PROBABILITY
AND
MATHEMATICAL STATISTICS

Vol. 3, Fasc. 1 (1982), p. 19-36

EXTENSION OF LIPSCHITZ INTEGRANDS AND MINIMIZATION OF NONCONVEX INTEGRAL FUNCTIONALS APPLICATIONS TO THE OPTIMAL RECOURSE PROBLEM IN DISCRETE TIME

BY

J.-B. HIRIART-URRUTY (TOULOUSE)

Abstract. A measurable integrand $f(s, x)$ satisfying a Lipschitz property in x on $\Gamma(s) \subset \boldsymbol{R}^n$ is extended to the whole of $\boldsymbol{R}^n$ preserving the Lipschitz condition in x. This extension is obtained by using the process developed in [6] for an arbitrary function f, Lipschitz on a given subset. The problem of minimizing the integral

$$I_f(x) = \int_S f(s, x(s))\, d\nu(s)$$

over a subset $\mathscr{X}$ of measurable functions x satisfying $x(s) \in \Gamma(s)$ almost everywhere is transformed into the problem of minimizing over $\mathscr{X}$ the integral functional $I_g(x)$ associated with the extended integrand g. Comparison results for optimal values as well as for solutions of the two problems are described. Finally, the results are applied to obtain necessary conditions for optimality for a class of multistage nonconvex stochastic programs.

0. INTRODUCTION

In [6] we studied the way of extending a function f satisfying a Lipschitz property on an arbitrary subset Γ of a metric space E. The problem was in finding a function $f_{\Gamma,k}$ defined and having the Lipschitz property on all of E, which was equal to f on Γ. The definition and the properties of the extension process as well as comparison results with regard to optimization are developed in [6]. In this paper, we begin by carrying out the extension procedure to measurable integrands $f(s, x)$ having a Lipschitz property in x on a subset $\Gamma(s) \subset \boldsymbol{R}^n$ "varying measurably with s". The new integrand $f_{\Gamma,k}(s, x)$ is constructed in such a way that it agrees with $f(s, x)$ on $\Gamma(s)$ and has a Lipschitz property on the whole $\boldsymbol{R}^n$, which is of importance with regard to optimization problems. Section 1 of this paper is devoted to translating, in terms of integrands, the properties developed in a "deterministic context" in [6]. All the definitions and concepts we need for the sequel, concerning measurability of multifunctions, integrands, measurable selections, are recalled in that section. For a full development of such topics in a finite-dimensional context, the reader is referred to the lecture notes by Rockafellar [8].

In many areas of mathematics, including optimization, variational problems, functional analysis, and best approximation, we have to minimize *integral functionals*, i.e. functionals of the following form:

$$x \mapsto I_f(x) = \int_S f(s, x(s))\, d\nu(s).$$

The admissible x are constrained in $\mathscr{X}$, the subspace of measurable functions defined on a measure space $(S, \mathscr{S}, \nu)$, and must satisfy $x(s) \in \Gamma(s)$ almost everywhere (a.e.). In a first step, one may drop the latter constraint by including it in the objective function. Setting $\bar{f}(s, x) = f(s, x)$ if $x \in \Gamma(s)$ and $\bar{f}(s, x) = +\infty$ otherwise, we reduce the above minimization problem to

$$(\bar{\mathscr{P}}) \quad \text{minimize } I_{\bar{f}}(x) = \int_S \bar{f}(s, x(s))\, d\nu(s), \ x \in \mathscr{X}.$$

In the last ten years, *convex* integral functionals have received a great deal of attention (see [1], [8], and the references therein). In that context, the key result was in calculating the conjugate function I_f^*. The properties of I_f as well as the determination of its subdifferential were then derived from the knowledge of I_f^*. In the absence of both differentiability and convexity assumptions on $f(s, \cdot)$, a step has been taken in the calculation of the *generalized gradient* ∂I_f of I_f, at least when $f(s, \cdot)$ and I_f have appropriate *Lipschitz* properties. The method consisted in reducing the problem of determining ∂I_f to that of evaluating the subdifferential of a certain "tangent integral functional" which turned out to be convex. In that respect, we mention [13], Chapter III, for the "tangentially convex" case and L_X^1 as underlying space, and [14] and [3] for the general Lipschitz case in L_X^p $(1 \leqslant p < \infty)$ and L_X^∞, respectively. The general situation will be treated in a forthcoming article by the author and R. J.-B. Wets. However, our method in deriving necessary conditions for $x_0 \in \mathscr{X}$ to be the solution of $(\bar{\mathscr{P}})$ (Section 2) *does not rely upon the direct calculation of* $\partial I_{\bar{f}}$. Our way of grappling with $(\bar{\mathscr{P}})$ consists in transforming it into an extended problem

$$(\mathscr{P})_{\Gamma,k} \quad \text{minimize } I_{f_{\Gamma,k}}(x) = \int_S f_{\Gamma,k}(s, x(s))\, d\nu(s), \ x \in \mathscr{X},$$

where $f_{\Gamma,k}$ is the extended (Lipschitz) integrand. Comparison results for optimal values as well as for solutions of $(\bar{\mathscr{P}})$ and $(\mathscr{P})_{\Gamma,k}$ will be provided. Afterwards, necessary conditions for optimality are deduced from the known results on $\partial I_{f_{\Gamma,k}}$ in the Lipschitz case.

In Section 3, the obtained results are applied to the study of a class of multistage nonconvex stochastic programs. We are more particularly concerned with the *optimal recourse problem* such as described in [11] and [12], from which we take most of the prerequisites. Roughly speaking, in a multistage stochastic program, we are in the presence of finitely many stages (N), at each of which a decision (or a recourse) is selected on the basis of prior observations of random events, and subject to costs and constraints depending on these observations as well as past decisions. The goal is to minimize an expected cost (hence an integral functional) taking into account the known probability distribution of random events. In the case of *convex* costs and constraints, necessary and sufficient conditions for a decision rule to be optimal are derived in [11] and [12] with the powerful tool which is the duality approach. For related work, especially from the Soviet literature in the framework of models for optimal economic developments, see the references in [11]. In the nonconvex but locally Lipschitz case, two-stage ($N = 2$) stochastic programs were studied in [4] in their *dynamic* formulation. Here, we consider a particular class of N-stage stochastic problems, namely those whose cost function has a Lipschitz property, but only in the admissible decision rules. Thus, the stochastic optimization problem is set in the framework of Section 2, with $\mathscr{X}$ representing the *nonanticipativity* constraint. As in the convex case, we show that it is possible (under certain constraint qualifications) to associate with nonanticipativity a price system (i.e. a system of Lagrange multipliers) having a martingale property.

1. EXTENSION OF LIPSCHITZ INTEGRANDS

1.1. In this Section 1, $(S, \mathscr{S})$ is a general measurable space. A *multifunction* Γ (or set-valued mapping) from S into $\boldsymbol{R}^n$ will be denoted by $\Gamma\colon S \rightrightarrows \boldsymbol{R}^n$. We also fix some other notation:

(i) the *graph* of Γ is $\operatorname{gr}\Gamma = \{(s, x) \in S \times \boldsymbol{R}^n \mid x \in \Gamma(s)\}$;

(ii) the *inverse* of Γ is the multifunction $\Gamma^{-1}\colon \boldsymbol{R}^n \rightrightarrows S$ defined by $\Gamma^{-1}(x) = \{s \in S \mid x \in \Gamma(s)\}$; so, for any $X \subset \boldsymbol{R}^n$,

$$\Gamma^{-1}(X) = \{s \in S \mid \Gamma(s) \cap X \neq \emptyset\}.$$

For the most part, the values of the multifunctions which will be considered in the sequel are either epigraphs of lower-semicontinuous (l.s.c.) functions or subsets defining the constraints of a certain optimization problem. So, without loss of generality, we shall be concerned with *closed-valued* multifunctions. Such a multifunction is said to be *measurable* (relatively to $\mathscr{S}$) if, for each closed set $X \subset \boldsymbol{R}^n$, the set $\Gamma^{-1}(X)$ belongs to $\mathscr{S}$. As a consequence, $\operatorname{gr}\Gamma$ is an $(\mathscr{S}\otimes\mathscr{B}_n)$-measurable subset of $S \times \boldsymbol{R}^n$ (where $\mathscr{B}_n$ is the σ-field of Borel sets in $\boldsymbol{R}^n$). The general properties of measurable multifunctions, the basic measurable selection theorem, and the description of operations on multifunctions that preserve measurability are fully developed in Part I of [8].

