Intc$) is that the point-to-set map Γ enjoys a cyclic monotonicity property (see for instance Rockafellar [1]). When transposed to the finite case, the property is as below.

Definition 2.1. The family $\{(x_i, x_i^*)\}_{i\in I}$ is said to be cyclically monotone if for all $j_0, j_1, \ldots, j_k, j_{k+1} \in I$ with $j_0 = j_{k+1}$, the following inequality holds:

$$\sum_{l=0,\cdots,k} \langle x_{j_l}^*, x_{j_{l+1}} - x_{j_l} \rangle \leq 0. \tag{CM}$$

This condition is necessary for the existence of solutions in our problem. The proof is the same as for the continuous problem and for the sake of completeness, we briefly reproduce it below. Assume that $f \in \mathcal{F}$. Then, because $x_i^* \in \partial f(x_i)$ for all $i \in I$, we have

$$\begin{array}{rcl} f(x_{j_1}) & \geq & f(x_{j_0}) + \langle x_{j_0}^*, x_{j_1} - x_{j_0} \rangle, \\ f(x_{j_2}) & \geq & f(x_{j_1}) + \langle x_{j_1}^*, x_{j_2} - x_{j_1} \rangle, \\ & \vdots & \\ f(x_{j_k}) & \geq & f(x_{j_{k-1}}) + \langle x_{j_{k-1}}^*, x_{j_k} - x_{j_{k-1}} \rangle, \\ f(x_{j_{k+1}}) & \geq & f(x_{j_k}) + \langle x_{j_k}^*, x_{j_{k+1}} - x_{j_k} \rangle. \end{array}$$

Condition (CM) follows by adding these inequalities and from the equality $x_{j_{k+1}} = x_{j_0}$.

3. Construction of a minimal function

Let us define the set $\mathcal{J}$ as the collection of all ordered subsets $J = \{j_0, \cdots, j_k\}$ of I such that $j_0 = 0$. We precise in this definition that $j_r \neq j_s$ when $r \neq s$.

Next, for each $J \in \mathcal{J}$, let us define, as in Rockafellar [1] p.238, the function on X

$$f_J(x) := \langle x_{j_0}^*, x_{j_1} - x_{j_0} \rangle + \cdots + \langle x_{j_{k-1}}^*, x_{j_k} - x_{j_{k-1}} \rangle + \langle x_{j_k}^*, x - x_{j_k} \rangle, \tag{1}$$

and then the function f^- as

$$f^-(x) := \lambda_0 + \max_{J\in\mathcal{J}} f_J(x). \tag{2}$$

Each function f_J is affine by construction. Hence, f^- is finite, convex and continuous on X since it is the maximum of a finite number of affine functions. In the following lemma, we are interested in the values of the function f^- at points x_i, $i \in I$.

Lemma 3.1. *Assume that condition (CM) holds. Let $i \in I$, then there exists $J = \{j_0, \cdots, j_k\} \in \mathcal{J}$ with $j_k = i$ such that*

$$f^-(x_i) = \lambda_0 + f_J(x_i). \tag{3}$$

In particular

$$f^-(x_0) = \lambda_0.$$

Proof. It results from (CM) that $f_J(x_0) \leq 0$ for all $J \in \mathcal{J}$. Hence $f^-(x_0) = \lambda_0$ because the maximum in (2) is reached for $J = \{0\}$. Next, for $i \neq 0$, let $J = \{j_0, \cdots, j_k\} \in \mathcal{J}$ be such that

$$f^-(x_i) = \lambda_0 + f_J(x_i). \tag{4}$$

If $i \notin J = \{j_0, \cdots, j_k\}$, change J into $J = \{j_0, \cdots, j_k, i\}$, then J still belongs to $\mathcal{J}$ and (4) holds as well. If $i \in J$ but $i \neq j_k$, i.e., if J is of the form $J = \{j_0, \cdots, j_{l-1}, j_l = i, j_{l+1}, \cdots, j_k\}$, then

$$f_J(x_i) = \sum_{r=0,\cdots,l-1} \langle x^*_{j_r}, x_{j_{r+1}} - x_{j_r} \rangle + \sum_{r=l,\cdots,k} \langle x^*_{j_r}, x_{j_{r+1}} - x_{j_r} \rangle$$

with $j_{k+1} = i$. Since $j_l = j_{k+1} = i$ and (CM) holds, the second sum is less or equal to 0. Change J into $J = \{j_0, \cdots, j_{l-1}, j_l = i\}$, then $J \in \mathcal{J}$ and (4) still holds. $\square$

In general, for a given $i \in I$, there does not exist a unique $J \in \mathcal{J}$ for which (3) holds. Moreover these J do not have necessarily the same cardinality. This motivates the following definition.

Definition 3.2. Given $i \in I$, we denote by $\mathcal{J}(i)$ the family of the ordered sets with the smallest cardinality $J = \{j_0, \cdots, j_k = i\} \in \mathcal{J}$ for which (3) holds. That cardinality, denoted by $r(i)$, is called the rank of i.

It follows that $r(0) = 1$ and $2 \leq r(i) \leq \operatorname{card}(I)$ for all $i \neq 0$. The following proposition is crucial for the construction of the function f^-.

Proposition 3.3. *Let $i \in I, i \neq 0$ and $J = \{j_0, j_1, \cdots, j_{k-1}, j_k = i\} \in \mathcal{J}(i)$. Then*

$$f^-(x_i) = f^-(x_{j_{k-1}}) + \langle x^*_{j_{k-1}}, x_i - x_{j_{k-1}} \rangle.$$

Moreover,

$$r(i) = r(j_{k-1}) + 1.$$

Proof. The assumptions imply that $r(i) = k + 1$ and

$$f^-(x_i) = [\lambda_0 + \sum_{r=0,\cdots,k-2} \langle x^*_{j_r}, x_{j_{r+1}} - x_{j_r} \rangle] + \langle x^*_{j_{k-1}}, x_i - x_{j_{k-1}} \rangle.$$

By definition of f^- and since $\{j_0, j_1, \cdots, j_{k-1}\} \in \mathcal{J}$, we have

$$f^-(x_i) \leq f^-(x_{j_{k-1}}) + \langle x^*_{j_{k-1}}, x_i - x_{j_{k-1}} \rangle. \tag{5}$$

Taking then some $\hat{J} = \{\hat{j}_0, \hat{j}_1, \cdots, \hat{j}_{r-1}, \hat{j}_r = j_{k-1}\} \in \mathcal{J}(j_{k-1})$, we consider two cases in succession.

