CHAPTER 3

CONVEX FUNCTIONS ON RIEMANNIAN MANIFOLDS

In this chapter we present, in a systematic manner, the basic concepts and theorems regarding the Riemannian convexity of real functions.

Let (M,g) be a complete Riemannian manifold. A subset A of M is called totally convex if A contains every geodesic γ_{xy} of M whose endpoints x and y belong to A (§1).

Let A be a totally convex set in M. A function $f:A \longrightarrow \mathbb{R}$ is called convex if

$$(f\circ\gamma_{xy})(t) \le (1 - t)f(x) + tf(y),$$

for every geodesic $\gamma_{xy}:[0,1] \longrightarrow A$, $\gamma_{xy}(0) = x$, $\gamma_{xy}(1) = y$, for all $t \in [0,1]$ and $x,y \in A$. This and equivalent definitions are analysed in §2.

A diffeomorphism does not destroy the convexity of a function. An advantage of this remark is that by suppressing unnecessary coordinates the full generality of the convexity theory becomes evident. Needless to repeat, the convexity property does not depend on nonlinear coordinate transformations (§3).

Geometrically, convex functions are important because (among other reasons) they give rise to totally convex sets in the sense that the sublevel sets of a convex function are totally convex. A convex function is also continuous on the interior of its domain (§3).

The notion of subgradient is essential to describe the properties of a continuous convex function (§4).

The convexity definition of a C^1 function f (§5) is equivalent to the fact that the totally geodesic hypersurfaces tangent to $G(f)$ lie below $G(f)$.

Reformulations of the convexity definition of a C^2 function (§6) lead to a generalization which does not ask that the domain be totally convex: a C^2 function $f:A \longrightarrow \mathbb{R}$ is called convex if $\mathrm{Hess}_g f$ is positive semidefinite. In case that A is a submanifold of M and $\mathrm{Hess}_g f$ is positive definite, the pair $(A, \mathrm{Hess}_g f)$ is a new Riemannian manifold. Solutions of the inequation $\mathrm{Hess}_g f > 0$ on a part of the sphere S^2 and on the Poincare plane (§6) are given as examples . Concerning the equation $\mathrm{Hess}_g f = 0$, whose solutions are called linear affine functions, we have the following result: An n-dimensional Riemannian manifold (M,g) is the Riemannian product of an $(n - p + 1)$ -dimensional Riemannian manifold and the Euclidean space $\mathbb{R}^{p-1}$ (locally at least) if and only if the vector space of all linear affine functions on M has dimension p (§6).

A program of type $\min_{x\in A} f(x)$ is called convex if there exists a Riemannian metric on $M \supset A$ such that the Riemannian manifold (M,g) is complete, the set A is totally convex in (M,g) and f is a convex function. Basic properties of convex programs are described in §7.

The theory of dual problem (§8) and the Kuhn-Tucker Theorem (§9) on Riemannian manifolds show that it is important to look for those Riemannian metrics which are able to produce the convexity of a function, if such a metric exists.

Quasiconvex functions on Riemannian manifolds are defined as those functions whose sublevel sets are totally convex. An equivalent definition and basic properties of such functions are presented in §10.

The convexity and nonpositive curvature make good house, which permits an interesting description of the distance from a point to a closed totally convex set (§11) and of the distance between two closed totally convex sets (§12).

References: [3], [5], [7], [10], [11], [13], [14], [16], [20], [24], [30], [39], [41], [43]-[46], [50] - [52], [55], [56],[58], [71]-[75], [77]-[83], [84]-[87], [94], [95], [97], [99], [100], [105], [107], [108], [119], [125], [133], [134], [136].

§1. Convex sets in Riemannian manifolds

Let (M,g) be a complete n-dimensional Riemannian manifold with Riemannian connection ∇. Let x and y be two points in M and

$$\gamma_{xy} : [0,1] \longrightarrow M$$

a geodesic joining the points x and y, i.e.,

$$\gamma_{xy}(0) = x \ , \ \gamma_{xy}(1) = y \ .$$

For a subset A in a Riemannian manifold M with the property that γ_{xy} is unique for x,y in A, the *convexity* of A must obviously be defined by: x , y in A implies $\gamma_{xy} \subset A$.

1.1. **Lemma**. *Fix $x_o \in M$ and consider a metric ball $B(x_o,r)$ with the boundary $S(x_o,r) = \partial B(x_o,r)$. Then there exists $b > 0$ such that whenever $r \in (0,b)$, any geodesic tangent to $S(x_o,r)$ at the point $x \in S(x_o,r)$ lies in the exterior of $B(x_o,r)$ at least in a neighborhood of x.*

Proof. If $(\xi^1,\dots,\xi^n)$ is a system of normal coordinates around the point x_o , then $\xi(x_o) = 0$ and

$$B(x_o \ , \ r) = \left\{ x \in M \,\middle|\, \sum_{i=1}^{n} (\xi^i(x))^2 < r^2 \right\} .$$

Let $\xi^i = \xi^i(t)$, $i = 1,\dots,n$, $t \in I$, the parametric equations of a geodesic which is tangent to $S(x_o,r)$ at the point $x = (\xi^1(0),\dots,\xi^n(0))$, $0 \in I$. We consider the function

$$F(t) = \sum_{i=1}^{n} (\xi^i(t)^2, \ t \in I.$$

By the tangency conditions we have

$$F(0) = r^2, \ 0 = \sum_{i=1}^{n} \xi^i(0) \frac{d\xi^i}{dt}(0) = \frac{1}{2} \frac{dF}{dt}(0).$$

Hence t = 0 is a critical point of F.

Taking into account the differential equations of the geodesics we find

$$\frac{1}{2} \frac{d^2F}{dt^2} = \sum_{j,k=1}^{n} \left(\delta_{jk} + \sum_{i=1}^{n} \Gamma^i_{jk} \xi^i(t) \right) \frac{d\xi^j}{dt} \frac{d\xi^k}{dt} .$$

Since $\Gamma^i_{jk}(x_o) = 0$, there exists $b > 0$ such that the matrix

$$\left[\delta_{jk} + \sum_{i=1}^{n} \Gamma^i_{jk} \xi^i(t) \right]$$

is positive definite on $B(x_o,b)$. If $0 < r < b$, there exists a neighborhood $J_o \subset I$ such that

$$\frac{d^2F}{dt^2} > 0, \ \forall \ t \in J_o.$$

It follows

$$F(t) > F(0) = r^2 \ , \ \forall \ t \in J_o - \{0\}$$

and hence an arc of the geodesic lies in the exterior of $B(x_o, r)$.

1.2. **Whitehead theorem**. *Each point* $x_o \in M$ *has a convex spherical normal neighborhood* $B(x_o, r)$.

Proof. According to Chapter 1, §5, each point x_o has a convex normal neighborhood W_{x_o}. Obviously there exists $B(x_o, r) \subset W_{x_o}$ and any two points $x, y \in B(x_o, r)$ can be joined by a unique geodesic $\gamma_{xy}(t)$, $t \in [0,1]$, i.e., $\gamma_{xy}(0) = x$, $\gamma_{xy}(1) = y$. Let us show that γ_{xy} is included in $B(x_o, r)$. For this we suppose that γ_{xy} is given by the parametric equations $\xi^i = \xi^i(t)$, $i = 1,\dots,n$, $t \in [0,1]$ and consider the function

$$F(t) = \sum_{i=1}^{n} \left(\xi^i(t)\right)^2, \quad t \in [0,1].$$

Obviously $F(o) < r^2$, $F(1) < r^2$. If γ_{xy} had some point in the exterior of $B(x_o, r)$, then $t \in [0,1]$ would exist such that $F(t) \geq r^2$. Let $t_o \in (0,1)$ be the smallest solution of the equation $F(t) = r^2$. Since $F(t_o) = r^2$ is a local maximum of F it follows

$$0 = \frac{dF}{dt}(t_o) = 2\sum_{i=1}^{n} \xi^i(t_o)\,\frac{d\xi^i}{dt}(t_o)$$

and hence γ_{xy} is tangent to $S(x_o, r)$ at $z = (\xi^1(t_o),\dots,\xi^n(t_o))$. On the other hand, the preceding lemma implies $F(t) > r^2$ for $t \neq t_o$, t in a neighborhood of t_o, which is in contradiction with $F(t) \leq F(t_o) = r^2$.

When the uniqueness of a geodesic joining two points fails there is no single best definition of convexity of selected subset [87]. An alternative would be: the subset A of M is *convex* if two arbitrary points x,y in A can be joined by a geodesic contained in A. However the geometrical properties of convex functions require the following

1.3. **Definition**. A subset A of M is said to be *totally convex* if A contains every geodesic γ_{xy} of M whose endpoints x and y are in A.

The whole of the manifold M is totally convex and, conventionally, so is the empty set. In a hyperboloid of revolution, the minimal circle is totally convex, but a single point is not; in a sphere, any proper subset is not totally convex.

