



HAL
open science

New size X curvature conditions for strict quasiconvexity of sets

Guy Chavent

► **To cite this version:**

Guy Chavent. New size X curvature conditions for strict quasiconvexity of sets. [Research Report] RR-1120, INRIA. 1989. inria-00075439

HAL Id: inria-00075439

<https://inria.hal.science/inria-00075439>

Submitted on 24 May 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

INRIA

**UNITÉ DE RECHERCHE
INRIA-ROQUENCOURT**

Institut National
de Recherche
en Informatique
et en Automatique

Domaine de Voluceau
Rocquencourt
BP 105
78153 Le Chesnay Cedex
France
Tél. (1) 39 63 55 11

Rapports de Recherche

N° 1120

*Programme 7
Calcul Scientifique,
Logiciels Numériques et Ingénierie Assistée*

**NEW SIZE \times CURVATURE
CONDITIONS FOR STRICT
QUASICONVEXITY OF SETS**

Guy CHAVENT

Novembre 1989



* RR - 1120 *

**NEW SIZE × CURVATURE CONDITIONS
FOR
STRICT QUASICONVEXITY OF SETS**

**DE NOUVELLES CONDITIONS DE
TAILLE × COURBURE
POUR QU'UN ENSEMBLE SOIT STRICTEMENT QUASICONVEXE**

Guy CHAVENT

Abstract

Given a closed, non necessarily convex set D of an Hilbert space, we consider the problem of the existence of a neighborhood \mathcal{U} on which the projection on D is uniquely defined and lipschitz continuous, and such that the corresponding minimization problem has no local minima. After having equipped the set D with a family \mathcal{P} of pathes playing for D the role the segments play for a convex set, we define the notion of strict quasi convexity of (D, \mathcal{P}) , which shall ensure the existence of such a neighborhood \mathcal{U} . Two constructive sufficient condition for the strict-quasiconvexity of D are given, the R_G -size \times curvature condition and the Θ -size \times curvature condition, which both amount to checking for the strict positivity of quantities defined by simple formulas in terms of arc length, tangent vectors and radii of curvatures along all pathes of \mathcal{P} . An application to the study of wellposedness and local minima of a non-linear least square problem is given.

Résumé

Etant donné un ensemble fermé D non nécessairement convexe, on considère le problème de l'existence d'un voisinage \mathcal{U} sur lequel la projection sur D soit définie de façon unique et lipschitz continue, et tel que le problème d'optimisation correspondant ne possède aucun minimum local. Après avoir équipé l'ensemble D d'une famille \mathcal{P} de chemins jouant pour D le rôle que les segments jouent pour un ensemble convexe, on introduit la notion de quasiconvexité stricte de (D, \mathcal{P}) , dont on montre qu'elle garantit l'existence du voisinage \mathcal{U} cherché. On donne ensuite deux conditions suffisantes constructive pour la quasiconvexité stricte de D , la condition R_G de taille \times courbure et la condition Θ de taille \times courbure, qui toutes deux se ramènent à tester la stricte positivité de quantités définies par des formules simples à partir de l'abscisse curviligne, des vecteurs tangents et des rayons de courbure le long de tous les chemins de \mathcal{P} . On termine par l'application à un problème de moindres carrés non linéaires, pour lequel on détermine des conditions assurant qu'il est bien posé et ne possède pas de minima locaux.

Keywords

Projection theory, approximation theory, non-linear least squares, inverse problems.

Mots clefs

Théorie de la projection, théorie de l'approximation, moindres carrés non linéaires, problèmes inverses.

Summary

1 - Introduction

2 - Equipping the set with paths.

3 - Quasiconvex and strictly quasiconvex sets.

4 - The γ -size \times curvature condition for quasiconvex sets.

5 - The R_G -size \times curvature condition : a new sufficient condition for strictly quasiconvex sets.

6 - Obtaining lower bounds on the global radius of curvature : The θ -size \times curvature Condition.

7 - One application to non-linear least squares.

1 - INTRODUCTION

Let

(1.1) F an Hilbert space,

(1.2) $D \subset F$ a (non necessarily convex) subset

be given. This paper is devoted to the study of the Hilbert projection on D . It is known (cf.[5]) that, when D is closed and bounded, the set of all $z \in F$ which admit a unique projection on D contains a dense countable intersection of open sets, called the Edelstein set, and defined by :

(1.3) $ES = \{z \in F \mid \forall \epsilon > 0, \exists \eta > 0 \text{ s.t.}$
 $X_j \in D, \|X_j - z\| \leq d(z, D) + \eta \text{ } j=0,1 \text{ implies } \|X_0 - X_1\| \leq \epsilon\}.$

This result is not precise enough for the application to non-linear least-square problems we have in mind, first because it gives only a generic result, and second because it does not give any insight into possible local minima on D of the "distance to z " function, whose presence or absence is determining when it comes to the actual numerical determination of the projection using an optimization algorithm.

Hence it would be very useful to find conditions on D which ensures i) that the Edelstein set contains some neighborhood \mathcal{U} of D , and ii) that ideally the $y \rightarrow d(y, z)$ function has no local minima on D whenever $z \in \mathcal{U}$. Necessary and sufficient conditions for point i), involving the local curvature of D , have been extensively studied, but point ii) has not received much attention.

The notion of the curvature of a set D is both delicate to define and difficult to apply to non-linear least square problems. Hence looking only for sufficient conditions for i) and ii) to holds makes it possible to work only with the elementary notion of the curvature of a curve, provided that the set D has been equipped with a collection \mathcal{P} of pathes P which shall play for D the role of the segments for a convex set. This will yield an easy application to non-linear least-square problems, where the pathes P can be very naturally defined as the image of the segments of the admissible set by the mapping to be inverted !

An answer along this line to i) and partially to ii) has been given in [3], where the notion of quasiconvex set was introduced together with a constructive sufficient condition (the γ -size \times curvature condition), and where an application the study of well posedness and local minima of abstract non-linear least square problems was given.

Closed quasiconvex sets, which possers a neighborhood \mathcal{U} on which the projection exists, is unique and lipchitz continue, satisfy fully point i), but they don't eliminate local minima, which may exist provided their value is "large enough". Moreover the γ -size \times curvature condition does not seem to be very sharp, as for example it is far from recognizing all quasiconvex arcs of circles (cf.figure 3.1 below).

We remedy in this paper to the above mentioned weaknesses :

We first introduce the slightly stronger notion of strictly quasiconvex set, which fully eliminates the possibility of local minima, as soon as the point to be projected is in the associated neighborhood \mathcal{U} .

Then we introduce the notion of global radius of curvature R_G associated to the family of pathes \mathcal{P} , and show that condition $R_G > 0$ (called the R_G -size \times curvature condition) implies strict quasiconvexity of the set. This condition is shown to be sharp (see figure 3.2 below) as it recognizes exactly all strictly quasiconvex curves.

Third, we calculate a lower bound to R_G in term of the usual radius of curvature R , the length Δ and the deflection Θ associated to the family of pathes \mathcal{P} ; this of course leads to another sufficient condition for strict quasiconvexity of the set, called the Θ -size \times curvature condition, which is also sharp as it recognizes exactly all strictly quasiconvex curves made of one arc of circle and one segment.

All geometrical quantities γ, R_G, R, Δ and Θ are defined by infimum or supremum, over the collection of pathes \mathcal{P} , of quantities easily computed on each path from the velocities and accelerations along the path.

Finally we give an application to a non-linear least square problem in the case where the derivative of the non-linear mapping to be inverted has a uniformly bounded pseudoinverse. A more detailed application to non-linear least square problems and their regularized versions can be found in reference [4].

2 - EQUIPPING THE SET WITH PATHES

The first step of our construction consists in choosing in the possibly non convex set D a collection \mathcal{P} of pathes P which shall play for D the role the segments play for a convex set. We summarize here the corresponding definitions and notation, following paragraph 2.1 of [3].

Definition 2.1 (pathes)

A mapping $P : [0, L] \rightarrow D$ is a path iff :

$$(2.1) \quad v \rightarrow P(v) \text{ is of class } \mathcal{C}^2,$$

$$(2.2) \quad \| P'(v) \|_F = 1 \quad \forall v \in [0, L].$$

We refer to [3] for sufficient condition for a mapping to be reparametrizable such as it is a path.

Definition 2.2 (attributes of a path)

Let a path P be given. Then,

$$(2.3) \quad v \in [0, L] \text{ is the arc-length along } P$$

$$(2.4) \quad \delta(P) \triangleq L \text{ is the length of } P$$

$$(2.5) \quad v(v) \triangleq P'(v) \text{ is the unit tangent vector to } P \text{ at } P(v)$$

$$(2.6) \quad a(v) \triangleq P''(v) \text{ is the acceleration vector to } P \text{ at } P(v)$$

$$(2.7) \quad \rho(v) \triangleq \|a(v)\|^{-1} \in \mathbb{R}^+ \cup \{+\infty\} \text{ is the radius of curvature of } P \text{ at } P(v)$$

$$(2.8) \quad R(P) \triangleq \inf_{v \in [0, \delta(P)]} \rho(v) \geq 0 \text{ is the smallest radius of curvature along } P.$$

As we shall use only the parametrization by the arc length in this paper, we denote it simply by v (notice that in reference [3], the (reduced) arc length was denoted by \bar{v} , and all corresponding quantities weared bars, as quoted at beginning of paragraph 3 of [3] ; hence $R(P)$ in this paper corresponds to $\bar{R}(\bar{P})$ in reference [3] etc...).

Definition 2.3 (collection of pathes)

A set of pathes \mathcal{P} is a collection of pathes iff :

$$(2.9) \quad \mathcal{P} \text{ is made of pathes}$$

$$(2.10) \quad \mathcal{P} \text{ is complete, i.e. } \forall X, Y \in D, X \neq Y, \exists P \in \mathcal{P}, \text{ such that } P(0) = X, \\ P(\delta(P)) = Y$$

$$(2.11) \quad \mathcal{P} \text{ is stable with respect to restriction, i.e. } \forall P \in \mathcal{P}, \forall v', v'' \in [0, \delta(P)], v' < v'', \\ \text{the path } \tilde{P} : v \in [0, v'' - v'] \rightarrow P(v' + v) \text{ belongs to } \mathcal{P}.$$

The set of pathes \mathcal{P} should contain the minimum number of pathes which allow to satisfy (2.10) (typically, one and only one path of \mathcal{P} connects two given distincts points X and Y of D , but we don't put this uniqueness in the requirement for a collection of pathes, as we want not to exclude the case of minimum length pathes for example).

Definition 2.4 (maximal pathes)

A subset \mathcal{P}_M of \mathcal{P} is said to be a collection of maximal pathes for \mathcal{P} if and only if :

$$(2.12) \quad \mathcal{P} = \cup \{P' \mid P' \text{ is a subpath of } P\} \\ P \in \mathcal{P}_M$$

Of course, such \mathcal{P}_M always exists, as $\mathcal{P}_M = \mathcal{P}$ satisfies (2.12) by virtue of (2.11) ! But a more interesting case will be found when there exists a smallest collection of maximal pathes : this will be the case in most of the applications to non-linear least squares (see [3] [4]), where \mathcal{P} is the image, by a mapping ϕ , of the segments of a closed convex set C , so that $\mathcal{P}_M = \{\phi([x, y]), x, y, \in \partial C\}$ satisfies obviously (2.12).

In all the sequel, we shall suppose that a collection of maximal pathes \mathcal{P}_M has been chosen.