- $i \notin \hat{J}$. Then $\{\hat{j}_0, \hat{j}_1, \cdots, \hat{j}_{r-1}, \hat{j}_r = j_{k-1}, \hat{j}_{r+1} = i\} \in \mathcal{J}$ and therefore

$$f^-(x_{j_{k-1}}) + \langle x^*_{j_{k-1}}, x_i - x_{j_{k-1}}\rangle = \lambda_0 + \sum_{l=0,\cdots,r} \langle x^*_{\hat{j}_l}, x_{\hat{j}_{l+1}} - x_{\hat{j}_l}\rangle \leq f^-(x_i). \tag{6}$$

  The equality follows from the definition of $\mathcal{J}(j_{k-1})$ and the inequality, from the definition of $f^-(x_i)$. Gathering (5) and (6), we obtain $f^-(x_i) = f^-(x_{j_{k-1}}) + \langle x^*_{j_{k-1}}, x_i - x_{j_{k-1}}\rangle$ and $r(i) \leq r(j_{k-1}) + 1$. Then, by definition of $r(i)$, $r(i) = r(j_{k-1}) + 1$.

- $i \in \hat{J}$. Then $\hat{J}$ is of the form

$$\hat{J} = \{\hat{j}_0, \cdots, \hat{j}_{s-1}, \hat{j}_s = i, \hat{j}_{s+1}, \cdots, \hat{j}_r = j_{k-1}\}.$$

  In order to prove that this case never occurs, set

$$A = \lambda_0 + \sum_{l=0,\cdots,s-1} \langle x^*_{\hat{j}_l}, x_{\hat{j}_{l+1}} - x_{\hat{j}_l}\rangle$$

  and

$$B = \sum_{l=s,\cdots,r} \langle x^*_{\hat{j}_l}, x_{\hat{j}_{l+1}} - x_{\hat{j}_l}\rangle + \langle x^*_{j_{k-1}}, x_i - x_{j_{k-1}}\rangle.$$

  Inequality (5) implies $f^-(x_i) \leq A + B$. Condition (CM) implies $B \leq 0$. Finally, $A \leq f^-(x_i)$ because the ordered set $\{\hat{j}_0, \cdots, \hat{j}_{s-1}, \hat{j}_s = i\}$ belongs to $\mathcal{J}$. It results first that $B = 0$ and $A = f^-(x_i)$; next

$$r(i) \leq s + 1 < k + 1 = r(i)$$

  which is not possible.

The proof is complete. □

Now, we can state the main result of this section.

Theorem 3.4. *Assume that Condition (CM) holds for the family $\{(x_i, x_i^*)\}_{i\in I}$. Then f^- satisfies the following properties:*

i) f^- *is convex and continuous on* X*;*
ii) $f^-(x_0) = \lambda_0$*;*
iii) $x_i^* \in \partial f^-(x_i) \quad \forall\, i \in I$*;*
iv) $f^-(x) = \max_{i\in I}[\, f^-(x_i) + \langle x_i^*, x - x_i\rangle \,] \quad \forall\, x \in X$*;*
v) $g(x) \geq f^-(x)$ *for all* $g \in \mathcal{F}$ *and* $x \in X$*.*

Proof. i) and ii) have already been shown.

iii) Let $x \in X$ and $i \in I$. Take $J = \{j_0, j_1, \cdots, j_k = i\} \in \mathcal{J}(i)$. Using (1), (2) and the definition of $\mathcal{J}(i)$, we obtain

$$\begin{aligned} f^-(x) &\geq \lambda_0 + \sum_{r=0,\cdots,k-1} \langle x^*_{j_r}, x_{j_{r+1}} - x_{j_r}\rangle + \langle x_i^*, x - x_i\rangle, \\ f^-(x) &\geq f^-(x_i) + \langle x_i^*, x - x_i\rangle. \end{aligned}$$

Thus, $x_i^* \in \partial f^-(x_i)$.

iv) Given $x \in X$, let $J = \{j_0, j_1, \cdots, j_k\} \in \mathcal{J}$ be such that

$$\begin{aligned} f^-(x) &= \lambda_0 + f_J(x) \\ &= [\lambda_0 + \sum_{r=0,\cdots,k-1} \langle x_{j_r}^*, x_{j_{r+1}} - x_{j_r} \rangle] + \langle x_{j_k}^*, x - x_{j_k} \rangle \\ &\leq f^-(x_{j_k}) + \langle x_{j_k}^*, x - x_{j_k} \rangle. \end{aligned}$$

Combining these relations with iii), we obtain iv).

v) Assume that $g \in \mathcal{F}$. First, let us prove that $g(x_i) \geq f^-(x_i)$ for all i. We proceed by induction on the ranks of the elements of I. If $r(i) = 1$, i.e., if $i = 0$ then, by definition, $g(x_0) = f^-(x_0)$ and the inequality holds. Assume that it holds for all i of rank less or equal to p and let any i of rank $p + 1$. Let $J = \{j_0, j_1, \cdots, j_k, j_{k+1} = i\} \in J(i)$. Then, since $x_{j_k}^* \in \partial g(x_{j_k})$, we have

$$g(x_i) \geq g(x_{j_k}) + \langle x_{j_k}^*, x_i - x_{j_k} \rangle.$$

Proposition 3.3 implies that, on the first hand, j_k is of rank p and thereby $g(x_{j_k}) \geq f^-(x_{j_k})$ and on the other hand $f^-(x_i) = f^-(x_{j_k}) + \langle x_{j_k}^*, x_i - x_{j_k} \rangle$. Hence v) holds for $x = x_i$.