1.4. **Theorem**. *Every intersection of totally convex sets is totally convex.*

Proof. Let A and B be two totally convex sets in M and $A \cap B$ their intersection. Let $x,y \in A \cap B$. Then $x,y \in A$ and $x,y \in B$. Since A is totally convex, the geodesics γ_{xy} joining x to y are included in A. Analogously γ_{xy} are included in B and hence they lie in $A \cap B$. This means that $A \cap B$ is totally convex. The generalization to every intersection is obvious.

Remarks. 1) In general, the union of totally convex sets is not totally convex.

2) The total convexity property is Riemannian metric dependent through geodesics, so that a subset of M may be totally convex in one Riemannian metric on M but not in another. The diffeomorphisms of (M,g) do not destroy the total convexity.

Total convexity of sets and the geometry of the space are intimately related. As example we quote a theorem in [13].

1.5. **Cheeger-Gromoll theorem**. *Let γ be a ray with $\gamma(0) = x$ and $B(\gamma(t),t)$ the open metric ball of radius t centered at $\gamma(t)$. Define the open half-space*

$$B_\gamma = \bigcup_{t>0} B(\gamma(t),t).$$

If M has nonnegative curvature, then the closed complement $M - B_\gamma$ of any half-space B_γ is totally convex.

§2. Convex functions on Riemannian manifolds

Suppose that (M,g) is a complete n-dimensional Riemannian manifold and ∇ is the Riemannian connection on M. Let A be a totally convex subset of M. For $x,y \in A$, γ_{xy} denotes the geodesic joining the points x and y. We denote by Γ the set of all geodesic arcs from x to y.

2.1. **Definition**. Let A be a totally convex set in M and $f:A \longrightarrow \mathbb{R}$ be a real-valued function.

1) If $f(\gamma_{xy}(t)) \le (1-t)f(x) + tf(y)$,

$\forall\, x,y \in A,\ \forall\, \gamma_{xy} \in \Gamma,\ \forall\, t \in [0,1]$,

then the function f is called *convex*.

2) If $f(\gamma_{xy}(t)) < (1-t)f(x) + tf(y)$,

$\forall\, x,y \in A,\ x \ne y,\ \forall\, \gamma_{xy} \in \Gamma,\ \forall\, t \in (0,1)$, then the function f is called *strictly convex*.

3) If -f is convex (strictly convex), then f is called *concave (strictly concave)*.

Remarks. 1) For the definition of a convex function $f:M \longrightarrow \mathbb{R}$ it is not necessary for M to be complete.

2) The convexity property is Riemannian connection dependent through geodesics, so that a function $f: M \longrightarrow \mathbb{R}$ may be convex with respect to one Riemannian connection on M but not to another.

All the theorems relating to convex functions have a correspondent for concave functions.

Let us formulate other propositions which are equivalent to the preceding definition.

2.2. **Theorem**. *The function $f : A \longrightarrow \mathbb{R}$ is convex if and only if $\forall\ x,y \in A$, $\forall\ \gamma_{xy} \in \Gamma$ the function $\varphi_{xy} = f\circ\gamma_{xy}$ is convex on [0,1].*

Equivalently, a function $f: A \longrightarrow \mathbb{R}$ is convex if and only if, for every geodesic $\gamma : [a,b] \longrightarrow A$, the function $f\circ\gamma : [a,b] \longrightarrow \mathbb{R}$ is convex in the usual sense, i.e.,

$$(f\circ\gamma)((1 - s)t_1 + st_2) \le (1 - s)(f\circ\gamma)(t_1) + s(f\circ\gamma)(t_2),$$

for all $t_1, t_2 \in [a,b]$ and $s \in [0,1]$.

Proof. Let us suppose that $\varphi_{xy} : [0,1] \longrightarrow \mathbb{R}$ is convex, i.e.,

$$\varphi_{xy}((1-s)t_1 + st_2) \le (1-s)\varphi_{xy}(t_1) + s\varphi_{xy}(t_2),\ \forall\ t_1, t_2 \in [0,1].$$

Particularly for $t_1 = 0$, $t_2 = 1$, we have

$$\varphi_{xy}(s) \le (1-s)\ \varphi_{xy}(0) + s\ \varphi_{xy}(1),\ \forall\ s \in [0,1],$$

i.e.,

$$f(\gamma_{xy}(s)) \le (1-s)\ f(x) + sf(y),\ \forall\ x,y \in A,\ \forall\ \gamma_{xy} \in \Gamma,\ \forall\ s \in [0,1].$$

Conversely, let f be a convex function. If $\gamma_{xy} : [0,1] \longrightarrow A$ is a geodesic joining the points x and y, then the restriction of γ_{xy} to $[t_1, t_2]$ joins the points $\gamma_{xy}(t_1)$ and $\gamma_{xy}(t_2)$. We reparametrize this restriction,

$$\alpha(s) = \gamma_{xy}(t_1 + s(t_2 - t_1)),\ s \in [0,1].$$

Since

$$f(\alpha(s)) \le (1-s)\ f(\alpha(0)) + sf(\alpha(1)),$$

i.e.,

$$f(\gamma_{xy}((1-s)t_1 + st_2)) \le (1-s)f(\gamma_{xy}(t_1)) + sf(\gamma_{xy}(t_2))$$

or

$$\varphi_{xy}((1-s)t_1 + st_2) \le (1-s)\ \varphi_{xy}(t_1) + \varphi_{xy}(t_2),$$

the function φ_{xy} is convex on [0,1].

Now, let $\gamma_1(t;\ x,y)$, $t \in [1,b]$, be a restriction of the natural extension of a geodesic γ_{xy}: $[0,1] \longrightarrow A$ such that

$$\gamma_1(t;\ x,y) \in A,\ \forall\ t \in [1,b].$$

2.3. **Theorem**. *The function $f\colon A \longrightarrow \mathbb{R}$ is convex if and only if $\forall\ x,y \in M$, $\forall\ t \ge 1$, such that $\gamma_1(t;\ x,y) \in A$, we have*

$$f(\gamma_1(t;\ x,y)) \ge (1-t)\ f(x) + tf(y).$$

Proof. Let γ_{xy}: $[0,1] \longrightarrow A$ be a geodesic joining x and y. We denote by $\gamma_1(u;\ x,y)$, $u \in [0,t]$, $t \ge 1$, a natural extension of γ_{xy}, beyond y, so that $\gamma_1(t;\ x,y) \in A$. Setting $u = st$, $s \in [0,1]$ we get the reparametrization $\gamma_1(st;\ x,y)$, $s \in [0,1]$. As f is convex we have

$$f(\gamma_1(st;\ x,y)) \le (1-s)f(x) + sf(\gamma_1(t;\ x,y)),\ \forall\ s \in [0,1].$$

Particularly, for $st = 1$ we find

$$f(\gamma_1(t;\ x,y)) \ge (1-t)f(x) + tf(y),\ \forall\ t \ge 1.$$

The converse is obvious.

Denoting by $\gamma_o(t;\ x,y)$, $t \in [a,0]$, a restriction of the natural extension of a geodesic γ_{xy} : $[0,1] \longrightarrow A$, such that $\gamma_o(t;\ x,y) \in A$, we obtain another definition.

2.4. **Theorem**. *The function $f\colon A \longrightarrow \mathbb{R}$ is convex if and only if $\forall\ x,y \in A$, $\forall\ t \le 0$, such that $\forall\ \gamma_o(t;\ x,y) \in A$, we have*

$$f(\gamma_o(t;\ x,y)) \ge (1-t)f(x) + tf(y).$$

2.5. **Corollary**. *If a convex function $f\colon M \longrightarrow \mathbb{R}$ is upper bounded then f is a constant.*

According to this corollary, the compact Riemannian manifolds do not admit nonconstant (nontrivial) convex functions.

Let $(\mathbb{R},h)$ be the 1-dimensional Euclidean space and $(M \times \mathbb{R},\ g + h)$ the Riemannian product manifold between (M,g) and (R,h). Locally,

$$g = g_{ij}dx^i \otimes dx^j,\ h = dt \otimes dt,\ g + h = g_{ij}dx^i \otimes dx^j + dt \otimes dt$$

and the Christoffel symbols $\bar{\Gamma}^{\alpha}_{\beta\gamma}$, $\alpha,\beta,\gamma = 1,2,\ldots,n,n+1$, attached to $g + h$ are all zero excepting (maybe) $\bar{\Gamma}^{i}_{jk} = \Gamma^{i}_{jk}$, $i,j,k = 1,\ldots,n$, where Γ^i_{jk} are the Christoffel symbols of g. Therefore we can prove that a geodesic which joins the points (x,u) and (y,v) in $M \times \mathbb{R}$ is of the form

$$(\gamma_{xy}(t),\ (1-t)u + tv),\ t \in [0,1],$$

where $\gamma_{xy}(t)$, $t \in [0,1]$, is a geodesic in M which joins x to y. So $A \subset M$ is convex if and only if $A \times \mathbb{R}$ is convex.

2.6. **Theorem**. *Let $A \subset M$ be a totally convex set. The function $f\colon A \longrightarrow \mathbb{R}$ is convex if and only if its epigraph*

$$E(f) = \{(x,u) \mid f(x) \le u\} \subset A \times \mathbb{R}$$

is a convex set.