3 - QUASICONVEX AND STRICTLY QUASICONVEX SETS

Let the set D be equipped with a collection of paths \mathcal{P} . To a given point $z \in F$, a path $P \in \mathcal{P}$, and an arc length value $v \in [0, \delta(P)]$, we associate the number :

$$(3.1) \quad k(z, P; v) = \langle z - P(v), a(v) \rangle_F$$

whose geometrical interpretation is given by (with the notations of figure 3.1) :

$$(3.2) \quad k(z, P; v) = \frac{d(v)}{\rho(v)} \cos \theta = \frac{\overline{MH}}{\overline{MC}}.$$

One may then associate to z and P the number :

$$(3.3) \quad k(z, P) = \text{Max}_{v \in [0, \delta(P)]} k(z, P; v)$$

and to z and $\eta > 0$ the number :

$$(3.4) \quad k(z, \eta) = \text{Sup}_{P \in \mathcal{P}(z, \eta)} k(z, P)$$

where the subcollection of paths $\mathcal{P}(z, \eta)$ is defined by :

$$(3.5) \quad \mathcal{P}(z, \eta) = \left\{ P \in \mathcal{P} \mid \|P(j) - z\|_F < d(z, D) + \varepsilon, j = 0, \delta(P) \right\}$$

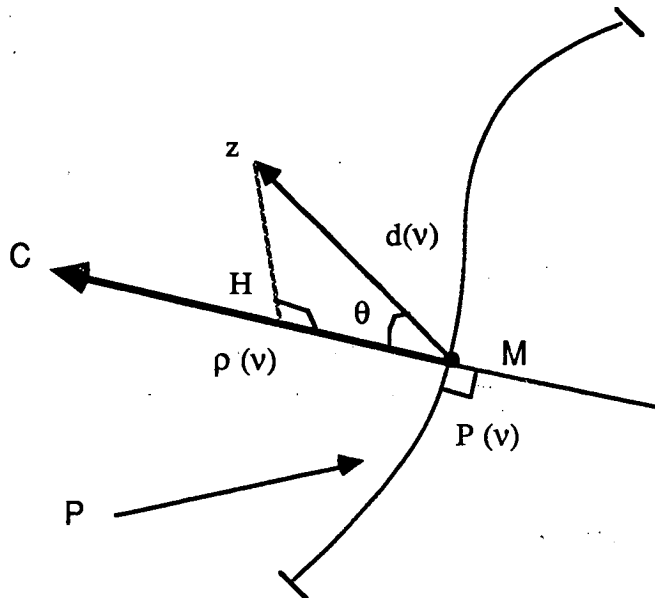


Figure 3.1. Notations for the geometrical interpretation of $k(z, P; v) = \langle z - P(v), a(v) \rangle_F$: C is the center of curvature of P at $P(v)$, and H is the orthogonal projection of z on the 1-D affine variety parallel to $a(v)$ and passing through $M = P(v)$.

We can now give the :

Definition 3.1 (quasiconvexity of sets)

A set (D, \mathcal{P}) is said to be quasiconvex iff :

- i) \mathcal{P} is a collection of pathes,
- ii) There exists a neighborhood \mathcal{U} of D in F , and a lower semi continuous function $\varepsilon : \mathcal{U} \rightarrow]0, +\infty[$ such that :

$$(3.6) \quad \begin{array}{l} z \in \mathcal{U} \\ 0 < \eta < \varepsilon(z) \end{array} \Rightarrow k(z, \eta) < 1$$

This is a slightly upgraded version of the definition of quasiconvexity given in [3] (where the $z \rightarrow \varepsilon(z)$ function was required to be continuous). All the results in [3] carry over when quasiconvexity is defined using definition 3.1, and this upgraded definition allows us to prove the

Proposition 3.1

Let (D, \mathcal{P}) be quasiconvex.

Then there exists a largest open neighborhood \mathcal{U} of D , and a largest ℓ .s.c. function $\varepsilon : \mathcal{U} \rightarrow]0, +\infty[$ satisfying the definition of quasiconvexity.

Proof

Let us denote by $\mathcal{U}_i, \varepsilon_i, i \in I$ all open neighborhood and ℓ .s.c. functions satisfying definition 3.1. Then :

$$(3.7) \quad \mathcal{U} = \bigcup_{i \in I} \mathcal{U}_i$$

is an open subset of F . If we denote by $\tilde{\varepsilon}_i$ the continuation of ε_i to \mathcal{U} by zero outside of \mathcal{U}_i , then $\tilde{\varepsilon}_i$ is ℓ .s.c. as ε_i is l.s.c. on \mathcal{U}_i and \mathcal{U}_i is open. Define then :

$$(3.8) \quad \varepsilon(z) = \text{Sup}_{i \in I} \tilde{\varepsilon}_i(z) \quad \forall z \in \mathcal{U}$$

which is ℓ .s.c. as supremum of a family of l.s.c. functions. Hence \mathcal{U}, ε will satisfy the definition of quasiconvexity as soon as they satisfy (3.6), which we prove now. Let $z \in \mathcal{U}$ and $0 < \eta < \varepsilon(z)$ be given, and set :

$$\alpha = (\varepsilon(z) - \eta) / 2 > 0$$

From the definition (3.8) of $\varepsilon(z)$, there exists $i_0 \in I$ such that :

$$\tilde{\varepsilon}_{i_0}(z) \geq \varepsilon(z) - \alpha > \eta > 0$$

This proves, as $\tilde{\varepsilon}_{i_0}(z) > 0$, that $z \in \mathcal{U}_{i_0}$. Hence \mathcal{U}_{i_0} and ε_{i_0} satisfy the hypothesis of (3.6), so that :

$$k(z, \eta) < 1$$

which proves that \mathcal{U} and $\varepsilon(z)$ satisfy (3.6). ■

We summarize now in Proposition 3.2 thru Theorem 3.6 the properties of quasiconvex sets, mostly duplicated from [3], to which we refer for their proof.

Proposition 3.2

Let (D, \mathcal{P}) be quasiconvex. Then the associated neighborhood \mathcal{U} is included in the Edelman set (1.3), and for any $z \in \mathcal{U}$, $0 < \eta < \varepsilon(z)$ and $P \in \mathcal{P}(z, \eta)$,

- i) $v \rightarrow f(v) \triangleq \|P(v)-z\|_F^2$ is strictly convex
- ii) $v \rightarrow d(v) \triangleq \|P(v)-z\|_F$ is strictly quasiconvex

Proof

For $z \in \mathcal{U}$ and $P \in \mathcal{P}(z, \eta)$ one has

$$f''(v) \geq 2(1-k(z, \eta)) > 0$$

which proves i) and ii), and also shows that the $v \rightarrow f(v) + v(\delta(P)-v)(1-k(z, \eta))$ function is convex. Hence :

$$f\left(\frac{\delta(P)}{2}\right) + \frac{\delta(P)^2}{4} (1-k(z, \eta)) \leq \frac{1}{2} f(0) + \frac{1}{2} f(\delta(P))$$

But :

$$f\left(\frac{\delta(P)}{2}\right) \geq d(z, D)^2$$

$$f(j) \leq (d(z, D) + \eta)^2 \quad j=0, \delta(P)$$

so that :

$$\delta(P)^2 \leq 4(1-k(z, \eta))^{-1} \eta (2d(z, D) + \eta)$$

If we choose for P the path connecting two points $X_j \in D$ $j=0,1$ and satisfying :

$$\|X_j - z\| \leq d(z, D) + \eta$$

we see that :

$$\|X_0 - X_1\|_F \leq \delta(P) \leq 2(1-k(z, \eta))^{-1/2} \eta^{1/2} (2d(z, D) + \eta)^{1/2} \rightarrow 0 \text{ when } \eta \rightarrow 0,$$

which prove that z belongs to the Edelman set. ■

Proposition 3.2. implies that the distance of any point z of \mathcal{U} to any path of \mathcal{P} having its extremities "not too far" from z is strictly quasiconvex.

Proposition 3.3

Let (D, \mathcal{P}) be quasiconvex.

Then, for any $P \in \mathcal{P}$, the $v \rightarrow P(v)$ mapping is injective.

Hence a path P of a quasiconvex set (D, \mathcal{P}) cannot "make a loop". We recall now the main properties of quasiconvex sets concerning projection :

Theorem 3.4 (uniqueness of the projection)

If :

(D, \mathcal{P}) is quasiconvex

$z \in \mathcal{U}$

Then :

The "distance to z " function has at most one global minimum over D . Moreover, all possible other local minima yield a value larger than or equal to $d(z,D) + \epsilon(z)$.

We see that quasiconvex sets ensure uniqueness of the global minimum of the "distance to z " function, but don't eliminate completely local minima, which are only rejected "far enough" from D .

Theorem 3.5 (continuity of the projection)

If :

$$(3.7) \quad (D, \mathcal{P}) \text{ is quasiconvex}$$

$$(3.8) \quad z_0, z_1 \in \mathcal{U} \text{ admit projections } X_0, X_1 \text{ on } D$$

$$(3.9) \quad \left\{ \begin{array}{l} z_0, z_1 \text{ one close enough, precisely if there exists } d \geq 0 \text{ such that} \\ \|z_0 - z_1\|_F + \text{Max}_{j=0,1} d(z_j, D) < d < \text{Min}_{j=0,1} \{d(z_j, D) + \epsilon(z_j)\} \end{array} \right.$$

Then :

$$(3.10) \quad \|X_0 - X_1\|_F \leq \delta(X_0, X_1) \leq (1-k)^{-1} \|z_0 - z_1\|_F$$

where :

$$(3.11) \quad \delta(X_0, X_1) = \begin{cases} \text{Sup}_{P \in \mathcal{P}} \delta(P) & \text{if } X_0 \neq X_1 \\ P: X_0 \rightarrow X_1 \\ 0 & \text{if } X_0 = X_1 \end{cases}$$

$$(3.12) \quad k = (k_0 + k_1)/2$$

$$(3.13) \quad k_j = k(z_j, \eta_j) \quad j = 0, 1$$

$$(3.14) \quad 0 < \eta_j = d - d(z_j, D) < \epsilon(z_j) \quad j = 0, 1$$

Remark first that lower semi continuity of $z \rightarrow \epsilon(z)$ is just enough to ensure the existence of a d satisfying (3.9) when z_0 and z_1 are close enough.

Let us also comment shortly on the definition (3.11) of the "curvilinear distance" $\delta(X_0, X_1)$ of two points of D : The supremum in (3.11) is not an error. In most of the case, the collection of pathes \mathcal{P} will be chosen in such a way that it contains one and only one path connecting any two distincts points of D , in which case $\delta(X_0, X_1) = \delta(P)$, which seems reasonable. In the case where \mathcal{P} has to be chosen rich enough so that there may exist many pathes connecting two distinct points of D , the majoration (3.10) has to hold for all pathes connecting X_0 to X_1 (and hence for the worst, i.e. the longest), as in the definition of quasiconvexity the $k(z, \eta)$ quantities involve a supremum over pathes with prescribed extremities, and hence take into account all pathes with the same ends. Notice also that if \mathcal{P} is made of the geodesic curves connecting two distinct points of D , there may be many pathes P connecting X_0 to X_1 , but all with the same length $\delta(P)$, so that definition (3.11) is also very natural in that case.

Theorem 3.6 (Existence of the projection)

If :

(3.15) (D, \mathcal{P}) is quasiconvex

(3.16) D is closed in F

Then :

(3.17) any $z \in \mathcal{U}$ has a unique projection \hat{X} on D

(3.18) $\left\{ \begin{array}{l} \text{any minimizing sequence is converging towards } \hat{X} \\ \text{for both the } \|X-Y\|_F \text{ and the } \delta(X,Y) \text{ distances} \end{array} \right.$

We introduce now a new, stronger definition which will keep all the above mentioned nice properties of quasiconvex sets with respect to projection, but will be much better behaved with respect to the problem of local minima :

Definition 3.7 (strict quasiconvexity of sets)

A set (D, \mathcal{P}) is said to be strictly quasiconvex iff :

i) (D, \mathcal{P}) is quasiconvex for some neighborhood \mathcal{U} and function $z \rightarrow \epsilon(z)$,

ii) \mathcal{U} and ϵ satisfy moreover :

(3.19) $\begin{array}{l} z \in \mathcal{U} \\ P \in \mathcal{P} \\ d(z,P) < d(z,D) + \epsilon(z) \end{array} \Rightarrow \begin{array}{l} \text{the "distance to z"} \\ \text{function is strictly} \\ \text{quasiconvex along the path P.} \end{array}$

The additional hypothesis (3.19) means (compare with proposition 3.2) that the distance of any point z of \mathcal{U} to any path of \mathcal{P} which is "not too far" of z is strictly quasiconvex. One checks easily that proposition 3.1 extends to strictly quasiconvex sets :

Proposition 3.8

Let (D, \mathcal{P}) be strictly quasiconvex. Then there exists a largest open neighborhood \mathcal{U} of D , and a largest l.s.c. function $\epsilon : \mathcal{U} \rightarrow]0, +\infty]$ satisfying the definition of strict quasiconvexity.

Of course, any neighborhood \mathcal{U} associated by definition 3.7. to a strictly quasiconvex set (D, \mathcal{P}) is included in the Edelstein set (1.3) by virtue of proposition 3.2.

The main interest of strict quasiconvexity comes from the :

Theorem 3.9 (Uniqueness of the projection)

If :

(D, \mathcal{P}) is strictly quasiconvex

$z \in \mathcal{U}$

Then :

The "distance to z " function has at most one global minimum over D and, provided this global minimum exists, no other local minima.