Next, let any $x \in X$. By iv) we know that there exists $i \in I$ such that $f^-(x) = f^-(x_i) + \langle x_i^*, x - x_i \rangle$. On another hand $g(x) \geq g(x_i) + \langle x_i^*, x - x_i \rangle$ because $x_i^* \in \partial g(x_i)$. Hence $g(x) \geq f^-(x)$ because $g(x_i) \geq f^-(x_i)$. □

Thus, in view of iv), the minimal function f^- of $\mathcal{F}$ is completely determined when the values $f^-(x_i)$ are known at points x_i. The following proposition shows that these values are the optimal solutions of a linear program.

Proposition 3.5. *The quantities $f^-(x_i)$ are the optimal solutions of the following linear program:*

$$\min_{\lambda} [\sum_{i \in I, i \neq 0} \lambda_i \ : \ \lambda_j - \lambda_i \geq \langle x_i^*, x_j - x_i \rangle \quad \forall\, i, j \in I]. \tag{LP^-}$$

(recall that λ_0 is fixed.)

Proof. i) Let $\bar{\lambda} \in \mathbb{R}^I$ be defined by $\bar{\lambda}_i = f^-(x_i)$ for all $i \in I$. Since f^- is convex and $x_i^* \in \partial f^-(x_i)$, then $\bar{\lambda}$ is a feasible solution of (LP^-).

ii) Next, let λ be a feasible solution of (LP^-). Define the following function p on X:

$$p(x) = \max_{i \in I} \{\lambda_i + \langle x_i^*, x - x_i \rangle\}.$$

By construction, p is finite and convex on X. Moreover,

$$p(x_j) \geq \lambda_j + \langle x_j^*, x_j - x_j \rangle = \lambda_j.$$

Assume, for contradiction, that $p(x_j) \neq \lambda_j$ for some j. Then, there exists an index i such that

$$p(x_j) = \lambda_i + \langle x_i^*, x_j - x_i \rangle > \lambda_j$$

in contradiction with λ being feasible. Thus, $p(x_j) = \lambda_j$ for all j. Next, for all $x \in X$

$$p(x) \geq \lambda_i + \langle x_i^*, x - x_i \rangle = p(x_i) + \langle x_i^*, x - x_i \rangle$$

and therefore $p \in \mathcal{F}$. Finally, since f^- is minimal in $\mathcal{F}$, $p \geq f^-$ and in particular

$$\sum_{i \in I} \lambda_i = \sum_{i = \in I} p(x_i) \geq \sum_{i \in I} f^-(x_i) = \sum_{i \in I} \bar{\lambda}_i.$$

□

Thus, the construction of f^- is equivalent to finding the optimal solution of the linear program (LP^-). Since the existence of f^- depends on Condition (CM), that condition is related to the feasibility of (LP^-).

Proposition 3.6. (CM) *holds if and only if* (LP^-) *is feasible.*

Proof. i) Assume that λ is feasible for (LP^-) and let $J = \{j_0, \cdots, j_k, j_{k+1}\} \subset I$ with $j_{k+1} = j_0$. Then,

$$\begin{array}{rcl} \lambda_{j_1} & \geq & \lambda_{j_0} + \langle x_{j_0}^*, x_{j_1} - x_{j_0} \rangle, \\ \lambda_{j_2} & \geq & \lambda_{j_1} + \langle x_{j_1}^*, x_{j_2} - x_{j_1} \rangle, \\ & \vdots & \\ \lambda_{j_{k+1}} & \geq & \lambda_{j_k} + \langle x_{j_k}^*, x_{j_{k+1}} - x_{j_k} \rangle. \end{array}$$

Adding these inequalities, and using the fact that $\lambda_{j_{k+1}} = \lambda_{j_0}$, we get immediately Condition (CM).

ii) Next, assume that (CM) holds. Then f^- exists. Take, as in the proof of the last proposition, $\bar{\lambda}_i = f^-(x_i)$ for all i. We have seen that such $\bar{\lambda}$ is a feasible (and even optimal) solution for (LP^-). □

Although a classical simplex algorithm for linear programming can be used to compute the optimal solution of (LP^-), we shall design a special algorithm in Section 5.

4. Construction of a maximal function

We begin this section by noticing that the variables x_i and x_i^* play a symmetric role in Condition (CM). This is the object of the following proposition:

Proposition 4.1. *The family* $\{(x_i, x_i^*)\}_{i \in I}$ *considered as a subset of* $X \times X^*$ *satisfies Condition* (CM) *if and only if the family* $\{(x_i^*, x_i)\}_{i \in I}$ *considered as a subset of* $X^* \times X$ *satisfies* (CM) *too.*

Proof. Assume that (CM) holds in $X \times X^*$. Let $J = \{j_0, \cdots, j_k, j_{k+1}\} \subset I$ with $j_{k+1} = j_0$. We must prove that the quantity

$$A = \sum_{r=0,\cdots,k} \langle x_{j_{r+1}}^* - x_{j_r}^*, x_{j_r} \rangle$$

is nonnegative. Rearranging the terms, we have

$$\begin{aligned} A &= \langle x^*_{j_1}, x_{j_0}\rangle + \langle x^*_{j_2}, x_{j_1}\rangle + \cdots + \langle x^*_{j_{k+1}}, x_{j_k}\rangle \\ &\quad -\langle x^*_{j_0}, x_{j_0}\rangle - \langle x^*_{j_1}, x_{j_1}\rangle - \cdots - \langle x^*_{j_k}, x_{j_k}\rangle. \end{aligned}$$

Take $I = \{i_0, \cdots, i_k, i_{k+1}\} \subset I$ where $i_s = j_{k+1-s}$ for $s = 0, 1 \cdots, k+1$. Then $i_{k+1} = i_0$ and

$$\begin{aligned} A &= \langle x^*_{i_0}, x_{i_1}\rangle + \langle x^*_{i_1}, x_{i_2}\rangle + \cdots + \langle x^*_{i_k}, x_{i_{k+1}}\rangle \\ &\quad -\langle x^*_{i_0}, x_{i_0}\rangle - \langle x^*_{i_1}, x_{i_1}\rangle - \cdots - \langle x^*_{i_k}, x_{i_k}\rangle. \end{aligned}$$