Proof. First suppose that f is a convex function. Let $(x,u) \in E(f)$, $(y,v) \in E(f)$. We have $u \ge f(x)$, $v \ge f(y)$ and hence

$$(1-t)u + tv \ge (1-t)f(x) + tf(y) \ge f(\gamma_{xy}(t)).$$

So

$$(\gamma_{xy}(t),\ (1-t)u + tv) \in E(f)$$

and hence $E(f)$ is a totally convex set.

Next suppose that $E(f)$ is a totally convex set. Let $x,y \in A$. We have $(x,f(x)) \in E(f)$ and $(y,f(y)) \in E(f)$. On the basis of total convexity of $E(f)$ it follows

$$(\gamma_{xy}(t),\ (1-t)f(x) + tf(y)) \in E(f),$$

i.e.,

$$f(\gamma_{xy}(t)) \le (1-t)f(x) + tf(y)$$

and hence f is convex.

2.7. **Corollary**. *Let A_i be a totally convex subset of M and let $f_i\colon A_i \longrightarrow \mathbb{R}$ be a convex function for every $i \in I$. Suppose $A = \bigcap_{i\in I} A_i \ne \emptyset$, define $f\colon A \longrightarrow \mathbb{R}$, $f(x) = \sup_{i\in I} f_i(x)$ and $S = \{x \in A \mid f(x) < \infty\}$. Then S is a totally convex set and f is a convex function on S.*

Let us now give a way of constructing convex functions on M starting from totally convex sets in $M \times J$, where $J = (a, \infty)$. For this, let (J,h) be the 1-dimensional Euclidean space and $(M \times J,\ g + h)$ be the product manifold of (M,g) and (J,h). A geodesic joining the points (x,u) and (y,v) of $M \times J$ is of the form $(\gamma_{xy}(t),\ (1 - t)u + tv)$, $t \in [0,1]$, where $\gamma_{xy}(t)$, $t \in [0,1]$, is a geodesic in M from x to y.

2.8. **Theorem**. *Let F be a totally convex (nonvoid) subset of $M \times J$. The function*

$$f(x) = \inf \{u \mid (x,u) \in F\}$$

is convex on the projection of F onto M.

Proof. The projection of F onto M is a totally convex subset of M.

Let (x,u) and (y,v) be two points in F. By the definition of f we have

(1) $f(x) \le u,\ f(y) \le v.$

But F is totally convex and hence,

$$\forall\ t \in [0,1],\ (\gamma_{xy}(t),\ (1-t)u + tv) \in F.$$

Taking again into account the definition of f we find

(2) $f(\gamma_{xy}(t)) \le (1-t)u + tv.$

We see that (2) holds whenever (1) holds and hence the epigraph E(f) is a totally convex set, i.e., f is convex.

Let $X \in T_xM$, $\|X\| = 1$ and $\gamma: (-a,a) \longrightarrow A$ be a geodesic such that $\gamma(0) = x$, $\dot{\gamma}(0) = X$. Let $f: A \longrightarrow \mathbb{R}$ be a continuous function. We define

$$Cf(x;\ X) = \liminf_{t \to 0} \frac{1}{t^2} [f(\gamma(t)) + f(\gamma(-t)) - 2f(\gamma(0))],$$

$$Cf(x) = \inf_{\|X\|=1} Cf(x;X).$$

The number Cf(x) measures the deviation of f from being convex.

2.9. **Theorem**. *f is convex if and only if $Cf \ge 0$.*

Proof. Suppose $f: A \longrightarrow \mathbb{R}$ is convex. Then $f \circ \gamma: (-a,a) \longrightarrow \mathbb{R}$ is convex, i.e.,

$$(f\circ\gamma)(\lambda t + (1-\lambda)s) \le \lambda(f\circ\gamma)(t) + (1-\lambda)(f\circ\gamma)(s),$$

for all $t,s \in (-a,a)$ and $\lambda \in [0,1]$. Fixing $\lambda = \frac{1}{2}$ and $t + s = 0$ we find

$$2f(\gamma(0)) \le f(\gamma(t)) + f(\gamma(-t)),\ \forall\ t \in (-a,a).$$

Consequently $Cf(x;X) \ge 0$ and $Cf(x) \ge 0$, where $x = \gamma(0)$.

The converse follows from the proposition " *$\varphi : (-a,a) \longrightarrow \mathbb{R}$ is convex if and only if φ is continuous and the lower second symmetric derivative,*

$$\liminf_{h \to 0} \frac{\varphi(t + h) + \varphi(t - h) - 2\varphi(t)}{h^2},$$

is nonnegative on (-a,a)".

Let $f: A \longrightarrow \mathbb{R}$ be a continuous function. A function $\bar{f}: A \longrightarrow \mathbb{R}$ is said *to support* f at $x \in M$ if and only if $\bar{f}$ is continuous near x and $\bar{f}(x) = f(x)$, $\bar{f} \le f$.

2.10. **Theorem** [136]. *If $\bar{f}$ supports f at x, then $C\bar{f}(x) \le Cf(x)$. If f is supported at every point of A by a convex function, then f is convex.*

The set $A \subset M$ is called *star-shaped* at $x_o \in A$ if $\gamma_{x_o x}(t) \in A$ whenever $x \in A$ and $t \in (0,1)$, where $\gamma_{x_o x}$ is any geodesic in A starting from x_o. Obviously, any totally convex set is star-shaped at each of its points. But there exist star-shaped sets which are not totally convex. For example the set in Figs. 4,5 are star-shaped in $\mathbb{R}^2$ (and are not totally convex); the set in Fig. 4 is star-shaped at any x_o in the

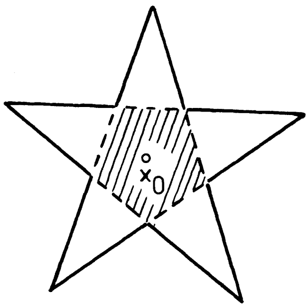

Fig. 4

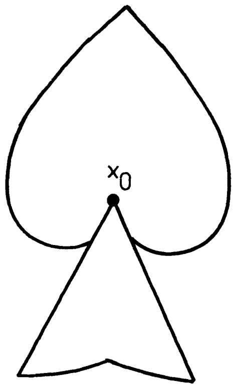

Fig. 5

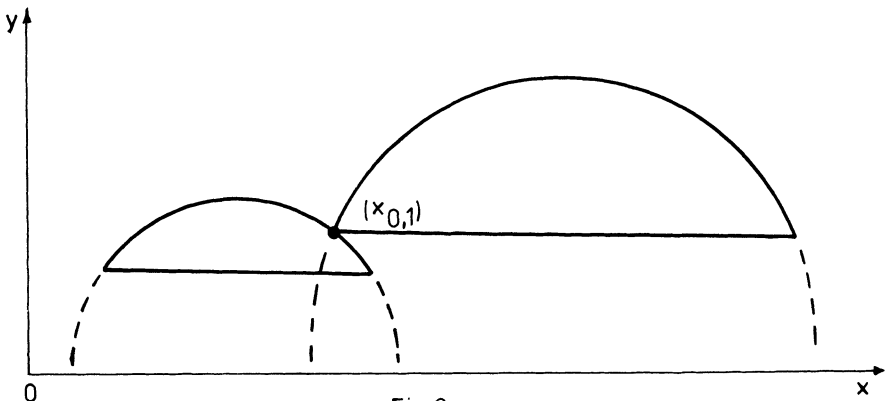

Fig. 6

shaded region; the set in Fig. 5 is star-shaped at x_o .

The set in Fig. 6 is star-shaped at $(x_o ,1)$ in the Poincare plane.

Let A be a star-shaped set at x_o and $f : A \longrightarrow \mathbb{R}$. The function f is called *convex at* x_o if

$$f(\gamma_{x_o x}(t)) \le (1-t)f(x_o) + tf(x), \ \forall\ x \in A,\ t \in (0,1).$$

If $-f$ is convex at x_o , then f is called *concave at* x_o .

Any convex function is convex at each point of its domain.

§3. Basic properties of convex functions

Let (M,g) be a complete finite-dimensional Riemannian manifold and ∇ be the Riemannian connection. Let $F\colon M \longrightarrow M$ be a diffeomorphism and $F_*\nabla$ be the connection which results by the transformation of ∇. If γ is a geodesic of (M,∇), then $F\circ\gamma$ is a geodesic of $(M,F_*\nabla)$.

3.1. **Theorem**. *Suppose that* $A \subset M$ *is a totally convex set and* $f\colon A \longrightarrow \mathbb{R}$ *is a convex function. If* $F\colon M \longrightarrow M$ *is a diffeomorphism then* $f\circ F^{-1}$ *is convex on the set* $F(A)$.

Proof. Let $x,y \in A$ and $\gamma_{xy}(t)$ be a geodesic which joins the points x and y. The set F(A) is totally convex and the geodesic $F\circ\gamma_{xy}$ joins the points F(x) and F(y). We have

$$(f\circ F^{-1})(F(\gamma_{xy}(t))) = f(\gamma_{xy}(t)) \le (1-t)f(x) + tf(y) =$$

$$= (1-t)(f\circ F^{-1})(F(x)) + t(f\circ F^{-1})(F(y)),$$

i.e., $f\circ F^{-1}$ is convex on the set F(A).