Now, let us consider functions from $S \times \boldsymbol{R}^n$ into $\bar{\boldsymbol{R}}$ (the extended reals). Such functions will be called *integrands* on $S \times \boldsymbol{R}^n$. An integrand f is completely determined by the epigraph multifunction $s \rightrightarrows \operatorname{epi} f(s, \cdot)$. We shall say that the integrand $f\colon S \times \boldsymbol{R}^n \to \bar{\boldsymbol{R}}$ is

(i) *measurable* if f is $(\mathscr{S} \otimes \mathscr{B}_n)$-measurable on $S \times \boldsymbol{R}^n$,

(ii) *normal* if the epigraph multifunction is a closed-valued measurable multifunction,

(iii) *proper* if $f(s, \cdot)$ is proper for every $s \in \mathscr{S}$, i.e. if $f(s, \cdot)$ does not take the value $-\infty$ and is not identically equal to $+\infty$ for every s.

A normal integrand is a measurable integrand. Conversely, if $(S, \mathscr{S})$ is complete with respect to some nonnegative σ-finite measure on $\mathscr{S}$, then a measurable integrand f yields a normal integrand by the "closure" operation: $\operatorname{cl} f$ defined by

$$\operatorname{cl} f(s, x_0) = \liminf_{x \to x_0} f(s, x) \quad \text{for all } (s, x_0) \in S \times \boldsymbol{R}^n$$

is a normal integrand.

We also fix another definition which expresses a kind of regular behavior of $f(s, \cdot)$ on a nonempty set $\Gamma(s)$ "varying measurably with s".

Definition 1.1. Let f be a measurable integrand on $S \times \boldsymbol{R}^n$ and let $\Gamma\colon S \rightrightarrows \boldsymbol{R}^n$ be a nonempty closed-valued measurable multifunction. Then f is said to be *Lipschitz* on Γ if for each $s \in S$ there exists $k(s) \in \boldsymbol{R}_+$ satisfying

$$(1.1) \qquad |f(s, x) - f(s, y)| \leqslant k(s) \|x - y\| \quad \text{for all } x \text{ and } y \text{ in } \Gamma(s).$$

Hence, "f Lipschitz on Γ" presupposes three ingredients: measurability of f as an integrand, measurability of the nonempty closed-valued multifunction Γ, and the Lipschitz property above.

The least positive real $k(s)$ such that (1.1) holds for $f(s, \cdot)$ is

$$|||f(s, \cdot)||| = \sup \left\{ \frac{|f(s, x) - f(s, y)|}{\|x - y\|} \;\middle|\; x, y \in \Gamma(s),\ x \neq y \right\}.$$

Clearly, k may be chosen to be measurable in relation (1.1).

The simplest example of the Lipschitz integrand is the indicator integrand δ_Γ of a measurable multifunction Γ, i.e.

$$\delta_\Gamma(s, x) = \begin{cases} 0 & \text{if } x \in \Gamma(s), \\ +\infty & \text{if } x \notin \Gamma(s). \end{cases}$$

For another example, let us consider a measurable proper convex integrand f (i.e. $\operatorname{epi} f(s, \cdot)$ is convex for each s); f is Lipschitz on Γ whenever Γ is a measurable compact-valued multifunction satisfying the condition

$$\Gamma(s) \subset \operatorname{int} \{x \in \boldsymbol{R}^n \mid f(s, x) \in \boldsymbol{R}\} \quad \text{for all } s.$$

By pasting together the results on composition of Lipschitz functions and the measurability techniques developed in [8], Section II, one gets new Lipschitz integrands via usual operations like addition, left-scalar multiplication, etc.

1.2. Let f be an integrand Lipschitz on Γ. In the way of extending f to the whole $\boldsymbol{R}^n$, we shall use the integrand $\bar{f}$ defined by

$$\bar{f}(s, x) = f(s, x) + \delta_{\Gamma(s)}(x).$$

$\bar{f}$ is clearly measurable; it is moreover normal whenever $(S, \mathscr{S})$ is complete ([8], Theorem 2A). Again $\bar{f}$ is normal in the case where f is proper normal and Γ nonempty-valued ([8], Proposition 2M). Therefore, *when considering an integrand f Lipschitz on Γ, it will be assumed throughout that $\bar{f}$ is proper and normal.*

THEOREM 1.1. *Let f be an integrand Lipschitz on Γ and let k be a measurable function such that $k(s) \geqslant |||f(s, \cdot)|||$ for all s. Then the integrand $f_{\Gamma,k}$ defined by*

$$f_{\Gamma,k}(s, x) = \inf_{u \in \Gamma(s)} \{f(s, u) + k(s)\|x-u\|\}$$

is Lipschitz on $\boldsymbol{R}^n$ with Lipschitz constant $k(s)$ and verifies

$$f_{\Gamma,k}(s, x) = f(s, x) \qquad \textit{for all } x \in \Gamma(s).$$

1.3. Tangent cones, generalized gradients. Let E be a real Banach space, let A be a nonempty closed subset of E, and let $u_0 \in A$.

Definition 1.2. δ is a *tangent direction* to A at u_0 if for every sequence $\{u_n\} \subset A$ converging to u_0 and for every $\{\lambda_n\} \subset \boldsymbol{R}_+^*$ converging to 0 there exists a sequence $\{\delta_n\}$ converging to δ such that $u_n + \lambda_n \delta_n \in A$ for all n.

The cone of all tangent directions to A at u_0 is the *tangent cone* to A at u_0 and will be denoted by $\mathscr{C}_A(u_0)$. Its polar cone, i.e. the set of $n \in E^*$ (topological dual space of E) satisfying $\langle n, \delta\rangle \leqslant 0$ for all $\delta \in \mathscr{C}_A(u_0)$, is called the *normal cone* to A at u_0 and will be denoted by $\mathscr{N}_A(u_0)$. The cones $\mathscr{C}_A(u_0)$ and $\mathscr{N}_A(u_0)$ are nonempty, closed and convex.