Condition (CM) implies $A \leq 0$. By symmetry, (CM) for $X^* \times X$ implies (CM) for $X \times X^*$. □

Assume that (CM) holds (it is no more necessary to precise if it holds for $X \times X^*$ or for $X^* \times X$). Set $\lambda_0^* = \langle x_0, x_0^*\rangle - \lambda_0$ and define $\mathcal{H}$ as the collection of lower semi-continuous convex functions on X^* such that

$$h(x_0^*) = \lambda_0^* \text{ and } x_i \in \partial h(x_i^*) \quad \forall\, i \in I.$$

It is easy to see that a function f belongs to $\mathcal{F}$ if and only if its Fenchel-conjugate belongs to $\mathcal{H}$. On the other hand, Theorem 3.4 implies the existence of a minimal function h^- in $\mathcal{H}$. Define the function f^+ as the Fenchel-conjugate of h^-. Then we have the following result.

Proposition 4.2. *f^+ is maximal in $\mathcal{F}$, i.e., $f \leq f^+$ for all $f \in \mathcal{F}$.*

Proof. We have already seen that f^+ belongs to $\mathcal{F}$. Let any $f \in \mathcal{F}$, then its conjugate f^* belongs to $\mathcal{H}$ and therefore $h^- \leq f^*$. Passing to the conjugates we obtain $(h^-)^* \geq f$. □

The next theorem shows that f^+ is completely characterized by its values at points x_i.

Theorem 4.3. *The maximal function f^+ of $\mathcal{F}$ is given by*

$$f^+(x) = \inf_r\{\sum r_i\mu_i : \; r \in \mathbb{R}^I,\; r_i \geq 0, \sum r_i = 1, \sum r_i x_i = x\}$$

where $\mu_i = f^+(x_i) = \langle x_i, x_i^*\rangle - \lambda_i^*$ and $\lambda_i^* = h^-(x_i^*)$.

In particular, $f^+(x) = +\infty$ for all x not in the convex hull of the points x_i.

Proof. Recall that, for a lower semi-continuous convex function f,

$$x^* \in \partial f(x) \Longleftrightarrow f(x) + f^*(x^*) = \langle x, x^*\rangle.$$

Hence, if $\mu_i = f^+(x_i)$ and $\lambda_i^* = h^-(x_i^*)$ and because h^- is the conjugate of f^+,

$$\lambda_i^* + \mu_i = \langle x_i, x_i^*\rangle.$$

Next, since h^- is minimal in $\mathcal{H}$ and in view of Theorem 3.4,

$$\begin{aligned} h^-(x^*) &= \max_{i\in I}\,[\; h^-(x_i^*) + \langle x_i, x^* - x_i^*\rangle \;] \\ &= \min_t\,[t \;:\; t \in \mathbb{R},\; \lambda_i^* + \langle x_i, x^* - x_i^*\rangle \leq t \;\forall\, i \in I\;]. \end{aligned}$$

Hence,

$$\begin{aligned} f^+(x) &= \sup_{x^* \in X^*} [\langle x, x^* \rangle - h^-(x^*)] \\ &= \sup_{x^*, t} [\langle x, x^* \rangle - t : \langle x_i, x^* \rangle - t \le \langle x_i, x_i^* \rangle - \lambda_i^* \ \forall\, i \in I] \\ &= \inf_r [\sum r_i \mu_i : r \in \mathbb{R}^I, r_i \ge 0, \sum r_i = 1, \sum r_i x_i = x \ \forall\, i \in I\ ]. \end{aligned}$$

Clearly, f^+ takes the value $+\infty$ outside the convex hull of points x_i. □

We have seen that the quantities $\lambda_i = f^-(x_i)$ correspond to the optimal solution of a linear program. This is also the case for the quantities $\mu_i = f^+(x_i)$.

Proposition 4.4. *The quantities μ_i are the optimal solutions of the following linear program:*

$$\max_\mu [\sum_{i \in I, i \ne 0} \mu_i \ : \ \mu_j - \mu_i \ge \langle x_i^*, x_j - x_i \rangle \quad \forall\, i, j \in I\,]. \tag{LP^+}$$

where $\mu_0 = \lambda_0$ is fixed.

Proof. Because h^- is minimal in $\mathcal{H}$, the quantities $\lambda_i^* = h^-(x_i^*)$ are the optimal solutions of the linear program

$$\min_{\lambda^*} [\sum_{i \in I, i \ne 0} \lambda_i^* \ : \ \lambda_i^* - \lambda_j^* \ge \langle x_i^* - x_j^*, x_j \rangle \quad \forall\, i, j \in I\,]. \tag{LP^+}$$

where $\lambda_0^* = \langle x_0^*, x_0 \rangle - \lambda_0$ is fixed. On the other hand, $\lambda_i^* + \mu_i = \langle x_i^*, x_i \rangle$ for all i, hence replacing λ_i^* by its value in term of μ_i, the problem is equivalent to

$$A = \min_\mu [\sum_{i \in I, i \ne 0} (\langle x_i, x_i^* \rangle - \mu_i) : \mu_j - \mu_i \ge \langle x_i^*, x_j - x_i \rangle \quad \forall i, j \in I\,],$$

i.e.,

$$A = \sum_{i \in I, i \ne 0} \langle x_i, x_i^* \rangle - \max_\mu [\sum_{i \in I, i \ne 0} \mu_i : \mu_j - \mu_i \ge \langle x_i^*, x_j - x_i \rangle \quad \forall i, j \in I\,].$$

As the first term is constant in μ, the result follows. □

It is worth noticing that the feasibility domains of (LP^+) and (LP^-) are the same. Here again, we shall give, in the next section, a specially designed algorithm for computing the values μ_i.