This theorem shows that the geodesic convexity of a function is independent of the particular system of coordinates selected.

Convex functions can often be combined in some way to determine other convex functions. The following two theorems are examples.

3.2. **Theorem**. *Let* $f\colon A \longrightarrow \mathbb{R}$ *be a convex function defined on the totally convex set A. Let I be a convex set in* $\mathbb{R}$ *that contains f(A). If* $\varphi : I \longrightarrow \mathbb{R}$ *is an increasing convex function, then* $\varphi\circ f$ *is a convex function on A.*

Proof. We have

$$f(\gamma_{xy}(t)) \le (1-t)f(x) + tf(y)$$

and

$$\varphi\circ f(\gamma_{xy}(t)) \le \varphi((1-t)f(x) + tf(y)) \le (1-t)\varphi\circ f(x) + t\varphi\circ f(y),$$

which proves the theorem.

Let us show that the set of convex functions is closed with respect to positively linear combinations.

3.3. **Theorem**. *If f_i, $i = 1,2,\dots,n$ are convex functions on $A \subset M$ and $c_i \ge 0$, then $\sum_i c_i f_i$ is convex on A.*

Proof. By hypothesis we have

$$f_i(\gamma_{xy}(t)) \le (1-t)f_i(x) + tf_i(y).$$

It follows

$$c_i f_i(\gamma_{xy}(t)) \le (1-t)c_i f_i(x) + tc_i f_i(y)$$

and

$$(\sum_i c_i f_i)(\gamma_{xy}(t)) \le (1-t)(\sum_i c_i f_i)(x) + t(\sum_i c_i f_i)(y).$$

3.4. **Theorem**. *If f is a convex function on A and c is a real number, then the sublevel set $A^c = \{z|z \in A,\ f(z) \le c\}$ is a totally convex subset of A. Particularly A^c is connected.*

Proof. Let x and y be two points of A which satisfy $f(x) \le c$, $f(y) \le c$ and $\gamma_{xy}(t) = z$ an arbitrary point on a geodesic which joins x and y. As f is convex we find

$$f(z) = f(\gamma_{xy}(t)) \le (1-t)f(x) + tf(y) \le (1-t)c + tc = c.$$

So, any point $z = \gamma_{xy}(t)$ satisfies $f(z) \le c$ when x and y satisfy this inequality.

3.5.**Corollaries**.*1) f is a constant on every closed geodesic in A.*

2) Let f_i, $i = 1,2,\dots,n$, be convex functions on A and c_i be real numbers. The subset

$$\{z|z \in A,\ f_i(z) \le c_i,\ i = 1,2,\dots,n\}$$

is totally convex in A.

Proof. 1) According to the preceding theorem, $A^c = \{x \in A|f(x) \le c\}$ is totally convex. It follows that for any closed geodesic we have $f(\gamma(s)) \le f(\gamma(t))$, $\forall\ s,t \in \mathbb{R}$. Hence $f(\gamma(t)) = \text{const}$.

2) We apply the fact that $\{z|f_i(z) \le c_i,\ i = \text{fixed}\}$ is totally convex and that an intersection of totally convex sets is totally convex.

The following continuity theorem seems to be the most important, although stronger results can be stated.

3.6. **Theorem**. *Let A be a totally convex set with the nonvoid interior and $f: A \longrightarrow \mathbb{R}$ be a convex function.*

1) f is continuous on int$A$;

2) if $x_o \in A$ is a boundary point, then

$$\liminf_{x \to x_o} f(x) \le f(x_o).$$

Proof. 1) Let $x_o \in \mathrm{int}A$ and $B(x_o; r)$ be an open ball centered at x_o and of sufficiently small radius $r > 0$. Choose c such that the totally convex set $A^c = \{x \in A | f(x) \le c\}$ contains $\bar{B}(x_o, r)$. Let $\gamma : [-1,1] \longrightarrow M$ be a geodesic in $\bar{B}(x_o, r)$, parametrized proportionally to the distance $d(x_o, x)$, and such that $\gamma(-1) = x_1$, $\gamma(0) = x_o$, $\gamma(1) = x_2$. We denote $\gamma(t) = x$, where $t = \dfrac{d(x_o, x)}{r} \in [0,1]$. As f is convex we have

$$f(\gamma(t)) \le (1-t)f(x_o)) + tf(x_2) \le (1-t)f(x_o) + tc, \ t \in [0,1],$$

i.e.,

(3) $$f(x) - f(x_o) \le t(c - f(x_o)).$$

The geodesic arc joining x_1 and x is the restriction $\gamma(u)$, $u \in [-1,t]$. Setting $u = -1 + s(t + 1)$, $s \in [0,1]$, we get the reparametrization

$$\alpha(s) = \gamma(-1 + s(t + 1)), \ s \in [0,1].$$

Obviously

$$\alpha(o) = \gamma(-1) = x_1, \ \alpha\left(\frac{1}{1 + t}\right) = \gamma(0) = x_o, \ \alpha(1) = \gamma(t) = x .$$

Due to the convexity of f we have

$$f(\alpha(s)) \le (1-s)f(x_1) + sf(x) \le (1-s)c + sf(x).$$

It follows

$$f(x_o) \le \frac{t}{1 + t} c + \frac{1}{1 + t} f(x),$$

or

(4) $$f(x) - f(x_o) \ge - t(c - f(x_o)).$$

The relation (3) and (4) imply

$$|f(x) - f(x_o)| \le t(c - f(x_o)).$$

As $t = \dfrac{d(x_o, x)}{r}$, we obtain $\lim_{x \to x_o} f(x) = f(x_o)$.

2) Let $x_o \in A$ be a boundary point and $\gamma_{xx_o} : [0,1] \longrightarrow A$ be a geodesic from x to x_o. We have

$$\liminf_{x \to x_o} f(x) \le \lim_{t \nearrow 1} f(\gamma_{xx_o}(t)) \le \lim_{t \nearrow 1} ((1-t)f(x) + tf(x_o)) = f(x_o).$$

3.7. **Corollaries**. *1) A convex function on M is necessarily continuous.*

2) Let T be an arbitrary set and f be a real-valued function on MxT. If f(x,t) is convex as a function of x for each t and bounded above as a function of t for each x, then

$$h(x) = \sup_{t \in T} \{f(x,t)\}, \quad x \in M,$$

depends continuously on x.

Indeed, h is convex, being a point-wise supremum of a collection of convex functions.

3.8. **Definition.** A function $f:M \longrightarrow \mathbb{R}$ is called *Lipschitz continuous* if there exists a real number B such that

$$|f(x) - f(y)| \le B\, d(x,y), \quad \forall\, x,y \in M,$$

where d(x,y) is the distance between the points x and y. The positive number B is called a *Lipschitz constant* for f.

The next lemma shows that Lipschitz continuity on a Riemannian manifold is a local property.

3.9. **Lemma**. *A function $f:M \longrightarrow \mathbb{R}$ is Lipschitz continuous with Lipschitz constant B if and only if f is Lipschitz continuous with Lipschitz constant B in a neighborhood of each point of M, i.e., for each point $x \in M$, there exists a neighborhood U_x of x such that*

$$|f(x_1) - f(x_2)| \le B\, d(x_1,x_2),$$

for every $x_1, x_2 \in U_x$.

Proof. Lipschitz continuity on M implies the local condition since one may take $U_x = M$ for every $x \in M$.

To show the converse, recall that for any $x_1, x_2 \in M$ we have $d(x_1, x_2) = \inf_\omega L(\omega)$, where ω ranges over all piecewise C^∞ regular curves ω: $[0,1] \longrightarrow M$, $\omega(0) = x_1$, $\omega(1) = x_2$. Thus to establish Lipschitz continuity of f with Lipschitz constant B it is sufficient to prove that for any such curve ω,

$$|f(x_1) - f(x_2)| \le B\, L(\omega).$$

Choose a finite subdivision of $[0,1]$ by points $t_o = 0 < t_1 < \dots < t_n = 1$ such that for all $i = 0,\dots,n-1$, the image $\omega([t_i,t_{i+1}])$ is contained in a neighborhood U_x satisfying the condition of the lemma for some $x \in M$. Such a choice is possible because the U_x - s form an open cover of M. Then

$$|f(x_1) - f(x_2)| \le \sum_{i=o}^{n-1} |f(\omega(t_i)) - f(\omega(t_{i+1}))| \le$$

$$\le B \sum_{i=o}^{n-1} d(\omega(t_i),\ \omega(t_{i+1})) \le B\ d(x_1,x_2).$$

3.10. **Corollary**. *A convex function $f\colon M \longrightarrow \mathbb{R}$ is Lipschitz continuous on any compact subset of M.*

Convex functions have many structural implications on manifolds. Let us give an example of implication.