PROPOSITION 1.1 ([4]). *Let $\Gamma\colon S \rightrightarrows \boldsymbol{R}^n$ be a closed-valued measurable multifunction and let $u\colon S \to \boldsymbol{R}^n$ be a measurable function such that $u(s) \in \Gamma(s)$ for all s. Then the multifunctions $s \mapsto \mathscr{C}_{\Gamma(s)}(u(s))$ and $s \mapsto \mathscr{N}_{\Gamma(s)}(u(s))$ are measurable.*

Let $f\colon E \to \bar{\boldsymbol{R}}$ be finite at x_0. Starting from the geometric concept of tangent cone, the *generalized directional derivative* of f at x_0 is defined by

$$d \mapsto f^{\Box}(x_0; d) = \inf\{\mu \in \boldsymbol{R} \mid (d, \mu) \in \mathscr{C}_{\operatorname{epi} f}(x_0, f(x_0))\} \tag{1.2}$$

with the usual convention that $\inf \emptyset = +\infty$. The definition of $f^{\Box}(x_0; \cdot)$ in (1.2) is "geometric" without any "analytic" formula involving limits of difference quotients of some kind. Rockafellar [9] gave recently the analytic form of $f^{\Box}(x_0; \cdot)$ by translating the construction of $\mathscr{C}_{\operatorname{epi} f}(x_0, f(x_0))$ in terms of sequences (Definition 1.2) into a statement in terms of "lim sup" of certain quotients. When f is Lipschitz around x_0, $f^{\Box}(x_0; d)$ has a simpler expression $f^{\circ}(x_0; d)$ which was first produced by Clarke [2],

$$f^{\circ}(x_0; d) = \limsup_{\substack{x \to x_0 \\ \lambda \to 0^+}} [f(x+\lambda d) - f(x)] \lambda^{-1}.$$

In the general case, the geometric definition (1.2) will be well adapted for measurability purposes. The generalized gradient of f at x_0, in Clarke's sense, is defined as follows:

$$\partial f(x_0) = \{x^* \in E^* \mid \langle x^*, d \rangle \leqslant f^{\Box}(x_0; d) \text{ for all } d \in E\}.$$

When f is convex, $\partial f(x_0)$ is the subdifferential in the sense of convex analysis; when f is C^1 at x_0, $\partial f(x_0)$ is reduced to one element, namely $\nabla f(x_0)$.

THEOREM 1.2. *Let f: $S \times \mathbf{R}^n \to \bar{\mathbf{R}}$ be a normal integrand and let x_0: $S \to \mathbf{R}^n$ be a measurable function such that $|f(s, x_0(s))| < \infty$ for all s. Then*

(a) *$(s, d) \mapsto f^{\Box}(s, x_0(s); d)$ is a normal convex integrand on $S \times \mathbf{R}^n$,*

(b) *the multifunction $s \mapsto\!\!\!\rightrightarrows \partial f(s, x_0(s))$ is measurable.*

Proof. (a) The function $s \mapsto (x_0(s), f(s, x_0(s)))$ is measurable. Therefore, according to Proposition 1.1, the multifunction Δ: $s \rightrightarrows \mathbf{R}^{n+1}$ which assigns to s the tangent cone to epi $f(s, \cdot)$ at $[x_0(s), f(s, x_0(s))]$ is measurable. Now, $\Delta(s)$ is nothing but

$$\{(d, \mu) \in \mathbf{R}^{n+1} \mid f^{\Box}(s, x_0(s); d) \leqslant \mu\}$$

(see [5]). Hence the normality of the integrand $(s, d) \mapsto f^{\Box}(s, x_0(s); d)$ is proved.

(b) The multifunction Δ°: $s \rightrightarrows \mathbf{R}^{n+1}$ which assigns to s the normal cone to epi $f(s, \cdot)$ at $[x_0(s), f(s, x_0(s))]$ is closed-valued and measurable; the constant multifunction $s \mapsto\!\!\!\rightrightarrows \mathbf{R}^n \times \{-1\}$ is measurable. Hence the multifunction

$$s \mapsto\!\!\!\rightrightarrows \Delta^{\circ}(s) \cap (\mathbf{R}^n \times \{-1\})$$

is measurable (closed-valued); see [8], Theorem 1M.

$\partial f(s, x_0(s))$ is the image of $\Delta^{\circ}(s) \cap (\mathbf{R}^n \times \{-1\})$ under the projection mapping $(x, \mu) \mapsto x$. Thus, following Corollary 1P in [8], the multifunction $s \mapsto\!\!\!\rightrightarrows \partial f(s, x_0(s))$ is measurable.

Concerning the generalized gradients of $\bar{f}(s, \cdot)$ and $f_{\Gamma,k}(s, \cdot)$ (such as defined in Section 1.2), we have general comparison results.

THEOREM 1.3. *Let f be an integrand Lipschitz on Γ, let x_0 be a measurable function such that $x_0(s) \in \Gamma(s)$ for all s, and let k be a measurable mapping.*

(a) *If $k(s) \geqslant |||f(s, \cdot)|||$ for all s, then*

$$\partial \bar{f}(s, x_0(s)) \subset \partial f_{\Gamma,k}(s, x_0(s)) + \mathcal{N}_{\Gamma(s)}(x_0(s)) \quad \text{for all } s. \tag{1.3}$$

(b) *If $k(s) > |||f(s, \cdot)|||$ for all s, then*

(1.4) $$\partial f_{\Gamma,k}(s, x_0(s)) \subset \partial \bar{f}(s, x_0(s)) \cap k(s)B,$$

where B denotes the closed unit ball in $\mathbf{R}^n$.

For the proof see [6], Theorem 2.

When $f(s, \cdot)$ is convex on $\Gamma(s)$ (i.e. $\bar{f}(s, \cdot)$ is a convex integrand), inclusions (1.3) and (1.4) become equalities. Actually, results (1.3) and (1.4) are strengthened to equalities for which we shall call "tangentially convex integrands".

Let A be a nonempty closed subset of a real Banach space E and let $u_0 \in A$. A classic way to get a conical approximation of A at u_0 is to consider the *contingent cone* to A at u_0 (also called *cone of adherent displacements* for A from u_0). This cone, denoted by $T_A(u_0)$, is defined by

$$T_A(u_0) = \{\delta \in E \mid \exists \lambda_n \downarrow 0,\ \delta_n \to \delta \text{ with } u_0 + \lambda_n \delta_n \in A \text{ for all } n\}.$$

Proposition 1.2. *Let Γ: $S \rightrightarrows \mathbf{R}^n$ be a closed-valued measurable multifunction and let u: $S \to \mathbf{R}^n$ be a measurable function such that $u(s) \in \Gamma(s)$ for all s. Then the multifunction $s \mapsto T_{\Gamma(s)}(u(s))$ is measurable.*

$\mathscr{C}_{\Gamma(s)}(u(s))$ is always included in $T_{\Gamma(s)}(u(s))$ (see [5]); if, in addition to the assumptions of Proposition 1.2, we suppose that $\mathscr{C}_{\Gamma(s)}(u(s)) = T_{\Gamma(s)}(u(s))$ for all s, we shall say that Γ is *tangentially convex at u*. Concerning the corresponding notion on integrands, a normal integrand f: $S \times \mathbf{R}^n \to \bar{\mathbf{R}}$ will be called *tangentially convex at x* if $f(s, x(s))$ is finite for all s and if $s \mapsto \operatorname{epi} f(s, \cdot)$ is tangentially convex at u: $s \mapsto (x(s), f(s, x(s)))$. In an analytic way, the tangential convexity of f at x is translated as

$$f^{\Box}(s, x(s); d) = \liminf_{\substack{\delta \to d \\ \lambda \to 0+}} [f(s, x(s) + \lambda\delta) - f(s, x(s))]\lambda^{-1} \quad \text{for all } d.$$

For a normal integrand f, the tangential convexity at x is in particular ensured whenever $f(s, \cdot)$ is convex in a neighborhood of $x(s)$ or $f(s, \cdot)$ is C^1 at $x(s)$.

Theorem 1.4. *Let f be an integrand Lipschitz on Γ, let x_0 be a measurable function such that $x_0(s) \in \Gamma(s)$ for all s, and let k be a measurable mapping such that $k(s) > |||f(s, \cdot)|||$ for all s. We suppose that the integrand $\bar{f}$ is tangentially convex at x_0. Then $f_{\Gamma,k}$ is tangentially convex at x_0 and*

$$\partial \bar{f}(s, x_0(s)) = \partial f_{\Gamma,k}(s, x_0(s)) + \mathscr{N}_{\Gamma(s)}(x_0(s)),$$

$$\partial f_{\Gamma,k}(s, x_0(s)) = \partial \bar{f}(s, x_0(s)) \cap k(s)B \quad \textit{for all } s.$$

The theorem follows from Theorem 1.3 and [6], Theorem 3.