Thus, we have shown in this section and the previous one the existence of a maximal and a minimal function in $\mathcal{F}$. We resume that below.

Theorem 4.5. *Assume that the family $\{(x_i, x_i^*)\}_{i \in I}$ satisfies Condition (CM). Then there exist two functions f^-, $f^+ \in \mathcal{F}$ such that*

$$f^- \le f \le f^+ \quad \textit{for all } f \in \mathcal{F}.$$

Furthermore $\mathrm{dom}(f^-) = X$ *and* $\mathrm{dom}(f^+) = \mathrm{co}(x_i, i \in I)$.

Remark 4.6. A geometrical remark on f^- and f^+.

In view of Theorems 3.4 and 4.3, the epigraphs of f^- and f^+ are two convex polyhedral sets. The first one is formulated as an intersection of a finite collection of closed convex half-spaces while the second one as a finitely generated convex set, the two dual formulations of convex polyhedra.

5. Algorithms

We have seen that f^- is completely determined when the quantities $\lambda_i = f^-(x_i)$ are known. We have also seen that these λ_i are the optimal solutions of a linear program and thereby they can be computed by the simplex algorithm. However, we propose a specially designed algorithm based on the result on the ranks in Proposition 3.3. In order to ease the understanding of the algorithm, explanations are included in the description of the algorithm.

An algorithm for computing the values $f^-(x_i)$.

The data are the family $\{(x_i, x_i^*)\}_{i\in I}$ with $I = \{0, 1, \cdots, n\}$ and the value λ_0. The algorithm computes the values $\lambda_i = f^-(x_i)$ for $i \neq 0$ when Condition (CM) holds or says that Condition (CM) does not hold.

Step k=1: Initialization

- **For** $i = 1, \cdots, n$ **do**

$$\lambda_i = \lambda_0 + \langle x_0^*, x_i - x_0\rangle \text{ and } m(i) = 1.$$

- **Do** $c(1) = 1$.
- **Do** $k = 2$**, go to the next step.**

Explanations:

- $m(i) = 1$ *means that* λ_i *has been modified during the step.*
- $c(1) = 1$ *means that some* λ_i *have been modified during step* $k = 1$.
- *In case where* (CM) *holds, the inequality* $\lambda_i \leq f^-(x_i)$ *holds for all* i *with equality for all* i *of rank less or equal to* 2.

Step k (for $k \geq 2$)

At the beginning of the step, in case where (CM) *holds, then the inequality* $\lambda_i \leq f^-(x_i)$ *holds for all* i *with equality for all* i *of rank less or equal to* k.

- **Do** $c(k) = 0$.
- **For** $i = 1, 2, \cdots, n$ **do**
  - **If** $m(i) = 0$**: go to the next** i.
  - **If** $m(i) = 1$ **and** $\lambda_i + \langle x_i^*, x_0 - x_i\rangle > \lambda_0$**: STOP, (CM) does not hold.**
  - **If** $m(i) = 1$ **and** $\lambda_i + \langle x_i^*, x_0 - x_i\rangle \leq \lambda_0$**:**
    * **Do** $m(i) = 0$.
    * **For all** $j = 1, \cdots, n$ **such that** $\lambda_j < \lambda_i + \langle x_i^*, x_j - x_i\rangle$
      - **Do** $\lambda_j = \lambda_i + \langle x_i^*, x_j - x_i\rangle$.
      - **Do** $m(j) = 1$ **and** $c(k) = 1$.
- **If** $c(k) = 0$**: STOP, (CM) holds,** $f^-(x_i) = \lambda_i$ **for all** i.

- **If $c(k) = 1$ and $k = n + 2$: STOP, (CM) does not hold.**
- **If $c(k) = 1$ and $k < n + 2$: Do $k = k + 1$, and go to step k.**

Explanations:

- *One works only on those i which have been modified since their last passage ($m(i) = 1$).*
- *In case where (CM) holds, the quantity λ_j stays less or equal to $f^-(x_j)$.*
- *In view of the last item, it is clear that if $\lambda_i + \langle x_i^*, x_0 - x_i\rangle > \lambda_0$, then it is not possible to construct f^-. Hence (CM) does not hold.*
- *At the end of step k, $\lambda_j = f^-(x_j)$ for all i of rank less or equal to $k+1$.*
- *If $c(k) = 0$, then no modifications on the λ_i have been done during step k. This means that $\lambda_j \geq \lambda_i + \langle x_i^*, x_j - x_i\rangle$ for all $i, j = 0, 1, \cdots, n$. λ is feasible for (LP-) and since $\lambda_j \leq f^-(x_j)$ we have got the optimal solution.*
- *Since the rank of each x_i cannot exceed the cardinality of I, the condition $k = n + 2$ means that it is not possible to construct f^-.*

The finite convergence of the algorithm, when (CM) holds, follows from Proposition 3.3.

Because the symmetry between the minimal functions f^- and h^- of $\mathcal{F}$ and $\mathcal{H}$ respectively, the same algorithm can be used to compute the values $\lambda_i^* = h^-(x_i^*)$. Then the values $\mu_i = f^+(x_i) = \langle x_i^*, x_i\rangle - \lambda_i^*$ are determined. Still, we can resume these two steps in one algorithm formulated directly in terms of μ_i.

An algorithm for computing the values $f^+(x_i)$.