A *closed curve* or *loop* in M at the point x is a curve $\alpha\colon[0,1] \longrightarrow M$ with $\alpha(0) = \alpha(1) = x$. The *product γ of two loops* α,β based at $x \in M$ is

$$\gamma = \alpha * \beta,\ \gamma(t) = \begin{cases} \alpha(2t) & \text{if } 0 \le t \le 1/2 \\ \beta(2t-1) & \text{if } 1/2 \le t \le 1 \end{cases}.$$

Two loops α and β based at x are *homotopic* if there exists a continuous map

$$H : [0,1] \times [0,1] \longrightarrow M,$$

with

$$H(t,o) = \alpha(t),\ H(t,1) = \beta(t),\ H(o,s) = H(1,s) = x.$$

Let us denote by $\pi_1(M,x)$ the set of homotopy classes of loops based at x. If $[\alpha]$ and $[\beta]$ belong to $\pi_1(M,x)$ then the *product* is defined by

$$[\alpha]\circ[\beta] = [\alpha*\beta],$$

i.e., the product of two homotopy classes is defined to be the class determined by the product of their representative elements. $(\pi_1(M,x),\circ)$ is a group named the *fundamental group* of the *curve-connected topological space* M based at x. The unit element of this group is the homotopy class of the constant loop at x.

The fundamental group $\pi_1(M,x)$ depends, up to some isomorphism, on the space M, and not on the base point x selected. Therefore sometimes we used the symbol $\pi_1(M)$.

3.11. **Theorem**. *Let $f\colon M \longrightarrow \mathbb{R}$ be a convex function and*

$$M^c = \{x \in M \mid f(x) \le c\}.$$

For each value c of f, the inclusion $M^c \subset M$ induces a homomorphism of the fundamental group $\pi_1(M^c)$ onto $\pi_1(M)$.

Proof. Each element of $\pi_1(M,x)$, $x \in M^c$, can be represented by a geodesic loop γ at x. As M^c is totally convex, γ lies in M^c.

§4. Directional derivatives and subgradients

Let (M,g) be a complete finite-dimensional Riemannian manifold and let A be a totally convex set in M. The set T_xA of tangent vectors to A at x is a convex cone in T_xM. Indeed, suppose $X_x, Y_x \in T_xA$. Obviously tX_x, $tY_x \in T_xA$, $\forall\, t \ge 0$. Suppose that $\alpha,\beta\colon [o,\varepsilon] \longrightarrow A$ satisfy $\alpha(o) = \beta(o) = x$, $\alpha'(o)=X_x$, $\beta'(o) = Y_x$. For sufficiently small $t > 0$ there is a unique minimal geodesic γ_t from $\alpha(t) = \gamma_t(o)$ to $\beta(t) = \gamma_t(1)$. Since A is totally convex, γ_t lies in A. Then for $0 \le s \le 1$, $\tau(t) = \gamma_t(s)$ defines a curve in A such that $\tau'(o) = (1 - s)X_x + sY_x$ (see also the proof of the Theorem 4.2).

Let $X_x \in T_xA$ and let $\gamma(t)$, $t \in I$, $0 \in I$ be the geodesic for which $\gamma(0) = x$, $\dot\gamma(0) = X_x$ and $\gamma(I) \subset A$. The place of 0 in I will be imposed by the context.

4.1. **Definition**. Let $f\colon A \longrightarrow \mathbb{R}$ be a real-valued function. The limit

$$Df(x;X_x) = \lim_{t \searrow o} \frac{f(\gamma(t)) - f(x)}{t}$$

(if any) is called the *unilateral directional derivative* of f at x with respect to X_x.

Since

$$-Df(x; -X_x) = \lim_{t \nearrow o} \frac{f(\gamma(t)) - f(x)}{t},$$

the unilateral directional derivative $Df(x; X_x)$ is *bilateral* if and only if $Df(x; -X_x)$ exists and $Df(x; -X_x) = -Df(x; X_x)$.

4.2. **Theorem**. *If $f\colon A \longrightarrow \mathbb{R}$ is a convex function and $\gamma(t)$, $t \in I$ is the geodesic for which $\gamma(o) = x$ and $\dot\gamma(o) = X_x$, then:*

1) *for a fixed* X_x , *the function* $Q : I \cap (o,\infty) \longrightarrow \mathbb{R}$ *defined by*

$$Q(t) = \frac{f(\gamma(t)) - f(x)}{t}$$

is nondecreasing;

2) $Df(x;\ X_x)$ *exists and is equal to* $\inf_t Q(t)$;

3) $Df(x;\ X_x)$ *is convex and positively homogeneous in* X_x .

Also $Df(x;\ 0_x) = 0$ *and* $-Df(x;\ -X_x) \le Df(x;\ X_x)$.

Proof. 1) Since f is convex, the function $g(t) = f(\gamma(t))$, $t \in I$, is convex. If $0 < t_1 < t_2$, then we have

$$\frac{g(t_1) - g(o)}{t_1} \le \frac{g(t_2) - g(o)}{t_2} .$$

This implies that

$$\frac{f(\gamma(t_1)) - f(x)}{t_1} \le \frac{f(\gamma(t_2)) - f(x)}{t_2} ,$$

i.e., Q is nondecreasing.

2) From 1) it follows that for any sequence of positive numbers $\{t_n\}$ which converges to zero, the sequence $\{Q(t_n)\}$ is nonincreasing so that $\{Q(t_n)\}$ has a limit which is $Df(x;\ X_x) = \inf_t Q(t)$.

3) Let $\tau(s,t)$, $0 \le s \le 1$, $0 \le t \le t_o$, be a C^∞ deformation of curves in A such that

(a) $\tau(s,o) = x$,

(b) $\tau(o,t) = \alpha(t) = \exp(tX_x)$, $\tau(1,t) = \beta(t) = \exp(tY_x)$,

$X_x, Y_x \in T_xA$,

(c) for each t, the function $s \longrightarrow \tau(s,t)$ is a geodesic parametrized proportionally to the arclength.

If t_o is sufficiently small, then it is obvious that we can construct such a deformation.

Condition (c) shows that $s \longrightarrow \partial_t\tau(s,t)$ is a Jacobi vector field along the geodesic $s \longrightarrow \tau(s,t)$. On the other hand, by (a), the Jacobi equation reduces at t = 0 to

$$\nabla_s \nabla_s \partial_t \tau(s,o) = 0.$$

This, together with the conditions

$$\partial_t\tau(o,o) = \dot\alpha(o) = X_x , \quad \partial_t\tau(1,o) = \dot\beta(o) = Y_x ,$$

obtained from (b), yield

$$\partial_t\tau(s,o) = (1 - s)X_x + sY_x .$$

Therefore

$$Df(x;(1-s)X_x + sY_x) = \lim_{t\searrow o} \frac{f(\tau(s,t)) - f(x)}{t} \le \text{(using the convexity of f)}$$

$$\lim_{t\searrow o} \frac{(1-s)f(\tau(o,t)) + sf(\tau(1,t)) - f(x)}{t} =$$

$$= \lim_{t\searrow o} \frac{(1-s)f(\alpha(t)) + sf(\beta(t)) - (1-s)f(x) - sf(x)}{t} =$$

$$= (1-s)Df(x; X_x) + sDf(x; Y_x),$$

i.e., $Df(x; X_x)$ is convex in X_x .

Prove now that $Df(x; X_x)$ is positively homogeneous in X_x .Indeed, $\forall$ s > o, we have

$$Df(x; sX_x) = \lim_{t\searrow o} \frac{f(\gamma(st)) - f(x)}{t} = s \lim_{u\searrow o} \frac{f(\gamma(u)) - f(x)}{u} =$$

$$= s\, Df(x; X_x).$$

One has $Df(x; 0_x) = 0$ by definition. Furthermore, by the convexity of $Df(x; X_x)$ we get

$$0 = Df(x; \tfrac{1}{2}(-X_x) + \tfrac{1}{2}X_x) \le \tfrac{1}{2}Df(x; -X_x) + \tfrac{1}{2}Df(x; X_x),$$

i.e.,

$$-Df(x; -X_x) \le Df(x; X_x).$$

Supplements. Let $f\colon (-a,a) \longrightarrow \mathbb{R}$ be continuous (here a may be ∞).

1) If f is convex, then the right and left derivatives f'_+ and f'_- exist on all of (-a,a) and are nondecreasing. Hence the second derivative exists almost everywhere and is nonnegative.

2) If f is convex on (-a,o] and [o,a), and $f'_-(o) \le f'_+(o)$, then f is convex.

3) If o is a minimum point of f and f is convex on (-a,o] and [o,a), then f is convex.

The set T_x^*A of cotangent vectors to A at a point x is a convex cone in T_x^*M. Let y be a generic point in A and $\gamma_{xy}(t)$, $t \in [0,1]$, be a geodesic such that

$$\gamma_{xy}(o) = x, \quad \gamma_{xy}(1) = y, \quad \dot\gamma_{xy}(o) = X_x \in T_xA.$$

We denote by Γ the set of all geodesics from x to y.