2. MINIMIZING INTEGRAL FUNCTIONALS

From now on, ν will denote a positive σ-finite measure on $(S, \mathscr{S})$. For any normal integrand f: $S \times \mathbf{R}^n \to \bar{\mathbf{R}}$ and any measurable function x: $S \to \mathbf{R}^n$, it follows that $s \mapsto f(s, x(s))$ is measurable, and therefore the integral functional

$$x \mapsto I_f(x) = \int_S f(s, x(s)) d\nu(s)$$

has a well-defined value in $\bar{\mathbf{R}}$ under the convention that $I_f(x) = +\infty$ whenever

$$\int_S [f(s, x(s))]^+ d\nu(s) = +\infty.$$

We denote by $L^0_{\mathbf{R}^n}$ the space of all measurable functions x: $S \to \mathbf{R}^n$. In optimization problems dealing with integral functionals I_f, there are typically two kinds of constraints on the admissible functions x. The first type of constraints concerns the "nature" of x: x must belong to a space $\mathscr{X} \subset L^0_{\mathbf{R}^n}$. For the second type of constraints, we are given a closed-valued multifunction Γ: $S \rightrightarrows \mathbf{R}^n$ which we will suppose, without loss of generality, is nonempty-valued on S. We will be concerned with measurable x such that $x(s) \in \Gamma(s)$ a.e. The assumed measurability of Γ ensures that such *measurable selections* do exist ([8], Corollary 1C). $\mathscr{T}$ will denote the set of all measurable selections of Γ.

Typically, the optimization problem we are concerned with is

($\mathscr{P}$) minimize $I_f(x)$ over $\mathscr{X} \cap \mathscr{T}$.

By the usual device which consists in "transferring the constraint into the objective", ($\mathscr{P}$) is equivalent to

($\bar{\mathscr{P}}$) minimize $I_{\bar{f}}(x)$ over $\mathscr{X}$.

In writing ($\bar{\mathscr{P}}$), we implicitly assume that the integrand $\bar{f}$ is normal, which is secured whenever f is proper. As a consequence of assumptions on f and Γ, the function

$$m_\Gamma\colon s \mapsto m_\Gamma(s) = \inf_{x \in \Gamma(s)} f(s, x)$$

is measurable ([8], Theorem 2K).

The integral functional I_f is said to be *proper on* $\mathscr{T}$ if $I_f(x) > -\infty$ for all $x \in \mathscr{T}$ and if $I_f(\bar{x}) < +\infty$ for at least one $\bar{x} \in \mathscr{T}$. The first general result relating minimization of I_f on $\mathscr{T}$ to pointwise minimization of $f(s, \cdot)$ on $\Gamma(s)$ is the following one:

PROPOSITION 2.1. *Suppose* $\mathscr{T} \subset \mathscr{X}$. *Then*

$$\inf_{x \in \mathscr{T}} \int_S f(s, x(s)) d\nu(s) = \int_S [\inf_{x \in \Gamma(s)} f(s, x)] d\nu(s). \tag{2.1}$$

If, moreover, the integral functional I_f is proper on $\mathscr{T}$, the following statements are equivalent:

(a) x_0 *minimizes* I_f *on* $\mathscr{T}$;

(b) $x_0 \in \mathscr{T}$ *and* $x_0(s)$ *minimizes* $f(s, \cdot)$ *on* $\Gamma(s)$ *a.e.*

Proof. The proof of (2.1) is contained in the parenthetical case of Theorem 3A in [8]. If I_f is proper on $\mathscr{T}$, then $\int_S f(s, x_0(s))\,dv(s)$ is finite for x_0 minimizing I_f on $\mathscr{T}$. Hence the desired equivalence is easily deduced from (2.1).

The assumption $\mathscr{T} \subset \mathscr{X}$ is satisfied when $\mathscr{X} = L^{\circ}_{\mathbb{R}^n}$ and, more generally, when $\mathscr{X} = L^p_{\mathbb{R}^n}$ (i.e. $L^p_{\mathbb{R}^n}(S, \mathscr{S}, v)$) and the function $s \mapsto \sup\{\|x\|;\ x \in \Gamma(s)\}$ is in $L^p_{\mathbb{R}^n}$. For results similar to those of Proposition 2.1 in the $L^1_{\mathbb{R}^n}$-case, see [16], Theorem 2 and Corollary 2. Actually, except in the case where $\mathscr{X}$ is "large enough", the constraint imposed on the nature of x ($x \in \mathscr{X}$) is not satisfied for all selections of Γ. We shall study the problem ($\mathscr{P}$) in the following particular framework:

($\mathscr{A}_1$) the underlying space is $L^p_{\mathbb{R}^n}$ $(1 \leqslant p \leqslant \infty)$;

($\mathscr{A}_2$) f is an integrand Lipschitz on Γ and the function $s \mapsto |||f(s, \cdot)|||$ is in $L^q_{\mathbb{R}}$ $(1/p+1/q = 1)$;

($\mathscr{A}_3$) $\mathscr{X}$ is a subspace of measurable functions in $L^p_{\mathbb{R}^n}$ and $\mathscr{T} \subset L^p_{\mathbb{R}^n}$.

As usually in the theory of integral functionals, we shall distinguish $\mathscr{X}$ by the presence or absence of a certain property of decomposability. Following the definition in [8], $\mathscr{X}$ is said to be *decomposable* if S can be expressed as the union of an increasing sequence of measurable subsets S_m $(m = 1, 2, \ldots)$ such that for every S_m and bounded measurable function x_α: $S \to \mathbb{R}^n$, and every $x_\beta \in \mathscr{X}$, the "decomposed" (measurable) function

$$s \mapsto x_m(s) = \begin{cases} x_\alpha(s) & \text{for } s \in S_m, \\ x_\beta(s) & \text{for } s \in S_m^c \text{ (complementary set of } S_m \text{ in } S) \end{cases}$$

belongs to $\mathscr{X}$.

As for ($\bar{\mathscr{P}}$), we will transfer the constraint $\mathscr{T}$ in the objective by redefining a penalized version of the objective function. For that purpose, we consider the extension $f_{\Gamma,k}$ of f such as defined in Section 1. Since v is σ-finite, there exists a function ε: $S \to \mathbb{R}^*_+$ in $L^q_{\mathbb{R}^n}$. We consider throughout the extension $f_{\Gamma,k}$ built up with k: $s \mapsto k(s) = |||f(s, \cdot)|||+\varepsilon(s)$. The extended version of ($\bar{\mathscr{P}}$) becomes now

($\mathscr{P})_{\Gamma,k}$ minimize $I_{f_{\Gamma,k}}$ over $\mathscr{X}$.

We note that, following assumption ($\mathscr{A}_2$), $I_{f_{\Gamma,k}}$ is Lipschitz on $\mathscr{T}$ with Lipschitz constant $\|k(\cdot)\|_{L^q_{\mathbb{R}}}$ whenever $I_{f_{\Gamma,k}}$ is finite at some $\bar{x} \in \mathscr{T}$.

THEOREM 2.1. *Suppose that* $I_{f_{\Gamma,k}}$ *is finite at some* $\bar{x}\in\mathscr{T}$ *and that* $\mathscr{X}$ *is decomposable. Then*

(a) $\inf_{x\in\mathscr{X}\cap\mathscr{T}} I_f(x) = \inf_{x\in\mathscr{X}} I_{f_{\Gamma,k}}(x) = \int_S [\inf_{x\in\Gamma(s)} f(s,x)]\,d\nu(s)$;

(b) *the following statements are equivalent*:

(i) $x_0\in\mathscr{X}\cap\mathscr{T}$ *minimizes* I_f *on* $\mathscr{X}\cap\mathscr{T}$,

(ii) $x_0\in\mathscr{X}$ *minimizes* $I_{f_{\Gamma,k}}$ *on* $\mathscr{X}$.