<u>**Step k=1: Initialization**</u>

- **For $i = 1, \cdots, n$ do**

$$\mu_i = \mu_0 + \langle x_i^*, x_i - x_0\rangle \textbf{ and } m(i) = 1.$$

- **Do $c(1) = 1$.**
- **Do $k = 2$, go to the next step.**

<u>**Step k** (for $k \geq 2$)</u>

- **Do $c(k) = 0$.**
- **For $i = 1, 2, \cdots, n$ do**
  - **If $m(i) = 0$: go to the next i.**
  - **If $m(i) = 1$ and $\mu_i + \langle x_0^*, x_0 - x_i\rangle < \mu_0$: STOP, (CM) does not hold.**
  - **If $m(i) = 1$ and $\mu_i + \langle x_0^*, x_0 - x_i\rangle \geq \mu_0$:**
    - **Do $m(i) = 0$.**
    - **For all $j = 1, \cdots, n$ such that $\mu_j > \mu_i + \langle x_j^*, x_j - x_i\rangle$:**
      - **Do $\mu_j = \mu_i + \langle x_j^*, x_j - x_i\rangle$.**
      - **Do $m(j) = 1$ and $c(k) = 1$.**
- **If $c(k) = 0$: STOP, (CM) holds, $f^+(x_i) = \langle x_i^*, x_i\rangle - \lambda_i^*$ for all i.**
- **If $c(k) = 1$ and $k = n + 2$: STOP, (CM) does not hold.**
- **If $c(k) = 1$ and $k < n + 2$: Do $k = k + 1$, and go to step k.**

6. The gap between f^- and f^+

In this section, we turn our interest to the gap $[f^-(x), f^+(x)]$. Because $f^+(x) = +\infty$ outside $C = \mathrm{co}(x_i, i \in I)$, we only consider the case where $x \in C$. We start the study with the particular case where x is one of the x_i.

Let us define

$$\epsilon_i = f^+(x_i) - f^-(x_i) = \mu_i - \lambda_i.$$

By construction, $\epsilon_0 = 0$.

Proposition 6.1. *For all $i, j \in I$,*

$$|\epsilon_i - \epsilon_j| \leq \langle x_i^* - x_j^*, x_i - x_j \rangle.$$

Proof. We have seen that

$$\begin{array}{rcl} \langle x_j^*, x_i - x_j \rangle & \leq & \lambda_i - \lambda_j \\ \langle x_i^*, x_j - x_i \rangle & \leq & \mu_j - \mu_i. \end{array}$$

Hence,

$$\langle x_j^* - x_i^*, x_i - x_j \rangle \leq \epsilon_j - \epsilon_i.$$

Next, by symmetry,

$$|\epsilon_i - \epsilon_j| \leq \langle x_i^* - x_j^*, x_i - x_j \rangle.$$

□

Since $\epsilon_0 = 0$, it follows directly that for all i

$$\epsilon_i \leq \langle x_i^* - x_0^*, x_i - x_0 \rangle. \tag{7}$$

However, a better upper bound can be obtained. For that, define $\hat{\epsilon}_0 = 0$ and for $i \neq 0$,

$$\hat{\epsilon}_i = \min_J \left[\sum_{k=0,\cdots,p} \langle x_{j_{k+1}}^* - x_{j_k}^*, x_{j_{k+1}} - x_{j_k} \rangle \ : \ \begin{array}{c} J = \{j_0, j_1, \cdots, j_{p+1}\} \\ j_0 = 0, \ j_{p+1} = i \end{array} \right]. \tag{8}$$

Next, set $\hat{\epsilon} = \max_i \hat{\epsilon}_i$. The following upper bound is a consequence of Proposition 6.1.

Proposition 6.2. *For all $i \in I$, $\quad \epsilon_i \leq \hat{\epsilon}_i \leq \hat{\epsilon}$.*

Proof. Let $i \in I$ and let $J = \{j_0, \cdots, j_{p+1}\}$ with $j_0 = 0$ and $j_{p+1} = i$. Then, since $\epsilon_0 = 0$,

$$\begin{array}{rcl} \epsilon_i & \leq & |\epsilon_i - \epsilon_{j_p}| + \cdots + |\epsilon_{j_1} - \epsilon_0| \\ & \leq & \displaystyle\sum_{k=0,\cdots,p} \langle x_{j_{k+1}}^* - x_{j_k}^*, x_{j_{k+1}} - x_{j_k} \rangle. \end{array}$$

Since J is arbitrary, the proof is complete. □

An algorithm similar to those described in Section 5 can be designed for computing the values of the upper bounds $\hat{\epsilon}_i$ of ϵ_i. These upper bounds are more efficient than those given by the inequalities (7). To see that, consider the following example:

Example 6.3. $X = \mathbb{R}$, the family is $\{(\frac{i}{q}, \frac{i}{q})\}_{i=0,1,\cdots,q}$.

The upper bound obtained from inequality (7) is $\epsilon_q \leq 1$, and from Proposition 6.2, $\epsilon_q \leq \hat{\epsilon}_q = \frac{1}{q}$.

Next, we consider some $x \in C$. Then x is a convex combination of the points x_i, i.e., there are $K \subseteq I$ and $t_k \geq 0$, $k \in K$ such that

$$x = \sum_{k \in K} t_k x_k, \text{ and } \sum_{k \in K} t_k = 1.$$

Define C_k as the convex hull of points $x_k,\ k \in K$.

For $x \in C_k$, it follows from Theorems 3.4 and 4.3 that

$$\begin{aligned} f^+(x) - f^-(x) &\leq \min_{i \in K} [\sum_{k \in K} t_k \mu_k - \lambda_i - \langle x_i^*, x - x_i \rangle], \\ &\leq \min_{i \in K} [\sum_{k \in K} t_k (\mu_k - \lambda_i - \langle x_i^*, x_k - x_i \rangle)]. \end{aligned}$$

On the other hand,

$$\begin{aligned} \mu_k + \langle x_k^*, x_i - x_k \rangle &\leq \mu_i, \\ \mu_k + \langle x_i^*, x_i - x_k \rangle &\leq \mu_i + \langle x_i^* - x_k^*, x_i - x_k \rangle. \end{aligned}$$

It follows that

$$f^+(x) - f^-(x) \leq \min_{i \in K} [\sum_{k \in K} t_k \left(\mu_i - \lambda_i + \langle x_i^* - x_k^*, x_i - x_k \rangle\right)]$$

and finally

$$f^+(x) - f^-(x) \leq \min_{i \in K} [\, \epsilon_i + \sum_{k \in K} t_k \langle x_i^* - x_k^*, x_i - x_k \rangle]. \tag{9}$$

Let us define

$$\sigma_K = \max_{i,k \in K} \langle x_i^* - x_k^*, x_i - x_k \rangle.$$

Then, we have the following upper bound for $f^+(x) - f^-(x)$:

Proposition 6.4. *For all $x \in C_k$,*

$$f^+(x) - f^-(x) \leq \sigma_K + \min_{i \in K} \epsilon_i.$$

Proof. The proof is a direct consequence of inequality (9). □

When $X = \mathbb{R}^n$, for any $x \in C$, there exists some K with cardinality less or equal to $n+1$ such that $x \in C_k$. As in finite element methods, C can be recovered by a finite collection of such C_k. By increasing the cardinality of I, it is possible to decrease the size of the subsets C_k and thereby the values of σ_K leading to a reduction of the gap $[f^-(x), f^+(x)]$.