4.3. **Definition**. Let $f\colon A \longrightarrow \mathbb{R}$ be a convex function. A 1-form $\omega_x \in T_x^*A$ is called the *subgradient* of f at x if

$$f(y) \geq f(x) + \omega_x(\dot{\gamma}_{xy}(o)), \ \forall\, y \in A, \ \forall\, \gamma_{xy} \in \Gamma.$$

4.4. **Definition**. Let $f: A \longrightarrow \mathbb{R}$ be a convex function. The set of all subgradients of f at x is called the *subdifferential of f at x* and is denoted by $\partial f(x)$. The multiform map $\partial f\colon x \longrightarrow \partial f(x)$ is called the *subdifferential* of f.

4.5. **Theorem**. *If $f: A \longrightarrow \mathbb{R}$ is a convex function, then the set $\partial f(x) \subset T_x^* A$ contains at least one element.*

Proof. Since f is convex on A, we have

$$f(y) - f(x) \geq \frac{f(\gamma_{xy}(t)) - f(x)}{t} \geq Df(x;\, X_x),$$

where $t > 0$ and $Df(x;\, X_x)$ is the unilateral directional derivative of f at x with respect to $\dot{\gamma}_{xy}(o) = X_x$. But $T_x^* A$ is a convex set in $T_x^* M$ and $X_x \longrightarrow Df(x;\, X_x)$ is a convex function. Therefore there exists $\omega_x \in T_x^* M$ such that $\omega_x(X_x) \leq Df(x;\, X_x)$ and hence

$$f(y) - f(x) \geq \omega_x(X_x),$$

i.e., $\omega_x \in \partial f(x)$.

4.6. **Theorem**. *If $f: A \longrightarrow \mathbb{R}$ is a convex function, then the set $\partial f(x)$ is convex and compact.*

Proof. Let ω_x^1 , $\omega_x^2 \in \partial f(x)$. Then

$$f(y) \geq f(x) + \omega_x^1(\dot{\gamma}_{xy}(o)), \ f(y) \geq f(x) + \omega_x^2(\dot{\gamma}_{xy}(o)), \forall\, y \in A, \ \forall\, \gamma_{xy} \in \Gamma.$$

Multiplying the first relation by $1-s$, the second by $s \in [0,1]$, and adding, we find

$$f(y) \geq f(x) + ((1-s)\omega_x^1 + s\,\omega_x^2)(\dot{\gamma}_{xy}(o)), \ \forall\, y \in A.$$

So

$$(1-s)\omega_x^1 + s\,\omega_x^2 \in \partial f(x)$$

and hence $\partial f(x)$ is convex.

We shall show that $\partial f(x)$ is closed and bounded, i.e., compact. Let $\{\omega_x^n\} \subset \partial f(x)$ be a sequence of subgradients convergent to a covector ω_x. Suppoose $\omega_x \notin \partial f(x)$, i.e., $\exists\, \varepsilon > 0$, $\exists\, z \in M$, $\exists\, \gamma_{xz}$, such that

$$f(z) + \varepsilon = f(x) + \omega_x(\dot{\gamma}_{xz}(0)).$$

On the other hand

$$f(z) \geq f(x) + \omega_x^n (\dot{\gamma}_{xz} (o)).$$

By subtraction we get

$$\varepsilon \leq (\omega_x - \omega_x^n)(\dot{\gamma}_{xz}(o)) \leq \|\omega_x - \omega_x^n\| \, \|\dot{\gamma}_{xz} (o))\|.$$

It follows

$$\left\|\omega_x - \omega_x^n\right\| \geq \frac{\varepsilon}{\left\|\dot{\gamma}_{xz}(o))\right\|} > 0, \ \forall \ n \in \mathbb{N},$$

which contradicts the fact that $\{\omega_x^n\}$ converges to ω_x. Hence $\omega_x \in \partial f(x)$ and $\partial f(x)$ is closed.

Let $\omega_x \in \partial f(x)$, i.e.,

$$\omega_x (\dot{\gamma}_{xy}(o)) + f(x) \leq f(y)$$

or

$$\|\omega_x\| \, \|\dot{\gamma}_{xy}(o)\| \cos \theta + f(x) \leq f(y), \ \forall \ y \in A, \ \forall \ \gamma_{xy} \in \Gamma .$$

It follows that $\|\omega_x\|$ is finite, and hence $\partial f(x)$ is bounded.

4.7. **Theorem**. *Let $f:A \longrightarrow \mathbb{R}$ be a convex function. $0_x \in \partial f(x)$ if and only if*

$$f(y) \geq f(x) + 0_x(\dot{\gamma}_{xy}(o)), \ \forall \ y \in A, \ \forall \ \gamma_{xy} \in \Gamma$$

i.e., if and only if x is a minimum point of f.

4.8. **Theorem**. *Let $f:A \longrightarrow \mathbb{R}$ be a convex function. $\omega_x \in \partial f(x)$ if and only if $Df(x;\, X_x) \geq \omega_x(X_x)$, $\forall \ X_x \in T_xA$.*

Proof. Suppose

$$\omega_x \in \partial f(x), \text{ i.e., } f(y) \geq f(x) + \omega_x(\dot{\gamma}_{xy}(o)), \ \forall \ y \in A, \ \forall \ \gamma_{xy} \in \Gamma .$$

Let $\gamma_{xy}(t)$, $t \in [0,1]$ be a geodesic such that $\dot{\gamma}_{xy}(o) = X_x$. Substituting y by $\gamma_{xy}(st)$, $s > o$, we find

$$\frac{f(\gamma_{xy}(st)) - f(x)}{s} \geq \omega_x(X_x),$$

and taking the limit as $s \longrightarrow o$, we conclude that

$$Df(x;\, X_x) \geq \omega_x(X_x).$$

Conversely, the relations

$$Df(x;\, X_x) \geq \omega_x(X_x), \quad f(y) - f(x) \geq \frac{f(\gamma_{xy}(t)) - f(x)}{t} \geq Df(x;\, X_x)$$

imply

$$f(y) - f(x) \geq \omega_x(X_x)$$

and hence

$\omega_x \in \partial f(x)$.

4.9. **Corollary**. *x is a minimum point of f if and only if $Df(x; X_x) \geq 0$, $\forall X_x \in T_xA$.*

4.10. **Theorem**. *Let A be an open totally convex set. The function $f: A \longrightarrow \mathbb{R}$ is convex if and only if whatever $x \in A$, there exists $\omega_x \in T^*_xA$ such that $\forall y \in A$, $\forall \gamma_{xy} \in \Gamma$ we have*

$$f(y) - f(x) \geq \omega_x(\dot{\gamma}_{xy}(o)).$$

Proof. Suppose that for each $x \in A$, there exists $\omega_x \in T^*_xA$ such that, $\forall y \in A$, $\forall \gamma_{xy} \in \Gamma$,

(5) $f(y) - f(x) \geq \omega_x(\dot{\gamma}_{xy}(o))$.

Changing y with x we obtain

(6) $f(x) - f(y) \geq \eta_y(\dot{\bar{\gamma}}_{yx}(o))$,

where $\bar{\gamma}_{yx}(t) = \gamma_{xy}(1 - t)$, $t \in [0,1]$ is a geodesic which joins y and x.

We fix t and so we obtain the point $\gamma_{xy}(t)$. The geodesic arc which joins $\gamma_{xy}(t)$ and y is the restriction $\gamma_{xy}(u)$, $u \in [t,1]$. Making $u = t + s(1-t)$, $s \in [0,1]$, we get the reparametrization

$$\alpha(s) = \gamma_{xy}(u(s)) = \gamma_{xy}(t + s(1-t)),\ s \in [0,1],$$

whence

$$\alpha(o) = \gamma_{xy}(t),\ \frac{d\alpha}{ds}(o) = (1-t)\,\dot{\gamma}_{xy}(t).$$

The restriction $\bar{\gamma}_{yx}(u) = \gamma_{xy}(1-u)$, $u \in [1-t, 1]$, is the geodesic arc which joins $\gamma_{xy}(t)$ and x. Making $u = 1-t + st$, $s \in [0,1]$, we find the reparametrization

$$\beta(s) = \bar{\gamma}_{yx}(1-t + st) = \gamma_{xy}(t-st),\ s \in [0,1],$$

whence

$$\beta(o) = \gamma_{xy}(t),\ \frac{d\beta}{ds}(o) = -\,t\,\dot{\gamma}_{xy}(t).$$

Replacing, in (5), x by $\gamma_{xy}(t)$ and $\dot{\gamma}_{xy}(o)$ by $\frac{d\alpha}{ds}(o)$ we obtain

(5′) $f(y) - f(\gamma_{xy}(t)) \geq (1-t)\omega_{\gamma_{xy}(t)}(\dot{\gamma}_{xy}(t))$.

Analogously, substituting, in (6), $\gamma_{xy}(t)$ for y and $\frac{d\beta}{ds}(o)$ for $\dot{\bar{\gamma}}_{yx}(o)$ we find

(6′) $f(x) - f(\gamma_{xy}(t)) \geq - t\, \omega_{\gamma_{xy}(t)}(\dot{\gamma}_{xy}(t))$.

Multiply (5′) by t and (6′) by (1-t). Adding we deduce

$$f(\gamma_{xy}(t)) \leq (1-t)f(x) + tf(y).$$

As t was arbitrarily choosen on [0,1], it follows that f is convex.