Proof. (a) Applying Theorem 3A from [8] to the integrands $\bar{f}$ and $f_{\Gamma,k}$, we get

$$\inf_{x\in\mathscr{X}}\int_S \bar{f}(s,x(s))\,d\nu(s) = \int_S m_\Gamma(s)\,d\nu(s)$$

and

$$\inf_{x\in\mathscr{X}}\int_S f_{\Gamma,k}(s,x(s))\,d\nu(s) = \int_S [\inf_{x\in\mathbf{R}^n} f_{\Gamma,k}(s,x)]\,d\nu(s).$$

Now, an easy consequence of the definition of $f_{\Gamma,k}$ is

$$m_\Gamma(s) = \inf_{x\in\mathbf{R}^n} f_{\Gamma,k}(s,x).$$

Hence the equality of minimum values is proved.

(b) Let $x_0\in\mathscr{X}$ minimizing $I_{\bar{f}}$ on $\mathscr{X}$. The integral $\int_S f(s,x_0(s))\,d\nu(s)$ is finite and, according to the proof of part (a), we have

$$\int_S f(s,x_0(s))\,d\nu(s) = \int_S [\inf_{x\in\mathbf{R}^n} f_{\Gamma,k}(s,x)]\,d\nu(s),$$

whence

$$f(s,x_0(s)) = \inf_{x\in\mathbf{R}^n} f_{\Gamma,k}(s,x) \text{ a.e.}$$

Conversely, let $x_0\in\mathscr{X}$ minimizing $I_{f_{\Gamma,k}}$ over $\mathscr{X}$. The only thing to prove is that x_0 necessarily belongs to $\mathscr{T}$. Let $\alpha(s)$ be the distance from $x_0(s)$ to $\Gamma(s)$; α is measurable and, furthermore, $\alpha\in L^p_{\mathbf{R}^n}$. Suppose that $\alpha(s)>0$ on a set A of positive measure. The multifunction Δ defined by

$$\Delta(s) = \left\{x\in\Gamma(s) \mid f(s,x)+k(s)\|x-x_0(s)\| \leqslant f_{\Gamma,k}(s,x_0(s))+\frac{\alpha(s)k(s)}{2}\right\}$$

is measurable ([8], Theorem 2I). Hence there is a measurable selection $\hat{x}$ of Δ, that is to say: $\hat{x}\in\mathscr{T}$ and

$$f(s,\hat{x}(s))+\frac{\alpha(s)k(s)}{2} \leqslant f_{\Gamma,k}(s,x_0(s)) \quad \text{for almost all } s. \tag{2.2}$$

Observe that $\hat{x}(s) = x_0(s)$ for almost all $s\in A^c$. Let S_m $(m = 1, 2, \ldots)$ be as in the definition of decomposability. Intersecting each S_m with the measurable set $\{s\in S \mid |\hat{x}(s)| \leqslant m\}$ if necessary, we can suppose $\hat{x}$ to be bounded on S_m. We define a new measurable function $\hat{x}_m$ as follows:

$$\hat{x}_m(s) = \begin{cases}\hat{x}(s) & \text{for } s\in S_m,\\ x_0(s) & \text{for } s\in S_m^c.\end{cases}$$

According to the decomposability assumption, $\hat{x}_m \in \mathscr{X}$. Now, following (2.2), we have

$$(2.3) \qquad \int_S f_{r,k}(s, \hat{x}_m(s))\,dv(s) + \int_{S_m} \frac{\alpha(s)k(s)}{2}\,dv(s) \leqslant \int_S f_{r,k}(s, x_0(s))\,dv(s).$$

For m sufficiently large, $v(S_m \cap A) > 0$; hence (2.3) yields

$$\int_S f_{r,k}(s, \hat{x}_m(s))\,dv(s) < \int_S f_{r,k}(s, x_0(s))\,dv(s),$$

which contradicts the optimality of x_0. Therefore, $\alpha(s) = 0$ a.e. and $x_0 \in \mathscr{T}$.

Remark. As noticed in the proof above, *if* x_0 *minimizes* $I_f(x)$ *on* $\mathscr{X} \cap \mathscr{T}$ *(*$\mathscr{X}$ *decomposable or not), then* x_0 *minimizes* $I_{f_{r,k}}(x)$ *on* $\mathscr{X}$. This is a weaker result than (b) in Theorem 2.1 but it will play a key role in the sequel.

Necessary conditions for $x_0 \in \mathscr{X} \cap \mathscr{T}$ to be a solution of $(\mathscr{P})$ will be derived via necessary conditions for optimality in $(\mathscr{P})_{r,k}$. For that, let us fix first the duality framework. The underlying space $L = L^p_{\mathbb{R}^n}$, endowed with the norm topology, is paired with its topological dual space L^* which is supposed to be equipped with a topology compatible with respect to the pairing. If $1 \leqslant p < \infty$, the canonical pairing between $L^p_{\mathbb{R}^n}$ and $L^q_{\mathbb{R}^n}$ $(1/p+1/q = 1)$ is simply

$$\langle x, x^* \rangle = \int_S x(s)\,x^*(s)\,dv(s)$$

and one can consider $L^q_{\mathbb{R}^n}$ equipped with the norm topology whenever $1 < p < \infty$. In the situations we will encounter, functionals on $L^\infty_{\mathbb{R}^n}$ will have nice properties with respect to the norm topology. The space $L^\infty_{\mathbb{R}^n}$ equipped with the norm topology, has a dual space which is identified as follows: $z^* \in (L^\infty_{\mathbb{R}^n})^*$ is said to be *singular* if there is an increasing sequence S_k $(k = 1, 2, \ldots)$ of measurable sets covering S and such that, whenever $x \in L^\infty_{\mathbb{R}^n}$ is a function vanishing a.e. outside of some S_k, we have $z^*(x) = 0$. The set of these singular functionals will be denoted by $L^{\mathrm{sing}}_{\mathbb{R}^n}$. For each $x^* \in (L^\infty_{\mathbb{R}^n})^*$, we exhibit the "absolutely continuous part" y^* and the "singular part" z^*. Under the pairing

$$\langle x, (y^*, z^*) \rangle = \langle x, y^* \rangle + z^*(x) \quad \text{for } x \in L^\infty_{\mathbb{R}^n},\ (y^*, z^*) \in L^1_{\mathbb{R}^n} \times L^{\mathrm{sing}}_{\mathbb{R}^n},$$

the relation

$$(L^\infty_{\mathbb{R}^n})^* = L^1_{\mathbb{R}^n} \oplus L^{\mathrm{sing}}_{\mathbb{R}^n}$$

holds isometrically (see [7] and [8]). Actually, the elements of $(L^\infty_{\mathbb{R}^n})^*$ we will produce as elements of the generalized gradient of a certain integral functional on $L^\infty_{\mathbb{R}^n}$ will be in $L^1_{\mathbb{R}^n}$.

The following statement is the synthesis of results in [14] and [3] concerning the generalized gradient of a Lipschitz integral functional on $L^p_{\mathbb{R}^n}$ $(1 \leqslant p \leqslant \infty)$.

PROPOSITION 2.2. *Let g be a measurable integrand on $S \times \mathbb{R}^n$ and let $x_0 \in L^p_{\mathbb{R}^n}$. We suppose that*

(a) *$f(s, \cdot)$ is Lipschitz in a neighborhood of $x_0(s)$ a.e.,*
(b) *the integral functional I_g is finitely defined at x_0,*
(c) *there exists a $k \in L^q_{\mathbb{R}^n}$ satisfying*

$$|g(s, x(s)) - g(s, y(s))| \leqslant k(s)\|x(s) - y(s)\| \text{ a.e.}$$

for all x, y in a neighborhood of x_0 in $L^p_{\mathbb{R}^n}$.