7. Changing the initial condition

In the previous study, we have considered for initial condition the equality $f(x_0) = \lambda_0$ with λ_0 fixed. We change this condition into the condition $f(\bar{x}) = \bar{\lambda}$ where $\bar{x}$ is an arbitrary point in X. Namely we consider the problem:

$$\text{Find a lsc convex function } g \text{ s.t. } g(\bar{x}) = \bar{\lambda} \text{ and } x_i^* \in \partial g(x_i) \ \forall\, i \in I$$

where $\bar{\lambda} \in \mathbb{R}$ is fixed. We denote by $\mathcal{G}$ the set of such functions.

It is clear that, given $g \in \mathcal{G}$, the function f defined by

$$f(x) = g(x) + \lambda_0 - g(x_0)$$

belongs to $\mathcal{F}$. Conversely, given $f \in \mathcal{F}$, the function g defined by

$$g(x) = f(x) + \bar{\lambda} - f(\bar{x})$$

belongs to $\mathcal{G}$. It follows that Condition (CM) is a necessary and sufficient condition for $\mathcal{G}$ to be not empty.

Through this section, we assume that (CM) holds. Let us define the following functions $g^+, g^- : X \to [-\infty, +\infty]$ by

$$g^+(x) = \sup_{g \in \mathcal{G}} g(x) \quad \text{and} \quad g^-(x) = \inf_{g \in \mathcal{G}} g(x).$$

Take some $g \in \mathcal{G}$, and let f be defined by $f(x) = g(x) + \lambda_0 - g(x_0)$. Since $f \in \mathcal{F}$, the following inequalities hold for all $x \in X$:

$$f^-(x) + g(x_0) - \lambda_0 \leq g(x) \leq f^+(x) + g(x_0) - \lambda_0.$$

On the other hand, since $x_0^* \in \partial g(x_0)$,

$$g(x_0) + \langle x_0^*, \bar{x} - x_0 \rangle \leq g(\bar{x}) = \bar{\lambda}.$$

It follows that for all $x \in X$,

$$g(x) \leq f^+(x) + (\bar{\lambda} - \lambda_0) - \langle x_0^*, \bar{x} - x_0 \rangle. \tag{10}$$

We can now prove the following result.

Theorem 7.1. *Assume that (CM) holds. Then $g^+ \in \mathcal{G}$. Furthermore the domain of g^+ is the convex hull of the points $\bar{x}$ and x_i, $i \in I$.*

Proof. g^+ is convex and lsc as the supremum of such functions. It is clear that Inequality (10) holds for g^+ as well. Hence, $\mathrm{dom}(g^+) \supseteq \mathrm{dom}(f^+) = \mathrm{co}(\bar{x}; x_i, i \in I)$. It is clear that $g^+(\bar{x}) = \bar{\lambda}$ and, because $g^+(x_i)$ is finite, that for all $x \in X$ and $i \in I$

$$g^+(x_i) + \langle x_i^*, x - x_i \rangle \leq g^+(x).$$

Hence $g^+ \in \mathcal{G}$. Finally, define $\hat{g}$ by $\hat{g}(x) = g^+(x)$ if $x \in \mathrm{co}(\bar{x}; x_i, i \in I)$ and $\hat{g}(x) = +\infty$ otherwise. Then $\hat{g}$ belongs to $\mathcal{G}$ and therefore coincides with g^+. It follows that $\mathrm{dom}(g^+) = \mathrm{co}(\bar{x}; x_i, i \in I)$. $\square$

The function g^- may not be convex when $\bar{x}$ is not one of the points x_i, as shown by the following examples:

Example 7.2. $X = \mathbb{R}$, the family is $\{(-1,-1),(1,1)\}$, the initial condition is $g(0) = 0$.

Then, it is easily seen that

$$g^+(x) = \begin{cases} |x| & \text{if } |x| \le 1, \\ +\infty & \text{otherwise} \end{cases} \quad \text{and } g^-(x) = \begin{cases} -|x| & \text{if } |x| \le 1, \\ -\infty & \text{otherwise} \end{cases}$$

Example 7.3. $X = \mathbb{R}$, the family is $\{(1,1),(2,2)\}$, the initial condition is $g(0) = 0$.

Here, we obtain

$$g^+(x) = \begin{cases} x & \text{if } 0 \le x \le 1, \\ 2x-1 & \text{if } 1 \le x \le 2, \\ +\infty & \text{otherwise} \end{cases} \quad \text{and } g^-(x) = \begin{cases} x & \text{if } x \le 0, \\ -\infty & \text{otherwise} \end{cases}$$

The function g^+ is completely determined when the quantities $\xi_i = g^+(x_i)$ are known. Indeed, since g^+ is maximal in $\mathcal{G}$, its epigraph is the polyhedral convex set of $X \times \mathbb{R}$ generated from the points $(\bar{x}, \bar{\lambda})$, (x_i, ξ_i), $i \in I$ and the direction $(0,1)$. Then, it is easily seen that the quantities ξ_i are the optimal solutions of the linear program:

$$\max \left[\sum_{i \in I} \xi_i \ : \ \begin{array}{ll} \xi_j - \xi_i \ge \langle x_i^*, x_j - x_i \rangle & \forall\, i,j \in I \\ \bar{\lambda} - \xi_i \ge \langle x_i^*, \bar{x} - x_i \rangle & \forall\, i \in I \end{array} \right].$$