Conversely, the convexity of $f: A \longrightarrow \mathbb{R}$ and the Theorem 4.5 provide the existence of $\omega_x \in T_x^*A$ such that

$$f(y) - f(x) \geq \omega_x(\dot{\gamma}_{xy}(o)).$$

4.11. **Definition**. A function $f: A \longrightarrow \mathbb{R}$ that is both convex and concave is said to be *linear affine*.

A function $f:A \longrightarrow \mathbb{R}$ that is both convex and concave at $x \in A$ is said to be *linear affine at x*.Let $A_x \subset A$ be a convex normal neighborhood with its center at the point $x \in A$. The functions defined by

$$y \longrightarrow \omega_x(\dot{\gamma}_{xy}(0)),\ y \in A_x$$

are linear affine at x, i.e., they have the property that their restrictions to every geodesic radiating from x are linear affine.

4.12. **Corollary**. *A convex function on a convex normal neighborhood A_x with center x is the superior of a family of functions which are linear affine at x.*

Remark. In this context we give the correct version of a problem proposed by Hermann in [45]. Let $\exp_x : T_xM \longrightarrow M$ be the exponential map. Let O be an open subset of the tangent space T_xM, containing the element "zero", such that $\exp_x$ is a diffeomorphism between O and an open subset $A = \exp_x(O)$ of M. If O is a sufficiently small neighborhood of zero and $f:O \longrightarrow \mathbb{R}$ is convex (relative to the Euclidean linear connection in T_xM), then the restrictions of $f\circ\exp_x^{-1}: A \longrightarrow \mathbb{R}$ to every geodesic radiating from x are convex. Obviously, this does not mean that $f\circ\exp_x^{-1}$ is convex on A.

§5. Convexity of functions of class C^1

Let (M,g) be a complete finite-dimensional Riemannian manifold and let A be an open subset of M. If $f: A \longrightarrow \mathbb{R}$ is differentiable at $x \in A$, then all directional derivatives $Df(x; X_x)$ are bilateral and

$$Df(x; X_x) = g(\operatorname{grad} f(x), X_x) = df(X_x), \ \forall\, X_x \in T_xA.$$

We give here necessary and sufficient conditions for a function of class C^1 to be convex.

5.1. **Theorem.** *Let $A \subset M$ be an open totally convex set and $f: A \longrightarrow \mathbb{R}$ be a function of class C^1.*

1) The function f is convex if and only if

$$f(x) + \dot{\gamma}_{xy}(f)(x) \le f(y), \ \forall\, x,y \in A, \ \forall\, \gamma_{xy} \in \Gamma,$$

where Γ is the set f all geodesics joining the points x and y.

2) The function f is strictly convex on A if and only if

$$f(x) + \dot{\gamma}_{xy}(f)(x) < f(y), \ \forall\, x \ne y.$$

Proof. 1) From

$$f(\gamma_{xy}(t)) \le (1-t)f(x) + tf(y), \ 0 < t \le 1,$$

we find

$$f(x) + \frac{f(\gamma_{xy}(t)) - f(x)}{t} \le f(y).$$

Taking into account that

$$\lim_{t \searrow o} \frac{f(\gamma_{xy}(t)) - f(x)}{t} = \frac{d}{dt} f(\gamma_{xy}(t)) \Big|_{t=o} =$$

$$= \dot{\gamma}_{xy}(o)(f) = \dot{\gamma}_{xy}(f)(x)$$

we obtain the relation

$$f(x) + \dot{\gamma}_{xy}(f)(x) \le f(y) \tag{7}$$

(In order to obtain this relation it suffices that f be differentiable at the point x only).

Conversely, we suppose that $\forall\, x,y \in A$, $\forall\, \gamma_{xy} \in \Gamma$, the relation (7) is valid. Changing y with x we deduce

$$f(y) + \dot{\gamma}^{-}_{yx}(f)(y) \le f(x), \tag{8}$$

where

$$\gamma^{-}_{yx}(t) = \gamma_{xy}(1-t), \ t \in [0,1]$$

is a geodesic which joins y and x.

We fix t and so we obtain the point $\gamma_{xy}(t)$. The geodesic arc which joins $\gamma_{xy}(t)$ and y is the restriction $\gamma_{xy}(u)$, $u \in [t,1]$. Setting $u = t + s(1-t)$, $s \in [0,1]$, we find the reparametrization

$$\alpha(s) = \gamma_{xy}(u(s)) = \gamma_{xy}(t + s(1-t)),\ s \in [0,1],$$

whence

$$\alpha(o) = \gamma_{xy}(t),\ \frac{d\alpha}{ds}(o) = (1-t)\frac{d\gamma_{xy}}{dt}(t).$$

The restriction $\gamma^{-}_{yx}(u) = \gamma_{xy}(1-u)$, $u \in [1-t, 1]$, is the geodesic arc which joins $\gamma_{xy}(t)$ and x.Setting $u = 1-t + st$, $s \in [0,1]$, we find the reparametrization

$$\beta(s) = \gamma^{-}_{yx}(1-t + st) = \gamma(t-st),\ s \in [0,1],$$

whence

$$\beta(o) = \gamma(t),\ \frac{d\beta}{ds}(o) = -\ t\frac{d\gamma_{xy}}{dt}(t).$$

Replacing, in (7), x by $\gamma_{xy}(t)$ and $\dot{\gamma}_{xy}(o)$ by $\frac{d\alpha}{ds}(o)$ we get

$$(7') \qquad f(\gamma_{xy}(t)) + (1-t)\frac{d\gamma_{xy}}{dt}(f)(\gamma_{xy}(t)) \le f(y).$$

Analogously, substituting, in (8), $\gamma_{xy}(t)$ for y and $\frac{d\beta}{ds}(o)$ for $\dot{\gamma}^{-}_{yx}(o)$ we obtain

$$(8') \qquad f(\gamma_{xy}(t)) - t\frac{d\gamma_{xy}}{dt}(f)(\gamma_{xy}(t)) \le f(x).$$

Multiply (7') by t and (8') by (1-t). Adding we deduce

$$f(\gamma_{xy}(t)) \le (1-t)f(x) + tf(y).$$

As t was arbitrarily choosen on [0,1], it follows that f is convex.

Remarks. 1) If $f\colon A \longrightarrow \mathbb{R}$ is differentiable at x and convex on A, then $\omega_x = df(x)$ is a subgradient of f at x.

2) Let $f\colon A \longrightarrow \mathbb{R}$ be of class C^1 and $x,y \in A$, x being a fixed point. The function

$$y \longrightarrow f(x) + \dot{\gamma}_{xy}(f)(x) = f(x) + df(\dot{\gamma}_{xy}(o))$$

is called the *linear radial approximation* of f because it is of degree at most one with respect to the parameter t on every geodesic starting from the fixed point x, i.e., it is a linear affine function at x.

3) Let $f\colon A \longrightarrow \mathbb{R}$ be a submersion of class C^1. The graph

$$G(f) = \{(x,\ f(x)) \mid x \in A\} \subset A \times \mathbb{R}$$

is a hypersurface of $A \times \mathbb{R}$.

Let $x \in A$ be a fixed point. The graph G_* of the function

$$y \longrightarrow f(x) + df(\dot{\gamma}_{xy}(o))$$

is a totally geodesic hypersurface at x, tangent to G(f) at (x, f(x)). Theorem 5.1 shows that f is convex if and only if the totally geodesic hypersurface G_* tangent to G(f) lies below G(f). This remark produces a new geometrical definition for the convexity of a C^1 function.

5.2. **Theorem**. *Let $A \subset M$ be an open totally convex set, and $f: A \longrightarrow \mathbb{R}$ be a function of class C^1. If f is strictly convex on A, then $\forall\, x \neq y$, $\forall\, \gamma_{xy} \in \Gamma$ we have*

$$\dot{\gamma}_{xy}(f)(x) \neq \dot{\gamma}_{xy}(f)(y).$$

Proof. We proceed by reductio ad absurdum. Suppose there exist $x \neq y$ such that $\dot{\gamma}_{xy}(f)(x) = \dot{\gamma}_{xy}(f)(y)$. As f is strictly convex, we have

$$f(x) + \dot{\gamma}_{xy}(f)(x) < f(y), \quad f(y) - \dot{\gamma}_{xy}(f)(y) < f(x).$$

These relations imply f(x) < f(x) which is contradictory.

5.3. **Theorem**. *Let $A \subset M$ be an open totally convex set, and $f: A \longrightarrow \mathbb{R}$ be of class C^1. The function f is convex if and only if*

$$df(\dot{\gamma}_{xy})(x) - df(\dot{\gamma}_{xy}(y) \leq 0, \ \forall\, x,y \in A, \ \forall\, \gamma_{xy} \in \Gamma$$

(i.e., the differential df is a monotone mapping).

Proof. Let us suppose that f is convex. It follows that

$$f(x) + df(\dot{\gamma}_{xy})(x) \leq f(y), \quad f(y) - df(\dot{\gamma}_{xy})(y) \leq f(x).$$

Adding these relations term by term, we obtain the desired inequality.