Then I_g is Lipschitz in a neighborhood of x_0 and

$$\partial I_g(x_0) \subset \{x^* \in L^q_{\mathbb{R}^n} \mid x^*(s) \in \partial g(s, x_0(s)) \text{ a.e.}\}. \tag{2.4}$$

Moreover, if g is tangentially convex at x_0, then I_g is tangentially convex at x_0 and equality holds in (2.4).

The intermediate step in deriving (2.4) is in proving that

$$I_g^\circ(x_0; d) \leqslant \int_S g^\circ(s, x_0(s); d(s))\, dv(s) \quad \text{for all } d \in L^p_{\mathbb{R}^n}.$$

Then the general statements giving the subdifferential of integral *convex* functionals yield the desired inclusion. We observe that $I_g^\circ\colon d \mapsto I_g^\circ(x_0; d)$ is *continuous* at $d = 0$ in the norm topology of $L^p_{\mathbb{R}^n}$; hence in the $L^\infty_{\mathbb{R}^n}$-case, the subdifferential of I_g° at 0 (which is nothing but the generalized gradient of I_g at x_0) consists of elements of $L^1_{\mathbb{R}^n}$ (see [7]).

Now, we turn back to the problem of deriving necessary conditions for optimality of $(\mathscr{P})$ under the assumptions described earlier.

THEOREM 2.2. *Suppose that I_f is finite at some $\bar{x} \in \mathscr{T}$. If $x_0 \in L^p_{\mathbb{R}^n}$ minimizes I_f on $\mathscr{X} \cap \mathscr{T}$, then there exists an $x^* \in L^q_{\mathbb{R}^n}$ satisfying*

(a) $-x^* \in \mathscr{N}_{\mathscr{X}}(x_0)$,
(b) $x^*(s) \in \partial\bar{f}(s, x_0(s))$ *and* $\|x^*(s)\| \leqslant k(s)$ *a.e.*

Proof. According to the Remark following Theorem 2.1, x_0 minimizing I_f on $\mathscr{X} \cap \mathscr{T}$ minimizes $I_{f_{\Gamma,k}}$ on $\mathscr{X}$. Since $I_{f_{\Gamma,k}}$ is Lipschitz on $L^p_{\mathbb{R}^n}$, we then necessarily have $0 \in \partial I_{f_{\Gamma,k}}(x_0) + \mathscr{N}_{\mathscr{X}}(x_0)$ (see [2] and [5]). Thus, there exists an $x^* \in L^q_{\mathbb{R}^n}$ such that

$$-x^* \in \mathscr{N}_{\mathscr{X}}(x_0) \quad \text{and} \quad x^*(s) \in \partial f_{\Gamma,k}(s, x_0(s)) \text{ a.e.}$$

The result (b) is then deduced from inclusion (1.4) in Theorem 1.3.

Remarks. One might try to go further in relation (b) and write, under suitable assumptions, that

$$\partial\bar{f}(s, x_0(s)) \subset \partial f(s, x_0(s)) + \mathscr{N}_{\Gamma(s)}(x_0(s)).$$

Conditions on the initial integrand f and on the multifunction Γ for the above inclusion to hold are fully developed in [10]. Moreover, when $\Gamma(s)$ has a representation in terms of equalities and inequalities, i.e.

$$\Gamma(s) = \{x \in \mathbf{R}^n \mid f_i(s, x) \leqslant 0,\, i = 1, \ldots, \alpha, \text{ and } h_j(s, x) = 0,\, j = 1, \ldots, \beta\},$$

regularity conditions do exist allowing us to compare results between $\mathcal{N}_{\Gamma(s)}(x_0(s))$ and the cones $\mathbf{R}\nabla h_j(s, x_0(s))$, $\mathbf{R}_+ \partial f_i(s, x_0(s))$ (see [5] and [10]).

3. NECESSARY CONDITIONS FOR OPTIMALITY IN A MULTISTAGE NONCONVEX STOCHASTIC PROGRAM

3.1. The problem and the data. For $k = 1, \ldots, N$, let $\xi_k \in \mathbf{R}^{\nu_k}$ and $u_k \in \mathbf{R}^{n_k}$ represent the *observation* and *decision* associated with stage k of a sequential decision process. The result of observations

$$\xi = (\xi_1, \ldots, \xi_N) \in \mathbf{R}^{\nu_1} \times \ldots \times \mathbf{R}^{\nu_N} = \mathbf{R}^{\nu}$$

and of the sequences of decisions

$$u = (u_1, \ldots, u_N) \in \mathbf{R}^{n_1} \times \ldots \times \mathbf{R}^{n_N} = \mathbf{R}^{n}$$

is a "cost" denoted by $f_0(\xi, u)$. The goal is to find a *decision rule* (or a *recourse function*) $\xi \mapsto u(\xi)$ which minimizes the expected value of this cost subject to certain constraints. An essential constraint on the nature of u is that u must be *nonanticipative*, i.e. the decision u_k at stage k depends only on the past observations $\xi_1, \ldots, \xi_k$, but not on the future $\xi_{k+1}, \ldots, \xi_N$. The problem of finding such an optimal u is called the *optimal recourse problem* (in discrete time). Our aim here is to derive necessary conditions for the optimality of a certain u_0 in the case of costs satisfying Lipschitz assumptions with respect to the decision variables. Actually, assumptions are decomposed into assumptions on the underlying probability space, on the class of decision rules to be considered, and on the objective f.

A. *The probability space.* The underlying probability space associated with the random elements of the problem is $(\Xi, \mathscr{B}_\Xi, P)$, where Ξ is a Borel subset of $\mathbf{R}^\nu$, $\mathscr{B}_\Xi$ is the Borel field on Ξ, and P is a regular Borel probability measure on $(\Xi, \mathscr{B}_\Xi)$.

B. *The class of decision rules.* A decision rule u is said to be *essentially nonanticipative* if it is Borel-measurable and differs only on a set of P-measure zero from some (Borel) measurable function $\tilde{u}$ of the form

$$(\xi_1, \ldots, \xi_N) = \xi \mapsto \tilde{u}(\xi) = \big(\tilde{u}_1(\xi_1), \tilde{u}_2(\xi_1, \xi_2), \ldots, \tilde{u}_N(\xi_1, \ldots, \xi_N)\big).$$

For $k = 1, \ldots, N$, let $\mathscr{B}_k$ be the σ-field generated by the "past" $\xi_1, \ldots, \xi_k$ and completed with respect to P. Then $\{\mathscr{B}_k\}$ is an increasing finite sequence of complete σ-fields with $\mathscr{B}_N = \hat{\mathscr{B}}_\Xi$ (completion of $\mathscr{B}_\Xi$ with respect to P) and *a function u is essentially nonanticipative if and only if for $k = 1, \ldots, N$ the component function $u_k\colon \Xi \to \mathbf{R}^{n_k}$ is $\mathscr{B}_k$-measurable* ([12], § I). As in [12], we adopt the latter property as the general definition of nonanticipativity and work in this notational framework.

In addition to the nonanticipativity constraint, we shall require that the decision rules satisfy almost surely (a.s.)

$$u(\xi) \in \Gamma(\xi), \tag{3.1}$$

where $\Gamma\colon \Xi \rightrightarrows \mathbf{R}^n$ is a nonempty closed-valued measurable multifunction. One can always reduce a state constraint to an abstract formulation like (3.1). The handiness of Γ and the determination of concepts associated with it (like normal cone) depend heavily on the representation of Γ. In that respect, $\Gamma(\xi)$ represented by equality and inequality constraints is easier to handle than $\Gamma(\xi)$ solution set of the variational equality $0 \in \Phi(x, \xi)$. In our multistage program, we are concerned with decision rules which are in $\mathbb{L}^p_{\mathbf{R}^n}$ $(1 \leqslant p \leqslant \infty)$. As in the previous section, we suppose that all the measurable selections of Γ are in $\mathbb{L}^p_{\mathbf{R}^n}$ $(\mathscr{T} \subset \mathbb{L}^p_{\mathbf{R}^n})$.