Proof

We have only to prove the last part of the conclusion. Hence let $X_0 \in D$ be one global minimum of the "distance to z " function, and suppose that this function admits a local minimum at $X_1 \neq X_0$. By definition of X_0 we have :

$$(3.20) \quad \|z - X_0\| = d(z, D)$$

Let then $P \in \mathcal{P}$ connect X_0 to X_1 . From (3.20) we get :

$$(3.21) \quad d(z, P) = \|z - X_0\| = d(z, D) < d(z, D) + \varepsilon(z)$$

Using then the property (3.19) of strictly quasiconvex sets we find that the "distance to z " function is strictly quasiconvex along the path P , which is impossible, as this function has a global minimum at $v = 0$ and a local minimum at $v = \delta(P)$! ■

We illustrate this paragraph on quasiconvex and strictly quasiconvex sets by considering the very simple case of a set D consisting in an arc of circle of radius R and of length L . Simple geometric considerations show that :

$$(3.22) \quad D \text{ is quasiconvex iff } \frac{L}{R} < 2\pi$$

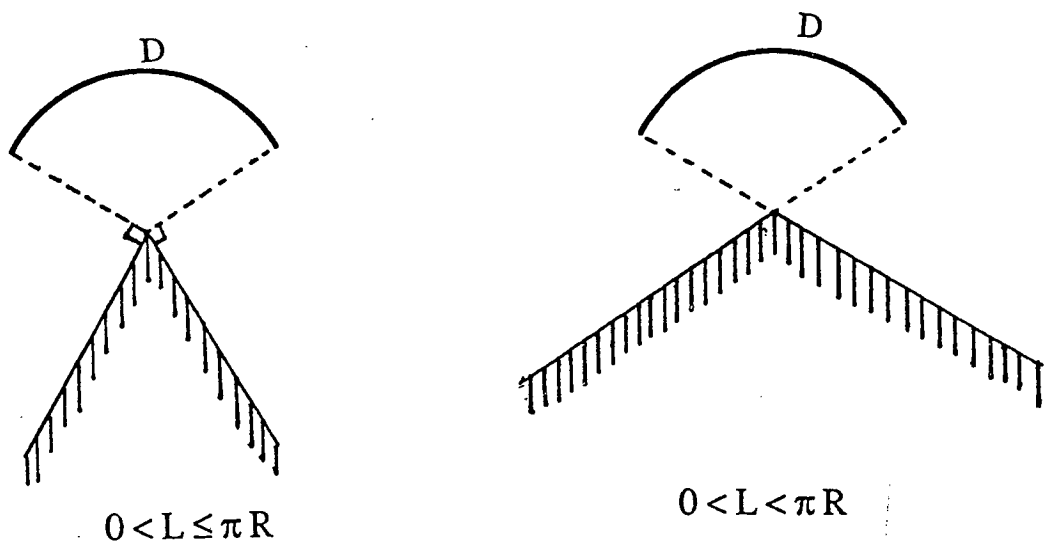
$$(3.23) \quad D \text{ is strictly quasiconvex iff } \frac{L}{R} < \pi$$

We illustrated on figure 3.2 for both cases the largest corresponding open neighborhood \mathcal{U} whose existence is asserted by proposition 3.1 and 3.8.

The first remark is that these neighborhoods are quite large, in particular they are "infinite" on the "convex side" of the arc of circle, and admit the center of the circle in their closure on the "concave side" of the arc of circle.

However, if the quasiconvex neighborhood (figure 3.2 left) catches all points admitting a unique projection on D when $\pi R \leq L < 2\pi R$, it misses many such points when $0 < L < \pi R$ (namely all the points of the dashed area which are not on the axis of symmetry of the figure), whereas the strictly quasiconvex neighborhood (figure 3.2 right) does a much better job, as it catches in all cases exactly all points z admitting a unique projection on D with no local minima on D of the distance to z .

However, if it is simple to figure out the shape of these neighborhoods for sets D as simple as an arc of circle, it becomes practically impossible for more complicated sets. Hence we need constructive sufficient conditions to recognize quasiconvexity and strict quasiconvexity, which will give us smaller cylindrical neighborhoods.



(\mathcal{U} = Complementary of dashed area)

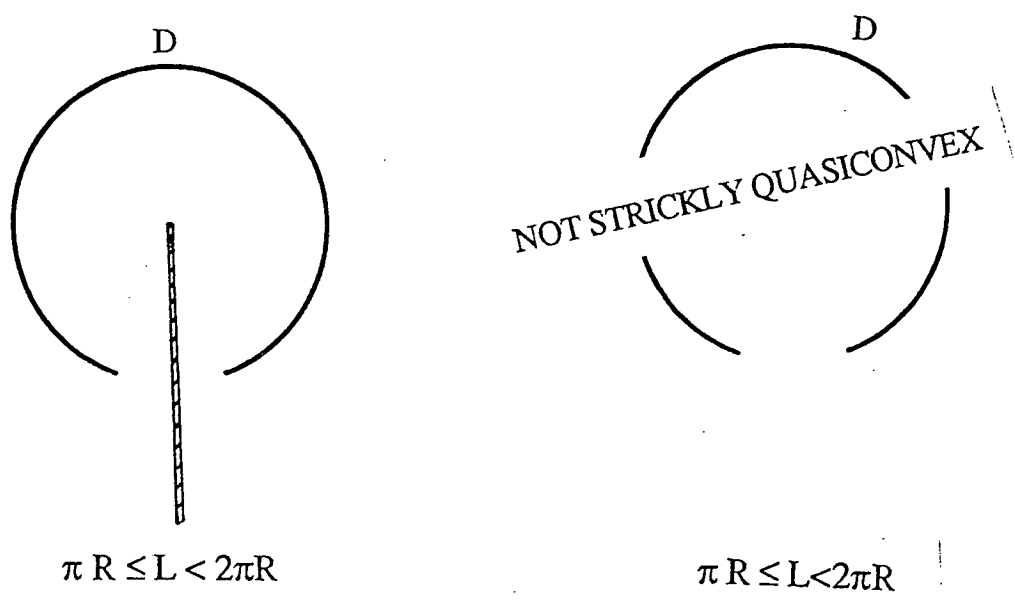


Figure 3.2. Largest open neighborhood \mathcal{U} for quasiconvex ($L < 2\pi R$, left) and strictly quasiconvex ($L < \pi R$, right) arcs of circle.

4 - THE γ -SIZE X CURVATURE CONDITION

We recall in this paragraph the sufficient condition for quasiconvexity introduced in [2],[3]. Given a set (D, \mathcal{P}) , we associate to any pseudosegment $P \in \mathcal{P}$ a number $\gamma(P)$ defined by :

$$(4.1) \quad \gamma(P) = \inf_{v \in [0, \delta(P)]} \{ \rho(v) - g(v) \}$$

where :

$$(4.2) \quad \rho(v) \text{ is the radius of curvature of } P \text{ at } P(v)$$

$$(4.3) \quad \begin{cases} g(v) \text{ is the solution of the elliptic problem} \\ -g''(v) = \rho^{-1}(v), \quad g(0) = g(\delta(P)) = 0 \end{cases}$$

This number $\gamma(P)$ measures in some sense the size \times curvature product of the path P : for a path P with given length, $\gamma(P) \rightarrow \infty$ when the radii of curvature along P go to infinity, and $\gamma(P) \rightarrow -\infty$ when the same radii of curvature tend to zero. Of course one has, for any $P \in \mathcal{P}$:

$$(4.4) \quad \gamma(P) \leq R(P)$$

and both $\gamma(P)$ and $R(P)$ are monotonous with respect to path inclusion.

We associate then to the set D the two numbers :

$$(4.5) \quad \gamma(D) = \inf_{P \in \mathcal{P}_M} \gamma(P)$$

$$(4.6) \quad R(D) = \inf_{P \in \mathcal{P}_M} R(P)$$

so that one always has :

$$(4.7) \quad \gamma(D) \leq R(D)$$

Definition 4.1 (γ -size \times curvature condition)

The set (D, \mathcal{P}) is said to satisfy the γ -size \times curvature condition iff :

$$(4.8) \quad \gamma(D) > 0$$

The main interest of this definition is :

Theorem 4.1

Let (D, \mathcal{P}) be given :

If :

$$\gamma(D) > 0 \quad (\gamma\text{-size } \times \text{ curvature condition})$$

Then

D is quasiconvex, with an open cylindrical neighborhood \mathcal{U} given by :

$$(4.9) \quad \mathcal{U} = \{ z \in F \mid d(z, D) < \gamma(D) \},$$

and an $\varepsilon(z)$ function given for any $z \in \mathcal{U}$, by

$$(4.10) \quad \varepsilon(z) = \gamma(D) - d(z,D) > 0$$

Hence theorems 3.4 thru 3.6 apply, and the projection on D , when it exists, satisfy the Lipschitz condition (compare with (3.10)) :

$$(4.11) \quad \|X_0 - X_1\|_F \leq \delta(X_0, X_1) \leq \left(1 - \frac{d}{R(D)}\right)^{-1} \|z_0 - z_1\|$$

as soon as z_0, z_1 are such that there exists d satisfying :

$$(4.12) \quad \|z_0 - z_1\|_F + \text{Max}_{j=0,1} d(z_j, D) \leq d < \gamma(D).$$

This sufficient condition is constructive, in the sense that one "only" has to test the sign of a number $\gamma(D)$ which can be estimated either analytically, or maybe also numerically-though the size of the computations becomes quickly unaffordable when the dimension of D increases.

We have only given in (4.1)-(4.3) the definition of $\gamma(P)$ when P is parametrized by its arc length. We refer to [2] [3] for the case of other parametrization of the pathes, which shall yield different numbers $\gamma(P)$. Hence the γ -size \times curvature (4.8) depends not only of the shape of the pathes on D , but also on their parametrization.

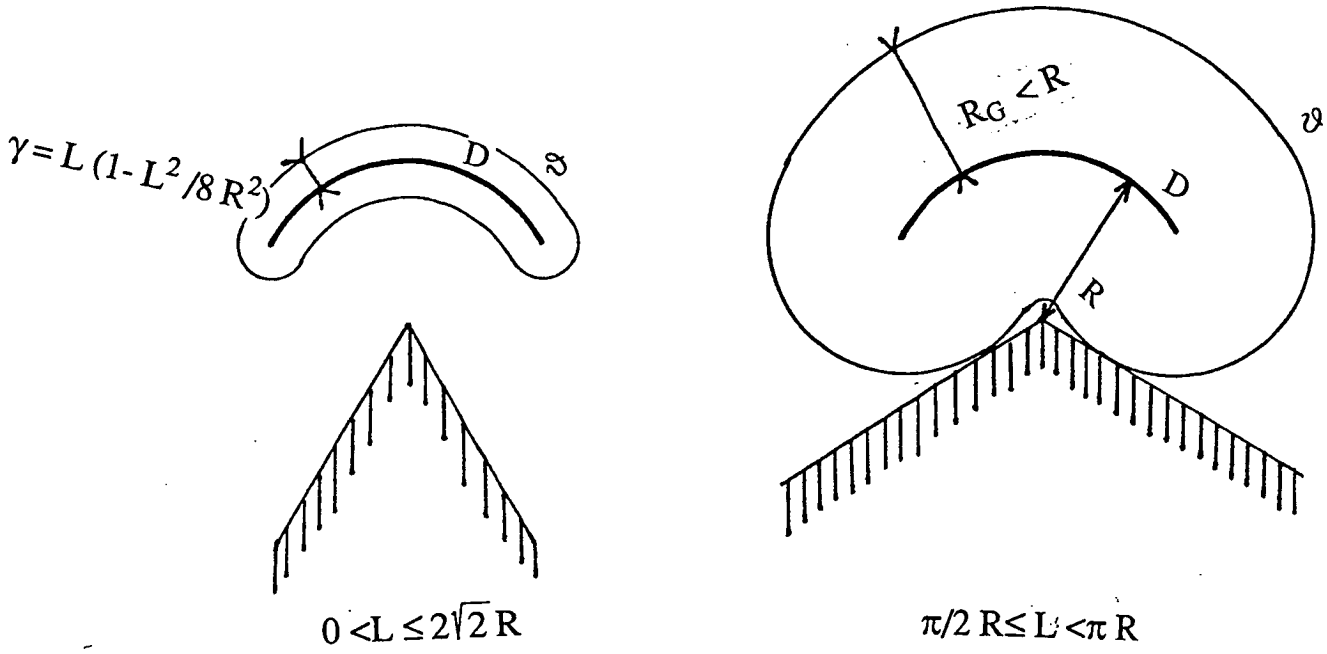
We conclude this paragraph by checking how well the γ -size \times curvature (4.8) recognizes quasiconvex arc of circles D of radius R and length L . A simple calculation shows that :

$$(4.13) \quad \gamma(D) = L \left(1 - \frac{L^2}{8R^2}\right)$$

which is positive as long as :

$$(4.14) \quad 0 < L < 2\sqrt{2}R$$

whereas the arc of circle is quasiconvex as long as $0 < L < 2\pi R$. Hence we see that the size \times curvature condition is quite far from recognizing all quasiconvex sets, even in the simple case of the arc of circle. We have depicted on figure 4.1, left, the size of the neighbourhood \mathcal{U} for $L = 2\pi/3 R < 2\sqrt{2} R$, which is quite small in comparison to the largest quasiconvex neighborhood represented by the complementary of the dashed area.