8. The particular case of one real variable

Within this section, we assume that $X = X^* = \mathbb{R}$. We shall show that in this case the cyclic monotonicity reduces to the following classical monotonicity condition

$$\langle x_i^* - x_j^*, x_i - x_j \rangle \ge 0 \text{ for all } i,j \in I. \tag{M}$$

The set $\mathbb{R}$, unlike $\mathbb{R}^n$, is totally ordered so that we can assume that

$$x_0 \le x_1 \le \cdots \le x_p$$

where $I = \{0, 1, \cdots, p\}$. Then (M) is equivalent to the condition

$$x_i^* \le x_{i+1}^* \quad \text{for } i = 0, 1, \cdots, p-1.$$

We consider for initial condition $g(\bar{x}) = \bar{\lambda}$ where $\bar{x} = x_k$ with $k \in \{0, 1, \cdots, p\}$. The case where $\bar{x}$ is not one of the points x_i is left to the reader (then g^- can be not convex). For simplicity, we set $x_{-1} = x_{-1}^* = -\infty$ and $x_{p+1} = x_{p+1}^* = +\infty$. Assume that (M) holds and $g \in \mathcal{G}$, then necessarily

$$\partial g(x) \subseteq \begin{cases} [x_i^*, x_{i+1}^*] & \text{if} \quad x_i < x < x_{i+1}, \quad i = -1, 0, \cdots, p \\ [x_{i-1}^*, x_{i+1}^*] & \text{if} \qquad x = x_i, \qquad i = 0, 1, \cdots, p \end{cases} \tag{11}$$

$$\partial g^-(x) = \begin{cases} x^*_{i+1} & \text{if } \; x_i < x < x_{i+1} \leq x_k, \\ x^*_i & \text{if } \; x_k \leq x_i < x < x_{i+1}, \end{cases} \tag{12}$$

$$\partial g^+(x) = \begin{cases} x^*_{i} & \text{if } \; x_i < x < x_{i+1} \leq x_k, \\ x^*_{i+1} & \text{if } \; x_k \leq x_i < x < x_{i+1}. \end{cases} \tag{13}$$

Clearly, ∂g^- and ∂g^+ are monotone. Hence they are the subdifferentials of two convex functions g^+ and g^- such that $g(\bar{x}) = \bar{\lambda}$. These two functions are piecewise linear functions, and are determined by to their values at the points x_i. Since, by definition $g^+(x_k) = g^-(x_k) = \bar{\lambda}$, we obtain for $i = k, k+1, \cdots, p-1$

$$\begin{aligned} g^+(x_{i+1}) &= g^+(x_i) + x^*_{i+1}(x_{i+1} - x_i), \\ g^-(x_{i+1}) &= g^-(x_i) + x^*_{i}(x_{i+1} - x_i), \end{aligned}$$

and for $i = k, k-1, \cdots, 1$

$$\begin{aligned} g^+(x_{i-1}) &= g^+(x_i) + x^*_{i-1}(x_{i-1} - x_i), \\ g^-(x_{i-1}) &= g^-(x_i) + x^*_{i}(x_{i-1} - x_i). \end{aligned}$$

Finally, for $x \neq x_i$,

$$g^+(x) = \begin{cases} +\infty & \text{if } \; x < x_0 \text{ or } x_p < x \\ g^+(x_i) + x^*_{i-1}(x - x_i) & \text{if } \; x_0 \leq x_{i-1} < x < x_i \leq x_k, \\ g^+(x_i) + x^*_{i+1}(x - x_i) & \text{if } \; x_k \leq x_i < x < x_{i+1} \leq x_p, \end{cases}$$

and

$$g^-(x) = \begin{cases} g^-(x_i) + x^*_{i}(x - x_i) & \text{if } \; -\infty \leq x_{i-1} < x < x_i \leq x_k, \\ g^-(x_i) + x^*_{i}(x - x_i) & \text{if } \; x_k \leq x_i < x < x_{i+1} \leq +\infty. \end{cases}$$

These two functions belong to $\mathcal{G}$, hence (CM) holds. Thus (M) is equivalent to (CM) when $X = \mathbb{R}$. It results from (11), (12) and (13) that $g^- \leq g \leq g^+$ for any $g \in \mathcal{G}$.

We close this section with a few words on the gap $[f^-(x), f^+(x)]$ when the initial condition is $f(x_0) = \lambda_0$. Let $t \in [0,1]$ and $x_t = tx_i + (1-t)x_{i+1}$. Then $x_t \in [x_i, x_{i+1}]$ and

$$tf^-(x_i) + (1-t)f^-(x_{i+1}) \leq f(x_t) \leq tf^+(x_i) + (1-t)f^+(x_{i+1}).$$

Thus, we are lead to consider the quantities $e_i = f^+(x_i) - f^-(x_i)$. By induction, for $i = 0, 1, \cdots, p-1$, and the fact that $e_0 = 0$, we have

$$\begin{aligned} e_{i+1} &= [f^+(x_{i+1}) - f^+(x_i)] + [f^+(x_i) - f^-(x_i)] + [f^-(x_i) - f^-(x_{i+1})] \\ &\leq \langle x^*_{i+1}, x_{i+1} - x_i \rangle + e_i + \langle x^*_i, x_i - x_{i+1} \rangle \\ &\leq e_i + \langle x^*_{i+1} - x^*_i, x_{i+1} - x_i \rangle \\ &\vdots \\ &\leq \sum_{j=0,\cdots,i} (x^*_{j+1} - x^*_j)(x_{j+1} - x_j). \end{aligned}$$

It is worth noticing the similarity of these two constructions with the Euler method for ordinary differential equation problems with initial conditions (see, for instance, Henrici [2]).

References

[1] R. T. Rockafellar: Convex Analysis, Princeton University Press, Princeton, New Jersey (1970).

[2] P. Henrici: Elements of Numerical Analysis, Wiley, New York (1964).