C. *The objective function.* The function $f_0\colon \Xi \times \mathbf{R}^n \to \bar{\mathbf{R}}$ is an integrand Lipschitz on Γ and we assume that the function $\xi \mapsto |||f_0(\xi, \cdot)|||$ is in $L^q_{\mathbf{R}^n}$ $(1/p+1/q = 1)$.

The multistage stochastic program can then be expressed as

$$(\mathscr{P}) \qquad \text{minimize } I_{f_0}(u) \text{ on } \mathscr{N}_p \cap \mathscr{T},$$

where $\mathscr{N}_p$ represents the constraint of nonanticipativity:

$$\mathscr{N}_p = \{u = (u_1, \ldots, u_N) \in \mathbb{L}^p_{\mathbf{R}^n} \mid u_k \text{ is } \mathscr{B}_k\text{-measurable for } k = 1, \ldots, N\}.$$

To avoid somewhat degenerate cases, it will also be assumed that $\mathscr{N}_p \cap \mathscr{T}$ is nonempty and that

$$\int_{\Xi} |f_0(\xi, \bar{u}(\xi))| \, dP(\xi) < +\infty \qquad \text{for some } \bar{u} \in \mathscr{T}.$$

For purposes of comparison with other hypotheses in stochastic programming, observe that under the assumptions described above, a decision rule which associates sequences of acceptable decisions with almost all ξ should have an expected cost. Thus, if $\bar{f}_0$ is the integrand defined as

$$(\xi, x) \mapsto \bar{f}_0(\xi, x) = f_0(\xi, x) + \delta_{\Gamma(\xi)}(x),$$

then

$$I_{\bar{f}_0}(u) < +\infty \Leftrightarrow u(\xi) \in \Gamma(\xi) \text{ a.s.}$$

3.2. Necessary conditions for optimality.

THEOREM 3.1. *If $u_0 \in L^p_{R^n}$ is an optimal decision rule for the multistage stochastic program $(\mathscr{P})$, then*

(a) *$u_0 \in \mathcal{N}_p$ and $u_0(\xi) \in \Gamma(\xi)$ a.s.,*

(b) *there exists a $\varrho^* = (\varrho_1^*, \ldots, \varrho_N^*) \in L^q_{R^n}$ satisfying*

$$\varrho^*(\xi) \in \partial f_0(\xi, u_0(\xi)) \quad \text{and} \quad \|\varrho^*(\xi)\| \leqslant k(\xi) \text{ a.s.} \tag{3.2}$$

and also

$$\mathrm{E}\{\varrho_k^* \mid \mathscr{B}_k\} = 0 \text{ a.s.} \quad \text{for } k = 1, \ldots, N. \tag{3.3}$$

Proof. $\mathcal{N}_p$ is a linear subspace of $L^p_{R^n}$. Therefore, the normal cone to $\mathcal{N}_p$ at $u_0 \in \mathcal{N}_p$ is independent of u_0 and amounts to the orthogonal subspace in the topological dual space of $L^p_{R^n}$ under the pairing described in Section 2. Let $1 \leqslant p < \infty$; as noticed in [11] and [12], the annihilator subspace $\mathcal{N}_p^{\perp}$ consists of elements $\varrho^* = (\varrho_1^*, \ldots, \varrho_N^*) \in L^q_{R^n}$ $(\varrho_k^* \in L^q_{R^{n_k}})$ which satisfy the *martingale property*

$$\mathrm{E}\{\varrho_k^* \mid \mathscr{B}_k\} = 0 \text{ a.s.} \quad \text{for } k = 1, \ldots, N.$$

In the case where $p = \infty$, those elements in $(L^{\infty}_{R^n})^*$ which are in $L^1_{R^n}$ can be described with the same martingale property.

According to Theorem 2.2, if u_0 is an optimal solution to problem $(\mathscr{P})$, then there exists a $\varrho^* \in L^q_{R^n}$ such that (3.2) holds and $-\varrho^* \in \mathcal{N}_p^{\perp}$. Hence the announced optimality conditions are derived from the above description of $\mathcal{N}_p^{\perp}$.

When the given problem has some more structural characteristics, necessary conditions may be made more precise and thereby more informatory. We shall especially examine the case where

$$\Gamma(\xi) = \{x \in R^n \mid f_i(\xi, x) \leqslant 0 \text{ for } i = 1, \ldots, \alpha \text{ and } f_i(\xi, x) = 0 \text{ for } i = \alpha+1, \ldots, \beta\} \cap \Gamma_0(\xi). \tag{3.4}$$

A. *The locally Lipschitz case.* We suppose here that for all $x \in \Gamma(\xi)$

$$\begin{aligned} &f_0(\xi, \cdot), f_1(\xi, \cdot), \ldots, f_\alpha(\xi, \cdot) \text{ are Lipschitz around } x, \\ &f_{\alpha+1}(\xi, \cdot), \ldots, f_\beta(\xi, \cdot) \text{ are continuously differentiable at } x. \end{aligned} \tag{3.5}$$

The multifunction Γ_0 entering in the definition of the additional constraint is supposed to be nonempty, closed-valued, and measurable. Note that the local Lipschitz property of $f_0(\xi, \cdot)$ in (3.5) ensures that $f_0(\xi, \cdot)$ is Lipschitz on $\Gamma(\xi)$ whenever $\Gamma(\xi)$ is bounded. In addition to the assumptions laid out above and earlier, representation (3.4) is supposed to satisfy the following *constraint qualification*:

(CQ) for all $u \in \mathcal{N}_p \cap \mathcal{T}$, $\nabla f_{\alpha+1}(\xi, u(\xi)), \ldots, \nabla f_\beta(\xi, u(\xi))$ are linearly independent a.s. and there exists a $d: \Xi \to \boldsymbol{R}^n$ such that

$$d(\xi) \in \operatorname{int} \mathscr{C}_{\Gamma_0(\xi)}(u(\xi)),$$

$$f_i^\circ(\xi, u(\xi); d(\xi)) < 0, \quad i = 1, \ldots, \alpha,$$

$$\langle \nabla f_i(\xi, u(\xi)), d(\xi) \rangle = 0 \text{ a.s.}, \quad i = \alpha+1, \ldots, \beta.$$

Under these conditions, relation (b) in Theorem 3.1 takes the form

(b′) *There exists a* $\varrho^* = (\varrho_1^*, \ldots, \varrho_N^*) \in L^q_{\boldsymbol{R}^n}$ *satisfying* (3.3) *and there exist* $\lambda_1, \ldots, \lambda_\beta$ *in* $L^q_{\boldsymbol{R}^n}$ *such that*

$$\lambda_i(\xi) \geqslant 0, \quad \lambda_i(\xi) f_i(\xi, u_0(\xi)) = 0 \ \textit{a.s.} \quad \textit{for } i = 1, \ldots, \alpha,$$

$$(3.6) \quad \varrho^*(\xi) \in \partial f_0(\xi, u_0(\xi)) + \sum_{i=1}^{\alpha} \lambda_i(\xi)\, \partial f_i(\xi, u_0(\xi)) + \\ + \sum_{i=\alpha+1}^{\beta} \lambda_i(\xi) \nabla f_i(\xi, u_0(\xi)) + \mathcal{N}_{\Gamma_0(\xi)}(u_0(\xi)) \ \textit{a.s.}$$

The proof is simply a matter of decomposition and representation of normal cones; all the necessary ingredients are laid out in [4]. Moreover, the reader will find himself the generalization of (CQ) for the special structure that would be the "directionally Lipschitz case"; all the material for that purpose may be found in [10].