(complementary of dashed area represent neighborhoods of figure 3.2)

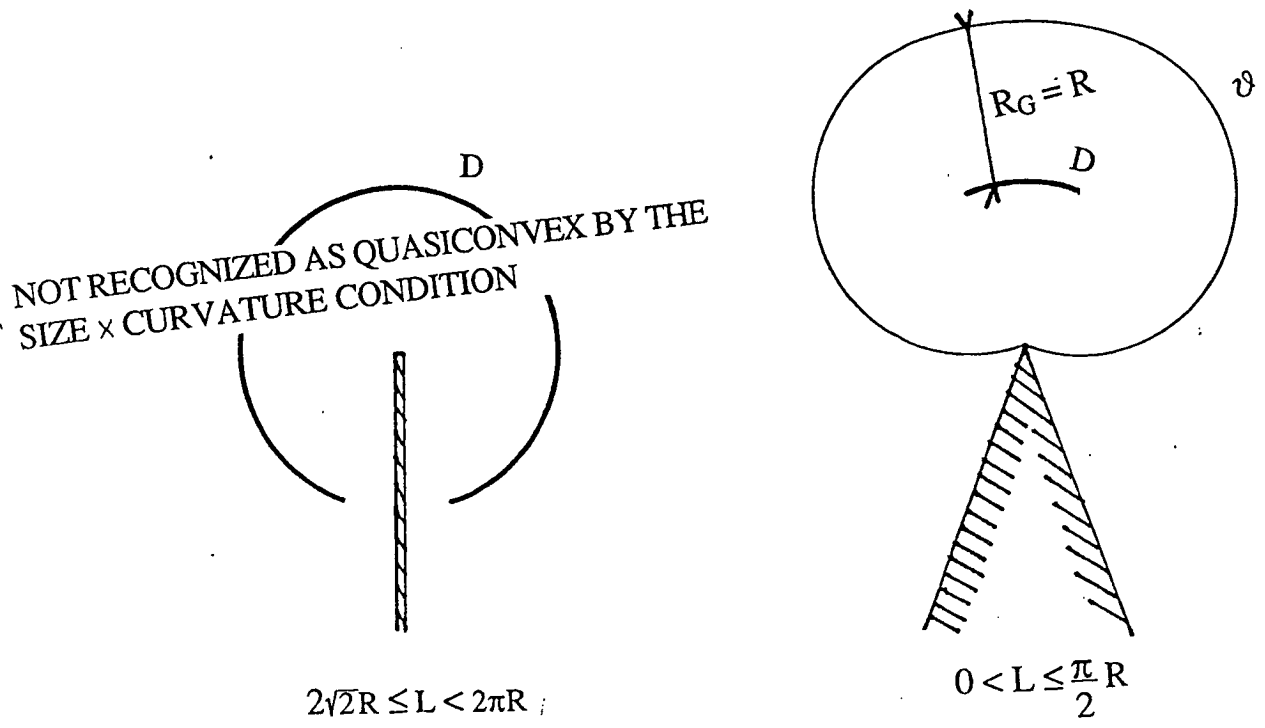


Figure 4.1. Cylindrical neighborhoods \mathcal{U} given by the γ -size \times curvature condition ($L < 2\sqrt{2}R$, left) and by the R_G size \times curvature condition ($L < \pi$, right).

5 - THE R_G SIZE \times CURVATURE CONDITION : A NEW SUFFICIENT CONDITION FOR STRICT QUASICONVEXITY OF SETS

We first introduce some new attributes of a pseudosegment :

Definition 5.1 (affine normal subspace - see figure 5.1)

Let a pseudosegment P be given. Then, for any $v, v' \in [0, \delta(P)]$, $v \neq v'$, we define the affine normal subspace $N(v, v')$ to P at v seen from v' by :

$$(5.1) \quad N(v, v') = \{z \in F \mid \langle z - P(v), \lambda v(v) \rangle_F \leq 0 \quad \forall \lambda \in \mathbb{R}, v + \lambda \in [\text{Min}(v, v'), \text{Max}(v, v')]\}$$

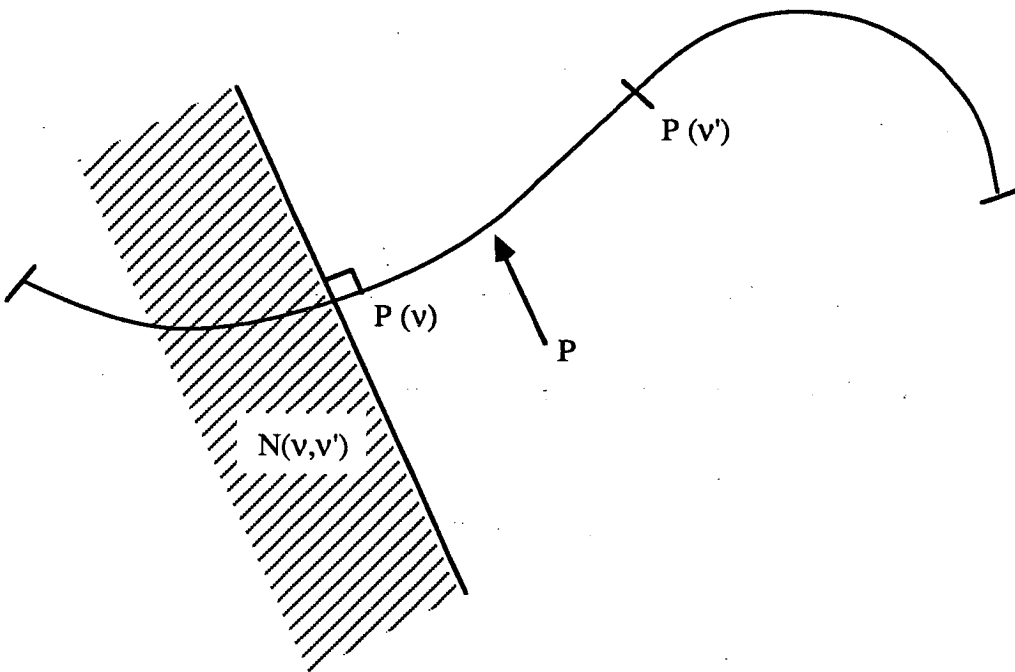


Figure 5.1. The normal subspace at v to a pseudosegment P seen from v' .

Proposition 5.2

Let a pseudosegment P be given. Then, for any $v, v' \in [0, \delta(P)]$, $v \neq v'$, one has :

$$P(v) + a(v) \in N(v, v')$$

Definition 5.3 (Global Radius of curvature)

Let a pseudosegment P be given. Then, for any $v, v' \in [0, \delta(P)]$, $v \neq v'$, we define the global radius of curvature of P at v seen from v' by :

$$(5.2) \quad \rho_G(v, v') = d(P(v), N(v, v') \cap N(v', v)) \in [0, +\infty]$$

with the natural convention that $\rho_G(v, v') = +\infty$ if $N(v, v') \cap N(v', v) = \emptyset$.

Notice that $\rho_G(v, v') \neq \rho_G(v', v)$ in general. This global radius of curvature can be easily calculated:

Proposition 5.4

Let the pseudo segment $P \in \mathcal{P}$ and $v, v' \in [0, \delta(P)]$, $v \neq v'$, be given, and denote :

$$(5.3) \quad \begin{aligned} X &= P(v) & , & & X' &= P(v') \\ v &= v(v) & , & & v' &= v(v') \end{aligned}$$

Then $\rho_G(v, v')$ is given by :

$$(5.4) \quad \rho_G(v, v') = \begin{cases} 0 & \text{if } \text{Sgn}(v'-v) \langle X'-X, v' \rangle \leq 0 \\ \frac{\text{Sgn}(v'-v) \langle X'-X, v' \rangle}{\sqrt{1 - \langle v, v' \rangle^2}} & \text{if } \text{Sgn}(v'-v) \langle X'-X, v' \rangle > 0 \text{ and } \langle v, v' \rangle \geq 0 \\ \text{Sgn}(v'-v) \langle X'-X, v' \rangle & \text{if } \text{Sgn}(v'-v) \langle X'-X, v' \rangle > 0 \text{ and } \langle v, v' \rangle \leq 0 \end{cases}$$

The proof of those formula is elementary, the basis ingredient being the projection of a point on an hyperplane. We have illustrated on figure 5.2 a few situations for the case of plane curve.

The relation with the usual (local) radius of curvature $\rho(v)$ is given by the :

Proposition 5.5

Let a pseudosegment P be given. Then, for any $v \in [0, \delta(P)]$, there exists an open neighborhood I of v in $[0, \delta(P)]$ such that, for any $v' \in I$:

$$(5.5) \quad \rho_G(v, v') = \frac{\text{Sgn}(v'-v) \langle X'-X, v' \rangle}{\sqrt{1 - \langle v, v' \rangle^2}}$$

$$(5.6) \quad \rho_G(v', v) = \frac{\text{Sgn}(v'-v) \langle X'-X, v \rangle}{\sqrt{1 - \langle v, v' \rangle^2}}$$

and

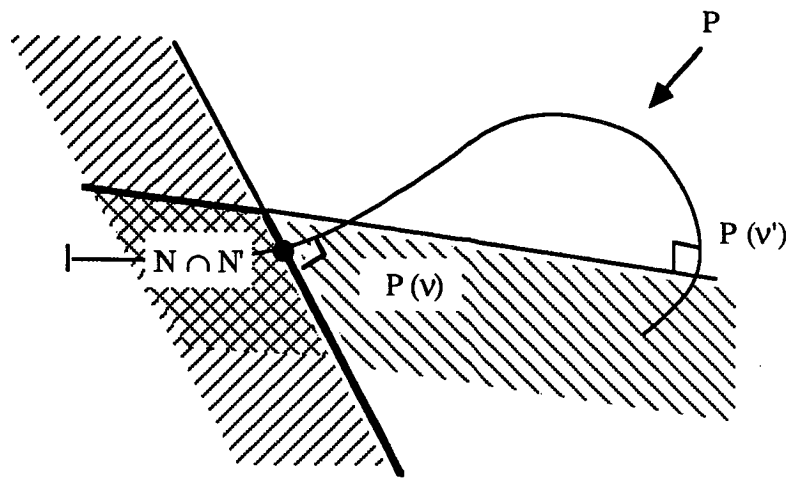
$$(5.7) \quad \rho_G(v, v') \rightarrow \rho(v)$$

$$(5.8) \quad \rho_G(v', v) \rightarrow \rho(v)$$

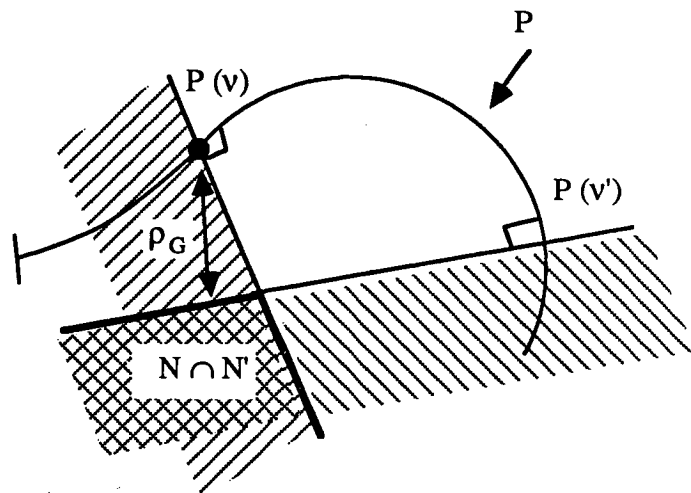
when $v' \rightarrow v$ in I .

Proof

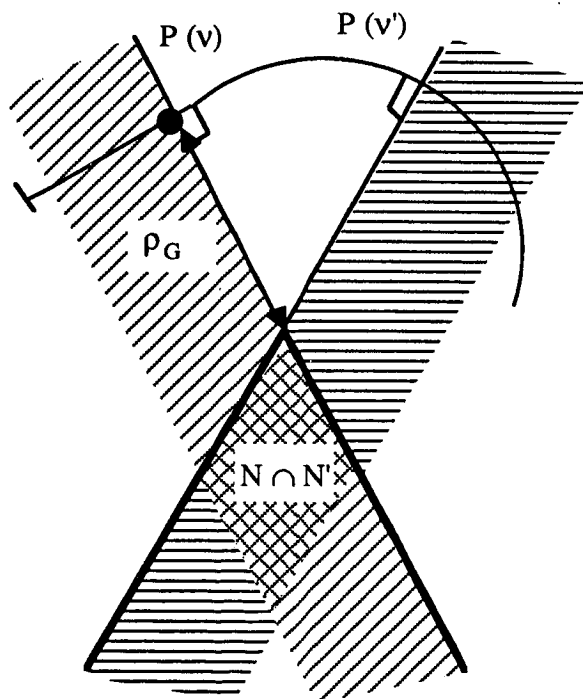
Formula (5.5) (5.6) results simply from proposition 5.4 and the \mathcal{C}^2 continuity of the path P .



$$\rho_G(v, v') = 0$$



$$\rho_G(v, v') = \text{Sgn}(v' - v) \langle X' - X, v' \rangle$$



$$\rho_G(v, v') = \frac{\text{Sgn}(v' - v) \langle X' - X, v' \rangle}{\sqrt{1 - \langle v, v' \rangle^2}}$$

Figure 5.2. Examples of global radii of curvature for a path P .