B. *The separable case.* In order to gain insight into this case without being wrapped in technical assumptions, we shall simplify our approach by presupposing we are in the locally Lipschitz situation (case A above) and by dropping the additional constraint $\Gamma_0(\xi)$. By a *separable problem* we mean that

(i) $f_i(\xi, u) = \sum_{k=1}^{N} f_{i,k}(\xi, u_k)$ for $i = 0, 1, \ldots, \beta$,

(ii) $f_{i,k}\colon \xi \mapsto f_{i,k}(\xi, u_k)$ are $\mathscr{B}_k$-measurable functions for $i = 0, 1, \ldots, \beta$ and $k = 1, \ldots, N$.

In such a situation, part (b) of necessary conditions for optimality in Theorem 3.1 takes a "decomposed" form.

THEOREM 3.2. *If* $u_0 \in L^p_{\boldsymbol{R}^n}$ *is an optimal decision rule for problem* ($\mathscr{P}$), *supposed separable, then*

(a) *for* $u_0 \in \mathcal{N}_p, f_i(\xi, u_0(\xi)) \leqslant 0$ *a.s. for* $i = 1, \ldots, \alpha$ *and* $f_i(\xi, u_0(\xi)) = 0$ *a.s. for* $i = \alpha+1, \ldots, \beta$;

(b″) *there exist* $\lambda_1, \ldots, \lambda_\beta$ *in* $L^q_{\boldsymbol{R}^n}$ *such that*

$$\lambda_i(\xi) \geqslant 0, \quad \lambda_i(\xi) f_i(\xi, u_0(\xi)) = 0 \ \textit{a.s.} \quad \textit{for } i = 1, \ldots, \alpha,$$

(3.7) $$0 \in \partial f_{0,k}(\xi, u_{0,k}(\xi)) + \sum_{i=1}^{\alpha} \mathrm{E}\{\lambda_i \mid \mathscr{B}_k\}(\xi)\, \partial f_{i,k}(\xi, u_{0,k}(\xi)) + \\ + \sum_{i=\alpha+1}^{\beta} \mathrm{E}\{\lambda_i \mid \mathscr{B}_k\}(\xi)\, \nabla f_{i,k}(\xi, u_{0,k}(\xi)) \quad a.s. \quad for\ k = 1, \ldots, N.$$

Proof. Due to the separable structure of all the functions involved in the problems, we have

$$\partial f_i(\xi, u) = \mathop{\mathsf{X}}_{k=1}^{N} \partial f_{i,k}(\xi, u_k) \quad \text{for } u = (u_1, \ldots, u_N).$$

By rewriting relation (3.6) of (b′) in such a context, we get

(3.8) $$\varrho_k^* \in \partial f_{0,k}(\xi, u_{0,k}(\xi)) + \sum_{i=1}^{\alpha} \lambda_i(\xi)\, \partial f_{i,k}(\xi, u_{0,k}(\xi)) + \\ + \sum_{i=\alpha+1}^{\beta} \lambda_i(\xi)\, \nabla f_{i,k}(\xi, u_{0,k}(\xi)) \quad \text{a.s.} \quad \text{for } k = 1, \ldots, N.$$

Since $\xi \mapsto u_{0,k}(\xi)$ is $\mathscr{B}_k$-measurable, so are the multifunction

$$\xi \rightrightarrows \partial f_{i,k}(\xi, u_{0,k}(\xi)) \quad \text{for } i = 1, \ldots, \alpha$$

and the mapping

$$\xi \mapsto \nabla f_{i,k}(\xi, u_{0,k}(\xi)) \quad \text{for } i = \alpha+1, \ldots, \beta.$$

Now, we use a result concerning the conditional expectation of multifunctions [15] which claims that if $\Delta\colon \Xi \rightrightarrows \boldsymbol{R}^n$ is a convex compact-valued $\mathscr{B}$-measurable multifunction (with $\mathscr{B}$ P-complete), then $\mathrm{E}(\Delta \mid \mathscr{B})$ is a.s. equal to Δ.

By plugging the relation $\mathrm{E}(\varrho_k^* \mid \mathscr{B}_k) = 0$ into (3.8), we get the desired condition.

Remarks. In the convex case (i.e. $f_i(\xi, \cdot)$ convex for $i = 1, \ldots, \alpha$ and $f_i(\xi, \cdot)$ affine for $i = \alpha+1, \ldots, \beta$), relation (3.7) can be translated in terms of pointwise minimization using information pertinent to stage k. For the salient features of these conditions in the "decomposed" form in the theoretical aspect as well as from the computational viewpoint, consult [12], § 3.A. In our approach, the "conditional multipliers" have been obtained in a rather mechanical way and this process does not have the flavour of the duality approach in the convex case [12].

REFERENCES

[1] M. Benamara, *Points extrêmaux, multiapplications et fonctionnelles intégrales*, Thèse de 3ème cycle, Université Scientifique et Médicale de Grenoble, 1975.

[2] H. F. Clarke, *A new approach to Lagrange multipliers*, Math. Operations Research 2 (1) (1976), p. 165-174.

[3] — *Generalized gradients of Lipschitz functionals*, Madison R.C. Technical Summary Report, University of Wisconsin, 1976.

[4] J.-B. Hiriart-Urruty, *Conditions nécessaires d'optimalité pour un programme stochastique avec recours*, SIAM J. Control Optimization 16 (2) (1978), p. 317-329.

[5] — *Tangent cones, generalized gradients and mathematical programming in Banach spaces*, Math. Operations Research 4 (1) (1979), p. 79-97.

[6] — *Extension of Lipschitz functions*, J. Math. Anal. Appl. 77 (1980), p. 539-554.

[7] R. T. Rockafellar, *Integrals which are convex functionals. II*, Pacific J. Math. 39 (1971), p. 439-469.

[8] — *Integral functionals, normal integrands and measurable selections*, in: *Nonlinear operators and the calculus of variations* (ed. L. Waelbroeck), Springer-Verlag Series, Lecture Notes in Math., 1976.

[9] — *Generalized directional derivatives and subgradients of nonconvex functions*, Canad. J. Math. 32 (1980), p. 257-280.

[10] — *Directionally Lipschitzian functions and subdifferential calculus*, Proc. London Math. Soc. 39 (1979), p. 331-355.

[11] — and R. J.-B. Wets, *Nonanticipativity and L^1-martingales in stochastic optimization problems*, Mathematical Programming Studies, No. 6 (1976), p. 170-187.

[12] — *The optimal recourse problem in discrete time: L^1-multipliers for inequality constraints*, SIAM J. Control Optimization 16 (1) (1978), p. 16-36.

[13] L. Thibault, *Propriétés des sous-différentiels de fonctions localement lipschitziennes définies sur un espace de Banach séparable*, Thèse de Doctorat de Spécialité, Université de Montpellier, 1976.

[14] — *Calcul sous-différentiel et calcul des variations en dimension infinie*, Journées d'Analyse Non Convexe (Mai 1977), Bull. Soc. Math. France, Mémoire 60 (1979), p. 57-85.

[15] B. Van Cutsem, *Eléments aléatoires à valeurs convexes compactes*, Thèse de Doctorat ès Sciences, Université Scientifique et Médicale de Grenoble, 1971.

[16] T. Zolezzi, *Well posed optimization problems for integral functionals*, J. Optimization Theory Appl. 31 (1980), p. 418-430.

Université Paul Sabatier (Toulouse III)
U.E.R. Mathématiques, Informatique, Gestion
118, route de Narbonne
31062 Toulouse Cédex, France

Received on 20. 9. 1979;
revised version on 2. 8. 1980