We turn now to the proof of (5.7). Let $dv = v' - v$. We rewrite first the numerator of $\rho_G(v, v')$ as :

$$(5.9) \quad \text{Sgn}(v' - v) \langle X' - X, v' \rangle = \text{Sgn}(v' - v) \langle v_\theta, v \rangle dv = \langle v_\theta, v \rangle |dv|$$

where

$$(5.10) \quad \begin{aligned} v_\theta &= v(v + \theta dv) & 0 \leq \theta \leq 1 \\ \langle v_\theta, v \rangle &\rightarrow 1 & \text{when } dv \rightarrow 0 \end{aligned}$$

Using then the theorem of mediane for v and v' , we obtain :

$$\langle v, v' \rangle = 1 - \frac{\delta^2}{2}$$

where :

$$(5.11) \quad \delta = \|v' - v\|$$

Hence the denominator of $\rho_G(v, v')$ rewrites :

$$(5.12) \quad \sqrt{1 - \langle v, v' \rangle^2} = \delta \left(1 - \frac{\delta^2}{4}\right)^{\frac{1}{2}}$$

But :

$$\delta = \left\| \int_0^{dv} a(v+\epsilon) d\epsilon \right\| = \|a(v)dv + \int_0^{dv} (a(v+\epsilon) - a(v)) d\epsilon\|$$

so that :

$$(\|a(v)\| - \eta(dv)) |dv| \leq \delta \leq (\|a(v)\| + \eta(dv)) |dv|$$

where one has set :

$$\eta(dv) = \sup_{\epsilon \in [0, dv]} \|a(v+\epsilon) - a(v)\|$$

which, as P is \mathcal{C}^2 , satisfies :

$$\eta(dv) \rightarrow 0 \quad \text{when } dv \rightarrow 0.$$

Hence :

$$(5.13) \quad \frac{\delta}{|dv|} \rightarrow \|a(v)\| \quad \text{when } dv \rightarrow 0$$

and $\rho_G(v, v')$ becomes, using (5.9) and (5.12) :

$$(5.14) \quad \rho_G(v, v') = \frac{|dv| \langle v_\theta, v \rangle}{\delta \left(1 - \frac{\delta^2}{4}\right)^{\frac{1}{2}}}$$

and from (5.10), (5.13) we see that :

$$\rho_G(v, v') \rightarrow \frac{1}{\|a(v)\|}$$

which is the sought result (5.7).

Then (5.8) follows immediately, as :

$$(5.15) \quad \rho_G(v, v') - \beta(dv) \leq \rho_G(v', v) \leq \rho_G(v, v') + \beta(dv)$$

with :

$$\beta(dv) = \frac{|\langle X'-X, v-v' \rangle|}{\sqrt{1 - \langle v, v' \rangle^2}} = \frac{|dv| |\langle v_\theta, v-v' \rangle|}{\sqrt{1 - \langle v, v' \rangle^2}}$$

where $0 < \theta' < 1$. Using (5.11), (5.12) we obtain :

$$(5.16) \quad \beta(dv) \leq \frac{|dv|}{\left(1 + \frac{\delta^2}{4}\right)^{\frac{1}{2}}} \rightarrow 0 \quad \text{when } dv \rightarrow 0.$$

This ends the proof of Proposition 5.5.

We are now naturally led to the :

Definition 5.6 (smallest global radius of curvature)

Let a pseudosegment P be given. Then we define its smallest global radius of curvature by :

$$(5.17) \quad R_G(P) = \inf_{v, v' \in [0, \delta(P)]} \rho_G(v, v')$$

Of course, $R_G(P)$ is easily related to $R(P)$ using proposition 5.5 :

Proposition 5.7

Let a pseudosegment P be given. Then :

$$(5.18) \quad 0 \leq R_G(P) \leq R(P) \leq +\infty$$

Remark 5.8

Unlike the number $\gamma(P)$ used in the size \times curvature condition in [2], [3], the number $R_G(P)$ depends only on the shape of the path P , and by no way of its parametrization (that is why we use only in this paper the parametrization by the arc length, which is the most convenient for the formulae). ■

The main interest of the smallest global radius of curvature comes from the :

Proposition 5.9

Let a pseudosegment $P \in \mathcal{P}$ and $z \in F$ be given.

If :

$$R_G(P) > 0$$

$$d(z, P) < R_G(P)$$

Then :

the $v \rightarrow d(v) = \|P(v) - z\|_F$ function is strictly quasiconvex.

Proof

Let $f(v) = d(v)^2$, and $v_0 \in [0, \delta(P)]$ be the value for which d and f attain their minimum on $[0, \delta(P)]$:

$$f(v_0) \leq f(v) \quad \forall v \in [0, \delta(P)]$$

which implies :

$$f'(v_0)\lambda \geq 0 \quad \forall \lambda \in \mathbb{R}, \quad v_0 + \lambda \in [0, \delta(P)]$$

i.e.

$$\langle z - P(v_0), \lambda v(v_0) \rangle \leq 0 \quad \forall \lambda \in \mathbb{R}, \quad v_0 + \lambda \in [0, \delta(P)]$$

which shows that :

$$z \in N(v_0, v') \quad \text{for any } v' \in [0, \delta(P)], v' \neq v_0.$$

Suppose now that $v \rightarrow d(v)$ is not strictly quasiconvex.

Then necessarily $d(v)$ has at least one local maximum for some value v_1 of the open interval $]0, \delta(P)[$. Of course $v \rightarrow f(v)$ has also a local maximum at the same $v_1 \in]0, \delta(P)[$, which implies that :

$$f'(v_1) = 0$$

i.e.

$$\langle z - P(v_1), v(v_1) \rangle = 0.$$

But, the preceding equation rewrites :

$$z \in N(v_1, v') \quad \text{for any } v' \in [0, \delta(P)], v' \neq v_1.$$

Hence we see that $z \in N(v_0, v_1) \cap N(v_1, v_0)$, which shows that :

$$\rho_G(v_0, v_1) \leq \|P(v_0) - z\|_F = d(z, P)$$

and, using the definition (5.6) :

$$R_G(P) \leq \rho_G(v_0, v_1) \leq d(z, P),$$

which contradicts the hypothesis $d(z, P) < R_G(P)$. This proves the strict quasiconvexity of d . ■

Definition 5.10 (smallest global radius of curvature of a set)

We associate to a set (D, \mathcal{P}) its smallest global radius of curvature defined by :

$$(5.19) \quad R_G(D) = \inf_{P \in \mathcal{P}_M} R_G(P)$$

which, using (4.7) and (5.18), satisfies :

$$(5.20) \quad 0 \leq R_G(D) \leq R(D).$$

This allows us to define a new "size \times curvature" condition for the set D :

Definition 5.11 (R_G - size \times curvature condition)

The set (D, \mathcal{P}) is said to satisfy a R_G - size \times curvature condition if and only if :

$$(5.21) \quad R_G(D) > 0$$

We give now the first main result of this paper :

Theorem 5.12

Let (D, \mathcal{P}) be given.

If :

$$(5.22) \quad R_G(D) > 0 \quad (R_G\text{- size} \times \text{curvature condition})$$

Then :

D is strictly quasiconvex, with a cylindrical neighborhood \mathcal{U} given by :

$$(5.23) \quad \mathcal{U} = \{z \in F \mid d(z, D) < R_G(D)\},$$

and an $\varepsilon(z)$ function defined, for any $z \in \mathcal{U}$, by :

$$(5.24) \quad \varepsilon(z) = R_G(D) - d(z, D) > 0.$$

Hence theorems 3.9 (the projection is unique, no local minima), 3.5 (continuity of the projection) apply, the lipschitz constant of the projection being given by (compare with (3.10) and (4.11)) :

$$(5.25) \quad \|X_0 - X_1\|_F \leq \delta(X_0 - X_1) \leq \left(1 - \frac{d}{R(D)}\right)^{-1} \|z_0 - z_1\|_F$$

as soon as z_0, z_1 are closed enough so that there exists d satisfying :

$$(5.26) \quad \|z_0 - z_1\|_F + \text{Max}_{j=0,1} d(z_j, D) \leq d < R_G(D)$$

If moreover

$$(5.27) \quad D \text{ is closed in } F$$

Then theorem 5.6 (existence of the projection) applies.

Proof

We check first that \mathcal{U} and ε defined by (5.23) and (5.24) satisfy the definition 3.1 of quasiconvex sets. Let z, η and P be given such that :

$$d(z, D) < R_G(D) \quad (\text{i.e. } z \in \mathcal{U})$$

(5.28)

$$0 < \eta < \varepsilon(z) = R_G(D) - d(z, D)$$

(5.29)

$$\|P(j) - z\| \leq d(z, D) + \eta \quad j = 0, \delta(P)$$

Hence $\|P(j) - z\| < R_G(D)$, $j = 0, j(P)$ which implies that $d(z, P) < R_G(D) \leq R_G(P)$. Then by proposition 5.9 we see that the :

$$v \rightarrow d(v) = \|P(v) - z\|_F$$

is strictly quasiconvex, and, using (3.2), we find that :

$$k(z, P; v) \leq \frac{\text{Max}_{v \in \{0, \delta(P)\}} d(v)}{\text{Min}_{v \in [0, \delta(P)]} \rho(v)}$$

and, using (5.29) and (2.8), (4.6) :

$$k(z, P; v) \leq \frac{d(z, D) + \eta}{R(D)}$$

which proves that (cf.3.4) :

$$(5.30) \quad k(z,\eta) \leq \frac{d(z,D)+\eta}{R(D)} < \frac{d(z,D)+\varepsilon(z)}{R(D)} = \frac{R_G(D)}{R(D)} < 1$$

which proves that D is quasiconvex. The strict quasiconvexity of D results then directly from proposition 5.9, and the formula (5.25), (5.26) follow immediately from (3.10), (3.12 thru 14) and (5.30). ■

In order to illustrate this new sufficient condition, we check how well it recognizes quasiconvex and strictly quasiconvex arc of circles of radius R and length L. A simple calculation yields :

$$(5.31) \quad R_G(D) = \begin{cases} R & 0 < \frac{L}{R} < \frac{\pi}{2} \\ R \sin \frac{L}{R} & \frac{\pi}{2} \leq \frac{L}{R} \leq \pi \\ 0 & \pi \leq \frac{L}{R} < 2\pi \end{cases}$$

which is strictly positive as long as :

$$(5.32) \quad 0 < \frac{L}{R} < \pi$$

Hence the R_G - size \times curvature condition recognizes exactly all strictly quasiconvex arc of circles ! We have illustrated in figure 4.1, right, the neighborhoods \mathcal{U} generated by formula (5.31), which turn out to be the largest cylindrical neighborhoods included in the largest strictly quasiconvex neighborhood (depicted as the complementary of the dashed area).

If we compare now the R_G - size \times curvature condition, driven by $R_G(D)$ (right of figure 4.1), to the previous size \times curvature condition driven by $\gamma(D)$ (left of figure 4.1) on arc of circles, we see that the new condition is much more precise than the old one, as it gives you more (strict quasi-convexity) for less ($L/R < \pi$ instead of $L/R < 2\sqrt{2}$) !

In the general case, the global size \times curvature condition cannot be expected to be optimal (i.e. necessary for strict quasiconvexity), as nothing prevents you to equip a convex set with pathes \mathcal{P} having small radii of curvature ! However it is reasonable to conjecture that the R_G -size \times curvature condition becomes necessary in the case where the collection of pathes \mathcal{P} is made of minimum length pathes in D. This matter will be discussed in a forecoming paper.

6 - OBTAINING LOWER BOUNDS ON THE GLOBAL RADIUS OF CURVATURE

Given a path $P \in \mathcal{P}$, we shall try in this paragraph to obtain lower bounds on its global radius of curvature $R_G(P)$.

In particular, one would like to substantiate the intuitive feeling that "arc of circles are the worst paths", i.e. that for any path P one should have $R_G(P) \geq R(P)$ as soon as $\delta(P) \leq \pi/2 R(P)$ and $R_G(P) > 0$ as soon as $\delta(P) < \pi R(P)$, as suggested by figure 4.1 (right) for the case of an arc of circle. In order to do that, we need to recall one (very natural) attribute of the path :

Definition 6.1 (deflection)

Given a path $P \in \mathcal{P}$, and $v, v' \in [0, \delta(P)]$, the deflection of P between v and v' is :

$$(6.1) \quad \theta(v, v') = \text{Arg cos} \langle v(v), v(v') \rangle$$

and the largest deflection of P is :

$$(6.2) \quad \Theta(P) = \text{Max } \theta(v, v') \\ v, v' \in [0, \delta(P)].$$

Of course, $\theta(v, v')$ satisfies :

$$(6.3) \quad \begin{aligned} \theta(v, v') &\in [0, \pi] & \forall v, v' \in [0, \delta(P)] \\ \theta(v, v) &= 0 & \forall v \in [0, \delta(P)] \\ \theta(v, v') &= \theta(v', v) & \forall v, v' \in [0, \delta(P)] \end{aligned}$$

and is related to the radius of curvature $\rho(v)$ at v by the

Lemma 6.2 : for any $v, v' \in [0, \delta(P)]$ one has

$$(6.4) \quad \left| \frac{\partial \theta}{\partial v}(v, v') \right| \leq \frac{1}{\rho(v)}$$

Proof : let $dv \in \mathbb{R}$ be chosen such that $v + dv \in [0, \delta(P)]$, and let

$$(6.5) \quad d\alpha = \text{Arg cos} \langle v(v), v(v + dv) \rangle$$

Using the triangular inequality for the curvilinear triangle on the unit sphere having $v(v)$, $v(v+dv)$ and v' as vertices yields :

$$(6.6) \quad \left| \theta(v+dv, v') - \theta(v, v') \right| \leq \left| d\alpha \right|$$

But by definition $d\alpha \in [0, \pi]$ and hence is positive ; so we see from (6.5) that $\left| d\alpha \right| = d\alpha = \text{Arg cos} (1 - \langle v(v), v(v) - v(v + dv) \rangle)$ and, as the $\epsilon \rightarrow \text{Arg cos} (1 - \epsilon)$ function is increasing :

$$(6.7) \quad \left| d\alpha \right| \leq \left| \langle v(v), v(v) - v(v + dv) \rangle \right| \leq \|v(v + dv) - v(v)\|$$

But we have seen with the proof of proposition 5.5 that

$$(6.8) \quad \delta = \|v(v + dv) - v(v)\| \leq (\|a(v)\| + \eta(dv)) \left| dv \right|$$

where $\eta(dv) \rightarrow 0$ when $dv \rightarrow 0$. Putting together (6.6) (6.7) (6.8) yields :

$$(6.9) \quad \left| \theta(v+dv, v') - \theta(v, v') \right| \leq (1/\rho(v) + \eta(dv)) \left| dv \right|$$

which proves (6.4) when $dv \rightarrow 0$.

We have drawn on figure 6.1, top, the function $v \rightarrow \theta(v, v')$ in the case wher P is arc of circle of length L and radius R, and on the bottom of the same figure the function $v \rightarrow \theta(v, v')$ for a general path P with smallest radius of curvative $R(P)$ and length $\delta(P)$: we see that in the general case the graph of $v \rightarrow \theta(v, v')$ may be quite complicated, as it may contains :

- points as $v, v_1, v_2 \in] 0, \delta(P) [$ which yield stationary points of the deflection θ , when the acceleration $a(v)$ happens to be orthogonal to the reference velocity $v(v')$
- points like v'' which correspond to values of v for which the velocity $v(v)$ happens to be parallele to the reference velocity $v(v')$.
- intervals like I where the deflection is constant, which correspond either to straight segments or to helices with axis $v(v')$
- points like v where $\theta(v, v')$ achieves its global maximum over $[0, \delta(P)]$. On the figure, $v \in] 0, \delta(P) [$ but it could as well be at $v = 0$ or $v = \delta(P)$.

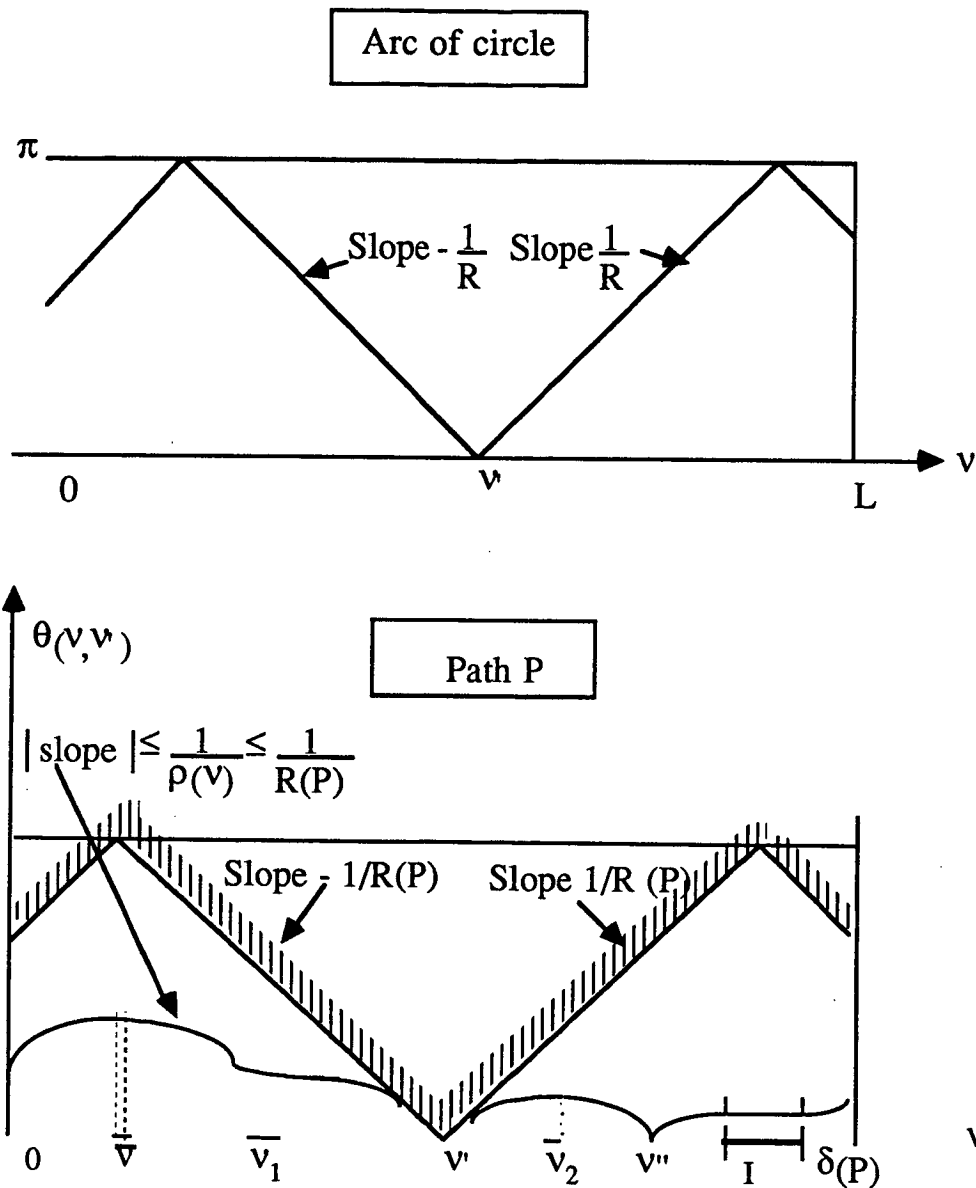


Figure 6.1 : The deflection function $\theta(v, v')$ for the case of on arc of circle (top) and in general (bottom).

So far for the properties of the deflection function. We come now to our purpose, namely, obtaining a lower bound on $R_G(P)$, which reduces to finding a lower bound on $\rho_G(v, v')$, independant of v and v' . Looking at the formula (5.4) giving $\rho_G(v, v')$, we see that the two pieces entering in this formula are :

$$(6.11) \quad \text{Sgn}(v' - v) \langle X' - X, v' \rangle \quad \text{and} \quad \sqrt{1 - \langle v, v' \rangle} > 2$$

which are related to the deflection θ by

$$(6.12) \quad \text{Sgn}(v' - v) \langle X' - X, v' \rangle = \int_{\text{MIN}(v, v')}^{\text{MAX}(v, v')} \cos \theta(t, v') dt$$

and :

$$(6.13) \quad \sqrt{1 - \langle v, v' \rangle} > 2 = \sin \theta(v, v').$$

We concentrate now on obtaining a lower bound for $\text{Sgn}(v', v) \langle X' - X, v' \rangle$. Our basic tool for that is the following lemma :

Lemma 6.3 : Let $v, v' \in [0, \delta(P)]$ be given. Then the following inequalities hold :

$$(6.14) \quad \text{Var } \theta(., v') \leq \int_v^{v'} \frac{dv}{\rho(v)}$$

$$(6.15) \quad \text{Sgn}(v' - v) \langle X' - X, v' \rangle \geq |v' - v| \cos \theta + R. \text{Var} \{ \sin \theta(., v') - \theta(., v') \cos \theta \}$$

where $\text{Var } g$ represents the total variation of the function g over the $[\text{Min}(v, v'), \text{Max}(v, v')]$ interval defined by :

$$(6.16) \quad \text{Var } g = \text{Sup} \sum_{i=1}^N |g(t_i) - g(t_{i-1})|$$

$$N \in \mathbb{N}$$

$$\text{Min}(v, v') \leq t_0 \leq t_1 \leq t_N \leq \text{Max}(v, v')$$

and where

$$(6.17) \quad \bar{R} = \inf_{t \in [v, v']} \rho(t)$$

$$(6.18) \quad \bar{\theta} = \text{Sup}_{t \in [v, v']} \theta(t, v')$$

Proof : Let us consider the case where $v \leq v'$ (the proof is similar if $v \geq v'$). We remark first, using (6.12) and (6.18), that :

$$(6.19) \quad \langle X' - X, v' \rangle = \int_v^{v'} (\cos \theta(t, v') - \bar{\cos} \theta) dt + (v' - v) \bar{\cos} \theta.$$

But from lemma 6.2 we have

$$(6.20) \quad \rho(t) \left| \frac{d\theta}{dv}(t, v') \right| \leq 1 \text{ for almost all } t \text{ of } [v, v']$$

which shows that

$$\int_v^{v'} \frac{dv}{\rho(v)} \geq \int_v^{v'} \left| \frac{d\theta}{dv}(t,v') \right| dt$$

Plugging then (6.20) in (6.19) yields, as $\cos \theta(t,v') - \cos \bar{\theta} \geq 0 \forall t \in [v,v']$ and $\rho(t) \geq R$
 $\forall t \in [v,v']$:

$$\begin{aligned} \langle X' - X, v' \rangle &\geq \bar{R} \int_v^{v'} (\cos \theta(t,v') - \cos \bar{\theta}) \left| \frac{\partial \theta}{\partial v}(t,v') \right| dt + (v' - v) \cos \bar{\theta} \\ &= \bar{R} \int_v^{v'} \left| \frac{\partial}{\partial v} (\sin \theta(t,v') - \theta(t,v') \cos \bar{\theta}) \right| dt + (v' - v) \cos \bar{\theta} \end{aligned}$$

These formula prove (6.14) (6.15) as, for any function g such that $g' \in L^1([v,v'])$ one has:

$$\text{Var } g = \int_v^{v'} \left| \frac{\partial g}{\partial v}(t) \right| dt. \quad \blacksquare$$

The lower bounds (6.14) (6.15) take advantage, through the total variation of the $\theta(.,v')$ and $\sin(.,v') + \theta(.,v') \cos \bar{\theta}$ functions, of all information contained in the deflection function $\theta(.,v')$.

In particular, they increase when the variation of the deflection increases, which shows that "oscillating paths" have more chance to have positive global radii of curvature.

If one retains only the maximum deflection $\bar{\theta}$ on $[v,v']$ seen from v' one has the

Corollary 6.4 : Let $v, v' \in [0, \delta(P)]$ be given. Then

$$(6.22) \quad \bar{\theta} \leq \int_v^{v'} \frac{dv}{\rho(v)} \leq \frac{|v' - v|}{\bar{R}}$$

$$(6.23) \quad \text{Sgn}(v' - v) \langle X' - X, v' \rangle \geq \bar{R} \sin \bar{\theta} + (|v' - v| - \bar{R} \bar{\theta}) \cos \bar{\theta}$$

where $\bar{R}, \bar{\theta}$ are defined in (6.17) (6.18).

Proof : suppose $v \leq v'$ for instance, and define $v \in [v,v']$ by :

$$\theta(v, v') = \bar{\theta} = \text{Sup } \theta(t, v').$$

$$t \in [v, v']$$

Then (6.22) (6.23) result immediately of (6.14) (6.15) and (6.16) with $N = 2, t_0 = v, t_1 = v'$ ■

We can put together the definition (5.4) of $\rho_G(v, v')$ and corollary 6.4 to obtain the sought lower bound on $\rho_G(v, v')$:

Proposition 6.5 : Let $P \in \mathcal{P}, v, v' \in [0, \delta(P)], v \neq v'$, be given, and R, θ be defined by (6.17) (6.18).

i) If :

$$(6.24) \quad 0 \leq \bar{\theta} \leq \pi/2$$

Then :

$$(6.25) \rho_G(v, v') \geq \bar{R} + (|v'-v| - \bar{R} \cdot \bar{\theta}) \cotan \bar{\theta}$$

and the right hand side of (6.25) is strictly positive as soon as

$$(6.26) \bar{R} > 0 \quad (\text{independantly of } |v'-v| !)$$

ii) If

$$(6.27) \frac{\pi}{2} \leq \bar{\theta} \leq \pi$$

Then :

$$(6.28) \rho_G(v, v') \geq R \sin \theta + (|v'-v| - R \cdot \theta) \cos \theta$$

and the right hand side of (6.28) is strictly positive as soon as the size x curvature condition

$$(6.29) |v'-v| / \bar{R} < \bar{\theta} - \text{tg } \bar{\theta}$$

is satisfied (notice that, because of (6.22), condition (6.29) can be satisfied only of $\bar{\theta} < \pi!$).

Proof : We prove first i).

As $\bar{\theta} \leq \pi/2$ one has $\cos \bar{\theta} \geq 0$ which shows by corollary 6.4 that

$$(6.30) \text{Sgn}(v'-v) \langle X'-X, v' \rangle \geq 0,$$

and by definition of θ we have

$$(6.31) \theta(v, v') \leq \bar{\theta} \leq \pi/2$$

which shows that

$$(6.32) \cos \theta(v, v') = \langle v, v' \rangle \geq 0$$

and

$$(6.33) 0 \leq \sin \theta(v, v') \leq \sin \bar{\theta}.$$

Using (6.30) (6.32) in the definition (5.4) of ρ_G yields then :

$$(6.34) \rho_G(v, v') = \frac{\text{Sgn}(v'-v) \langle X'-X, v' \rangle}{\sin \theta(v, v')}$$

and using (6.33) :

$$(6.35) \rho_G(v, v') \geq \frac{\text{Sgn}(v'-v) \langle X'-X, v' \rangle}{\sin \bar{\theta}}$$

which yields the sought result (6.25) using the lower bound (6.23) for the numerator.

We prove now ii).

As $\pi/2 \leq \bar{\theta} < \pi$, one has no information on the sign of $\cos \theta(v, v') = \langle v, v' \rangle$. But we see in (5.4) that, whatever the signs of $\text{Sgn}(v'-v) \langle X'-X, v' \rangle$ and $\langle v, v' \rangle$ may be, one always has

$$(6.36) \rho_G(v, v') \geq \text{Sgn}(v'-v) \langle X'-X, v' \rangle,$$

the equality holding when $\text{Sgn}(v'-v) \langle X'-X, v' \rangle \geq 0$ and $\langle v, v' \rangle \leq 0$. Then the sought minoration (6.28) results immediately from (6.36) and the formula (6.23) of corollary 6.4, and the size x curvature condition (6.29) results immediately from (6.28) by noticing that $\cos \theta \leq 0$.

■

Formula (6.25) gives us also a more precise insight in the way $\rho_G(v, v')$ approaches $\rho(v)$ when $v' \rightarrow v$: if $t \rightarrow \rho(t)$ is for example an increasing function , then

$$(6.37) \quad \rho_G(v, v') \geq \rho_G(v) \quad \text{when } v' \geq v$$

$$\rho_G(v, v') \geq \rho_G(v') \quad \text{when } v' \leq v$$

as soon as v' and v are close enough so that the maximum deflection between v and v' is smaller than $\pi/2$.

We deduce now from proposition 6.5 a lower bound for $R_G(P) = \inf_{v, v' \in [0, \delta(P)]} \rho_G(v, v')$, which is our second main result :

Theorem 6.6 :

Let a path $P \in \mathcal{P}$ be given.

Then its maximum deflection $\Theta(P)$ (defined in (6.2)) is related to its length $\delta(P)$ and its smallest radius of curvature $R(P)$ by :

$$(6.39) \quad \Theta(P) \leq \int_0^{\delta(P)} \frac{dv}{\rho(v)} \leq \frac{\delta(P)}{R(P)}$$

and the following lower bounds for $R_G(P)$ defined in (5.17) hold :

i) Low deflection pathes : if

$$(6.40) \quad 0 \leq \Theta(P) \leq \pi/2$$

then

$$(6.41) \quad R_G(P) = R(P)$$

independantely of the length $\delta(P)$ of the path.

ii) Large deflection pathes : if

$$(6.42) \quad \pi/2 \leq \Theta(P) \leq \pi$$

then

$$(6.42) \quad R_G(P) \geq R(P) \sin \Theta(P) + \{\delta(P) - R(P) \Theta(P)\} \cos \Theta(P)$$

which is strictly positive as soon as the length $\delta(P)$ of the path satisfies the size x curvature condition :

$$(6.43) \quad \delta(P) / R(P) < \Theta(P) - \tan \Theta(P)$$

(once again, this condition can be satisfied only if $\Theta(P) < \pi$!).

Proof : Notice first that (6.39) results immediately from (6.22) of corollary (6.4) by taking $v = v_0, v' = v_1$ where $v_0, v_1 \in [0, \delta(P)]$ are such that $\theta(v_0, v_1) = \Theta(P)$, and by noticing that :

$$\int_{v_0}^{v_1} \frac{dv}{\rho(v)} \leq \int_0^{\delta(P)} \frac{dv}{\rho(v)} .$$

We prove now the results of i).

If $\Theta(P) \leq \pi/2$, one has obviously for any $v, v' \in [0, \delta(P)]$:

$$0 \leq \bar{\theta} \leq \Theta(P) \leq \pi/2$$

so that part i) of proposition 6.5 applies. Hence

$$\rho_G(v, v') \geq \bar{R}$$

which implies, as $\bar{R} \geq R(P)$:

$$R_G(P) \geq R(P)$$

But we know from proposition 5.7 that

$$R_G(P) \leq R(P)$$

which ends the proof of (6.41).

We prove now the results of ii).

As now we know only that $\pi/2 \leq \Theta(P) \leq \pi$, two cases may happen for given $v, v' \in [0, \delta(P)]$:

. either $0 \leq \bar{\theta} \leq \pi/2 \leq \Theta(P)$, and we find as above :

$$(6.44) \quad \rho_G(v, v') \geq R(P)$$

. or $\pi/2 \leq \bar{\theta} \leq \Theta(P) \leq \pi$ and then we find from part ii) of proposition 6.5 that

$$\rho_G(v, v') \geq \bar{R} \sin \bar{\theta} + (|v' - v| - \bar{R} \cdot \bar{\theta}) \cos \bar{\theta}$$

But as $\cos \bar{\theta} \leq 0$, this implies as $|v' - v| \leq \delta(P)$ and $\bar{R} \geq R(P)$ that

$$\rho_G(v, v') \geq R(P) \sin \bar{\theta} + (\delta(P) - R(P)\bar{\theta}) \cos \bar{\theta}.$$

But the mapping $\alpha \rightarrow R(P) \sin \alpha + (\delta(P) - R(P)\alpha) \cos \alpha$ is decreasing over the $[0, \Theta(P)]$ interval, as its derivative is the $\alpha \rightarrow -\sin \alpha (\delta(P) - R(P)\alpha)$ function, which is negative over the $[0, \Theta(P)]$ interval because of (6.39). Hence we obtain, as $\bar{\theta} \leq \Theta(P)$:

$$(6.45) \quad \rho_G(v, v') \geq R(P) \sin \Theta(P) + (\delta(P) - R(P)\Theta(P)) \cos \Theta(P)$$

When taking the minimum over all $v, v' \in [0, \delta(P)]$ we obtain (6.42) as the right-hand side of (6.45) is smaller than the right hand side of (6.44).

Finally, (6.43) is obtained immediately by writing that the right-hand side of (6.42) is strictly positive. ■

We illustrate first on figure 6.2, for a path P with given smallest radius of curvature $R(P)$, the largest (according to theorem 6.6) admissible length $\delta_{\text{Max}}(P)$ and the corresponding lower bound $R_{G \text{ Min}}(P)$ for $R_G(P)$, as functions of the maximum deflection $\Theta(P)$ of \mathcal{P} . Notice that $\delta_{\text{Max}}(P) > \pi R(P)$ for $\Theta(P) < \pi$, which proves the intuitive result announced at the beginning of this paragraph, namely that paths P such that $\delta(P)/R(P) < \pi$ satisfy $R_G(P) > 0$!

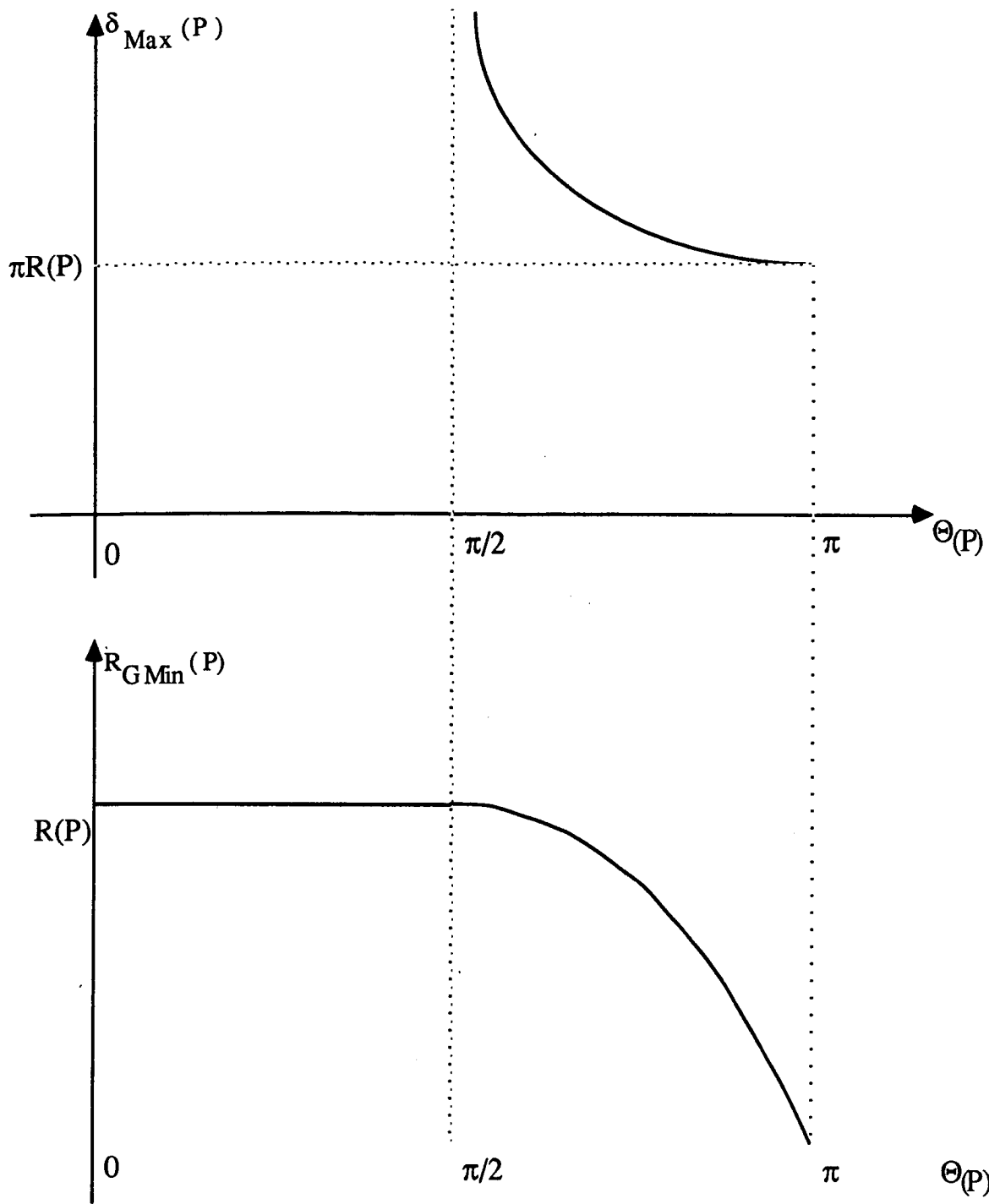


Figure 6.2 : The largest length $\delta_{\text{Max}}(P)$ recognized by theorem 6.6 and the corresponding lower bound $R_{\text{GMin}}(P)$ for a path P with given smallest radius of curvature $R(P)$, as a function of the maximum deflection $\Theta(P)$.

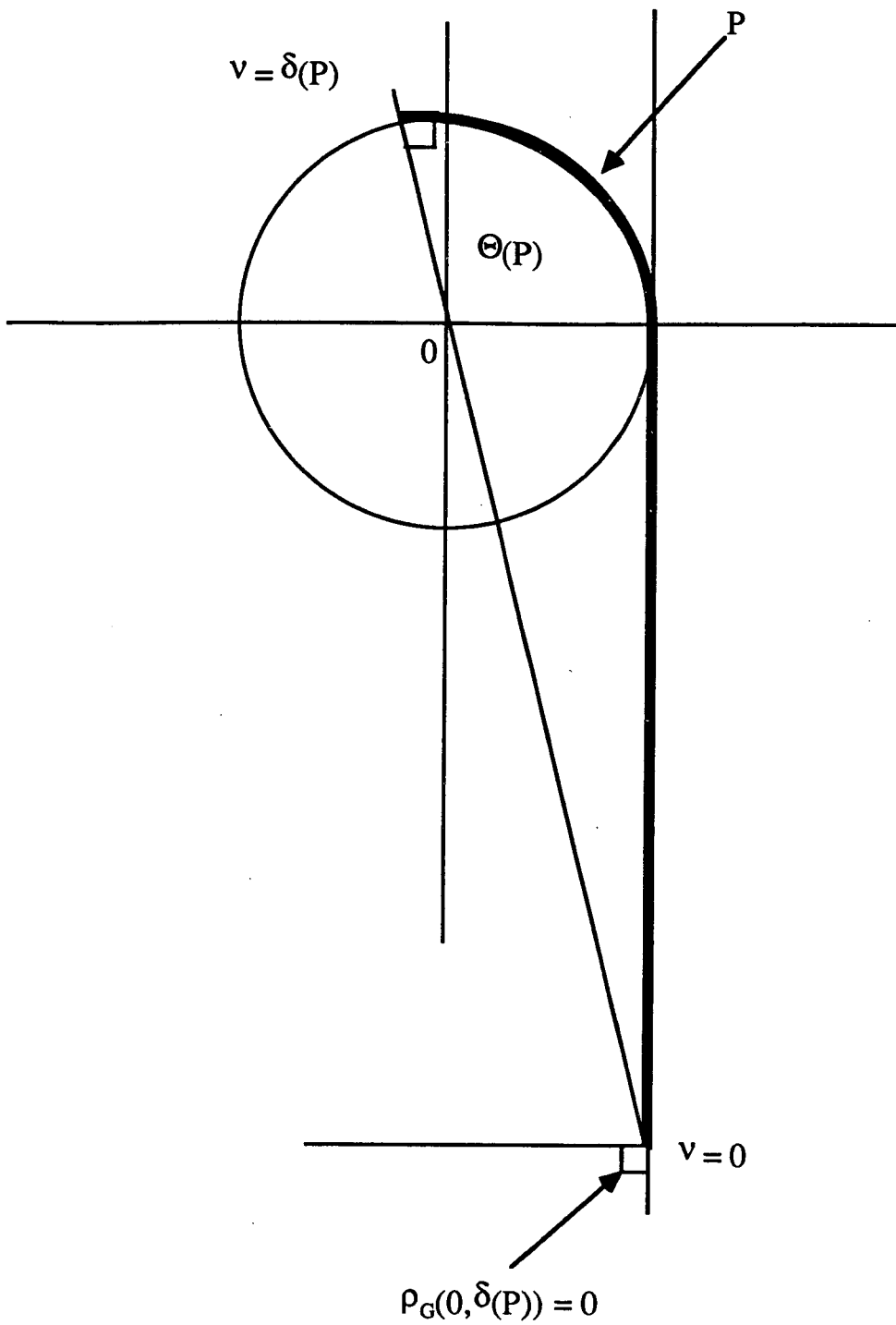


Figure 6.3 : Illustration of the sharpness of the upper bound to the length of paths with maximum deflection larger than $\pi/2$.

We notice also that the bounds given in theorem 6.6 are sharp : when P is an arc of circle, inequalities in (6.39) and (6.42) become equality, and the theorem recognizes exactly all strictly quasiconvex arcs of circles, as the equality in (6.39) implies that (6.43) is satisfied as soon as $\delta(P)/R(P) < \pi$. But the upper bound (6.43) for the admissible length $\delta(P)$ of a path with a maximum deflection larger than $\pi/2$ is also sharp, as may be seen from figure 6.3 : given a maximum deflection $\pi/2 \leq \Theta(P) < \pi$, the worst thing you can do (up to a change of scale of course) is

- first go straight ahead with a length $|\tan \Theta(P)|$
- then turn an angle $\Theta(P)$ using an arc of circle of radius one.

The resulting path has a length $\Theta(P) - \tan \Theta(P)$, and a zero $R_G(P)$ as obviously $\rho_G(0, \delta(P)) = 0!$

Of course, one may use theorem 6.6 to construct various sufficient conditions for the strict quasiconvexity of a set, as we know from theorem 5.12 and definition 5.10 that the only thing we have to do is to find a lower bound for $R_G(P)$, $P \in \mathcal{P}_M$.

We state here one such condition, which will use the following lower and upper bounds to attributes of the set (D, \mathcal{P}) :

Lower bound R to radii of curvature along paths :

$$(6.46) \quad R(P) \geq R \quad \forall P \in \mathcal{P}_M$$

Upper bound Δ to path length :

$$(6.47) \quad \delta(P) \leq \Delta \quad \forall P \in \mathcal{P}_M$$

Upper bound Θ to path maximal deflection :

$$(6.48) \quad \Theta(P) \leq \Theta \quad \forall P \in \mathcal{P}_M$$

(Notice that such an upper bound can be obtained by either formula

$$(6.40) \quad \Theta = \sup_{P \in \mathcal{P}_M} \int_0^{\delta(P)} \frac{dv}{\rho(v)} \quad \text{or} \quad \Theta = \sup_{P \in \mathcal{P}_M} \frac{\delta(P)}{R(P)}.$$

Theorem 6.7 : (Θ - size \times curvature condition)

Let (D, \mathcal{P}) be given such that :

$$(6.50) \quad R > 0$$

i) If

$$(6.51) \quad 0 \leq \Theta \leq \pi/2$$

then

$$(6.52) \quad R_G(D) \geq R > 0$$

and (D, \mathcal{P}) is strictly quasiconvex for a cylindrical neighborhood of size R at least.

ii) If

$$(6.53) \quad \pi/2 \leq \Theta < \pi$$

then

$$(6.54) \quad R_G(D) \geq R \sin \Theta + (\Delta - R\Theta) \cos \Theta$$

and (D, \mathcal{P}) is strictly quasiconvex, with a cylindrical neighborhood of the size of the r.h.s. of (6.54) at least, as soon as the maximum length of pathes satisfy

$$(6.55) \quad \Delta/R < \Theta - \tan \Theta$$

7. One application to non-linear least squares :

We consider the problem :

$$(7.1) \quad \text{find } x \in C \text{ such that } J(x) = \|\varphi(x) - z\|_F^2 = \min \text{ over } C$$

where

E is a Banach space, F is an Hilbert space

$$(7.2) \quad C \subset E \text{ is a closed convex set}$$

$\varphi : C \rightarrow F$ is a C^2 - mapping .

$z \in F$ is a given point (data)

As φ is of class C^2 there exists β such that :

$$(7.3) \quad \|\varphi''(x)(y, y)\|_F \leq \beta \|y\|_E^2 \quad \forall x \in C, \forall y \in E,$$

and we suppose that the derivative $\varphi'(x)$ is uniformly invertible, ie that there exists α_m and α_M such that

$$(7.4) \quad \alpha_m \|y\|_E \leq \|\varphi'(x) \cdot y\|_F \leq \alpha_M \|y\|_E \quad \forall x \in C, \forall y \in E.$$

We may then apply the results of paragraph 6 to the set $D = \varphi(C)$ equipped with the pathes

$$(7.5) \quad \mathcal{P} = \{\varphi([x, y]), x, y \in C\}$$

for which the subset

$$(7.6) \quad \mathcal{P}_M = \{\varphi([x, y]), x, y \in \partial C\}$$

is obviously a subfamily of maximal pathes.

One checks easily (cf [3]) that, given a path $P = \varphi([x, y]), x, y \in C$, one has

$$(7.7) \quad v(t) = \int_0^t \|\varphi'(x_t)(y-x)\|_F dt \geq \alpha \|y-x\|_E$$

$$(7.8) \quad \rho(v) \geq \frac{\|\varphi'(x_t)(y-x)\|_F^2}{\|\varphi''(x_t)(y-x, y-x)\|_F} \geq \frac{\alpha^2}{\beta}$$

where

$$(7.9) \quad x_t = (1-t)x + ty \quad \forall t \in [0, 1].$$

From (7.7) (7.8) we find that

$$(7.10) \quad \int_0^{\delta(P)} \frac{dv}{\rho(v)} \leq \int_0^1 \frac{\|\varphi''(x_t)(y-x, y-x)\|_F}{\|\varphi'(x_t)(y-x)\|_F} dt \leq \frac{\beta}{\alpha} \|y-x\|_E$$

These formula give us the following lower and upper bounds for the attributes of the paths of $D = \varphi(C)$:

$$(7.11) \quad \begin{aligned} R &= \alpha_m^2 / \beta \text{ (lower bound to radii of curvature)} \\ \Delta &= \alpha_M \text{ diam } C \text{ (upper bound to path length)} \\ \Theta &= (\beta / \alpha_m) \text{ diam } C \text{ (upper bound to path deflection).} \end{aligned}$$

Using theorem 6.7 yields then the

Theorem 7.1 : Let (7.2) (7.3) (7.4) hold, and let R, Δ, Θ be defined by (7.11).

i) If

$$(7.12) \quad \Theta = \frac{\beta}{\alpha_m} \text{ diam } C \leq \frac{\pi}{2}$$

Then, for any z such that :

$$(7.13) \quad d(z, \varphi(C)) < R = \alpha_m^2 / \beta,$$

the non-linear least square problem (7.1) has the following properties :

(7.14) existence of a unique global minimum \hat{x}

(7.15) Any minimizing sequence is a Cauchy sequence in E converging towards \hat{x} .

(7.16) no local minimum

(7.17) lipschitz-continuity of the $z \rightarrow \hat{x}$ mapping :

$$\alpha_m \|\hat{x}_1 - \hat{x}_2\| \leq (1 - \frac{d}{R})^{-1} \|z_1 - z_2\| \text{ as soon as}$$

(7.18)

$$\|z_1 - z_2\| + \text{Max}_{j=1,2} d(z_j, \varphi(C)) \leq d < R$$

ii) If :

$$(7.19) \quad \frac{\pi}{2} \leq \Theta = \frac{\beta}{\alpha_m} \text{ diam } C < \pi$$

and

$$(7.20) \quad \left(\frac{\alpha_M}{\alpha_m} - 1 \right) \Theta < |\tan \Theta|$$

Then, for any z such that

$$(7.21) \quad d(z, \varphi(C)) < R \sin \Theta + (\alpha_M - \alpha_m) \text{ diam } C \cos \Theta$$

the non linear least square problem (7.1) has still properties (7.14 thru 18) (with R in the r.h.s of the second line of (7.18) replaced by the r.h.s of (7.21)).

Remark that, under the same hypothesis, the γ -size curvature condition would read

$$(7.22) \quad \frac{\beta}{\alpha_m} \text{ diam } C < 2\sqrt{2}$$

and imply (7.14) (7.15) (7.17) (7.18) only.

REFERENCES

- [1] G. Chavent, "Local stability of the output least square parameter estimation techniques", Mat. Applic. Comp., Vol.2, n° 1, pp.3-22, 1983.
- [2] G. Chavent, "On the uniqueness of local minima for general abstract non-linear least square problems", Inverse Problems 4, pp.417-433, 1988.
- [3] G. Chavent, "Quasiconvex sets and size \times curvature condition ; application to non-linear inversion", INRIA Report 1017, Avril 1989.
- [4] G. Chavent, "A new sufficient condition for the wellposedness of non-linear least-square problems arising in identification and control", submitted to the 9th Int. Conf. on Analysis and Optimization of systems, June 12-15, 1990.
- [5] J.P. Aubin, "Mathematical Methods of Game and Economic Theory", North Holland, 1979.