Conversely, if

$$df(\dot{\gamma}_{xy})(x) - df(\dot{\gamma}_{xy})(y) \leq 0, \ \forall\, x,y \in A, \ \forall\, \gamma_{xy} \in \Gamma\ ,$$

the functions $\varphi_{xy}: [0,1] \longrightarrow \mathbb{R}$, $\varphi_{xy}(t) = f(\gamma_{xy}(t))$ have increasing derivatives,

$$\frac{d\varphi_{xy}}{dt}(t) = df(\dot{\gamma}_{xy})(\gamma_{xy}(t)).$$

Indeed, for $t_1 < t_2$, we have

$$(t_1-t_2)\Big(\frac{d\varphi_{xy}}{dt}(t_1) - \frac{d\varphi_{xy}}{dt}(t_2)\Big) = (t_1-t_2)\Big(df(\dot{\gamma}_{xy})(\gamma_{xy}(t_1)) -$$

$$- df(\dot{\gamma}_{xy})(\gamma_{xy}(t_2)\Big) = - \Big(df\Big(\frac{d\alpha_{pq}}{ds}\Big)(p) - df\Big(\frac{d\alpha_{pq}}{ds}\Big)(q)\Big) \geq 0,$$

where

$$\alpha_{pq}(s) = \gamma_{xy}(t_1 + s(t_2 - t_1)), \quad s \in [0,1],$$

and hence

$$\frac{d\alpha_{pq}}{ds}(o) = (t_2 - t_1)\,\dot{\gamma}_{xy}(t_1), \quad \frac{d\alpha_{pq}}{ds}(s) = (t_2 - t_1)\,\dot{\gamma}_{xy}(t_2).$$

Thus the functions φ_{xy} are convex and in particular

$$\varphi_{xy}(t) \le (1-t)\,\varphi_{xy}(o) + t\,\varphi_{xy}(1), \ \forall\, t \in [0,1], \ \forall\, x,y \in A,$$

or

$$f(\gamma_{xy}(t)) \le (1-t)f(x) + tf(y), \ \forall\, x,y \in A, \ \forall\, \gamma_{xy} \in \Gamma, \ \forall t \in [0,1].$$

§6. Convexity of functions of class C^2

Let (M,g) be a complete finite-dimensional Riemannian manifold.

Now we give necessary and sufficient conditions for a function of class C^2 to be convex.

6.1. **Theorem.** *Let $A \subset M$ be an open totally convex set and $f: A \longrightarrow \mathbb{R}$ be a function of class C^2. The function f is convex if and only if*

$$\varphi_{xy}(t) = f(\gamma_{xy}(t)), \ t \in [0,1]$$

satisfies

$$\frac{d^2\varphi_{xy}}{dt^2} \ge 0, \ \forall\, t \in [0,1], \quad \forall\, x,y \in A, \quad \forall\, \gamma_{xy} \in \Gamma,$$

where Γ is the set of all geodesics joining the points x and y.

Proof. Let

$$\frac{d^2\varphi_{xy}}{dt^2} \ge 0.$$

Taylor formula gives

$$\varphi_{xy}(o) - \varphi_{xy}(t) = \frac{d\varphi_{xy}}{dt}\Big|_t (-t) + \frac{1}{2}\frac{d^2\varphi_{xy}}{dt^2}\Big|_{t'} t^2, \quad t' \in [o,t],$$

$$\varphi_{xy}(1) - \varphi_{xy}(t) = \frac{d\varphi_{xy}}{dt}\Big|_t (1-t) + \frac{1}{2}\frac{d^2\varphi_{xy}}{dt^2}\Big|_{t''} (1-t)^2, \quad t'' \in [t,1].$$

Hence it follows

$$(1-t)(\varphi_{xy}(o) - \varphi_{xy}(t)) + t(\varphi_{xy}(1) - \varphi_{xy}(t)) =$$

$$= \frac{1}{2} \left.\frac{d^2\varphi_{xy}}{dt^2}\right|_{t'} (1-t)t^2 + \frac{1}{2} \left.\frac{d^2\varphi_{xy}}{dt^2}\right|_{t''} (1-t)^2 t \geq 0$$

or

$$f(\gamma_{xy}(t)) \leq (1-t)f(x) + tf(y) ,$$

and consequently f is convex.

For the converse we suppose that there exist $x,y \in M$, $\gamma_{xy} \in \Gamma$ such that $\frac{d^2\varphi_{xy}}{dt^2} \geq 0$ is not true on $[0,1]$. Therefore there exists at least one point on $[0,1]$ at which $\frac{d^2\varphi_{xy}}{dt^2}$ is < 0. Taking into account the first part of the proof it follows that $\varphi_{xy}(t)$ is not convex for $t \in [a,b] \subset \subset [0,1]$ and hence nor for $t \in [0,1]$. In other words, f cannot be convex, which is contradictory.

Remarks. 1) Let $\varphi\colon [0,1] \longrightarrow \mathbb{R}$ be a function of class C^2. The function φ is *convex* if and only if $\frac{d^2\varphi}{dt^2} \geq 0$ on $[0,1]$.

If $\frac{d^2\varphi}{dt^2} > 0$, $\forall\, t \in (0,1)$, then φ is *strictly convex*. The converse is not true.

2) If

$$\frac{d^2\varphi_{xy}}{dt^2} > 0,\ \forall\, x \neq y,\ \forall\, \gamma_{xy} \in \Gamma ,\ \forall\, t \in (0,1),$$

then f is *strictly convex*. The converse is not true.

Let $f\colon M \longrightarrow \mathbb{R}$ be a function of class C^2. We recall that the Hessian of f is defined by

$$\operatorname{Hess} f(X,Y) = \nabla_X(df)(Y),\ \forall\, X,Y \in \mathcal{X}(M).$$

6.2. **Theorem.** *Let $A \subset M$ be an open totally convex set and $f{:}A \longrightarrow \mathbb{R}$ be a function of class C^2. The function f is convex if and only if Hess f is positive semidefinite on A.*

Proof. Let $\varphi_{xy}(t) = f(\varphi_{xy}(t))$. As $\nabla_{\dot\gamma_{xy}} \dot\gamma_{xy} = 0$ we find

$$\frac{d^2\varphi_{xy}}{dt^2} = \dot{\gamma}_{xy}\left(\dot{\gamma}_{xy}(f)\right)(\gamma_{xy}(t)) = \nabla_{\dot{\gamma}_{xy}}(df)(\dot{\gamma}_{xy}) = \text{Hess } f(\dot{\gamma}_{xy}, \dot{\gamma}_{xy}).$$

Theorem 6.1 shows that f is convex if and only if

$$\text{Hess } f \geq 0.$$

Remarks. 1) The convexity property is invariant under nonlinear coordinate transformations.

2) If Hess f, then f is *strictly convex*. The converse is not true.

3) We observe that Hess f depends only on the function $f:A \longrightarrow \mathbb{R}$ and on the Riemannian connection ∇, but it does not require that A is totally convex. So we have a possibility to define the concept of convexity for a C^2 function on an open subset of a manifold M endowed with a torsion free linear connection: the function f is called *convex* if Hess $f \geq 0$; the function f is called *strictly convex* if Hess $f > 0$. Note, however, that the convexity property is linear connection dependent, such that a function f may be convex for a given torsion free linear connection on M, but not for another one.

4) The convexity of a C^2 function f with respect to a family of C^2 curves is equivalent to $\text{Hess}_{\nabla} f \geq 0$ if and only if the curves are geodesics attached to the torsion free linear connection ∇ .

5) If a C^2 function f is convex at x_o , then $\text{Hess } f(x_o) \geq 0$. The converse is not true.

6) A C^2 function f is *linear affine* if and only if Hess f = 0.

If a C^2 function f is linear affine at x_o , then $\text{Hess } f(x_o) = 0$. The converse is not true.

Convexity of Rosenbrock banana function

Let $(\mathbb{R}^2,\ g_o = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix})$ be the Euclidean plane. The function

$$F:\ \mathbb{R}^2 \longrightarrow \mathbb{R},\ F(y) = 100\ y_2^2 + y_1^2\ ,\ y = (y_1,y_2)$$

is convex with respect to g_o.

Now we consider the nonlinear coordinate transformation

$$\begin{cases} y_1 = 1 - x_1 \\ y_2 = x_2 - x_1^2 \end{cases} .$$

The Riemannian manifold $(\mathbb{R}^2,g_o)$ is changed into

$$(\mathbb{R}^2,\ g(x) = \begin{pmatrix} 4x_1^2 + 1 & -2x_1 \\ -2x_1 & 1 \end{pmatrix}),\quad x = (x_1, x_2)$$

and F is changed into *Rosenbrock banana function*

$$f\colon \mathbb{R}^2 \longrightarrow \mathbb{R},\ f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2,$$

which is convex with respect to g.

Open problem. *The existence of a function $f\colon M \longrightarrow \mathbb{R}$ whose Hessian is positive definite means the existence of a Riemannian metric h on M of the form $h = \mathrm{Hess}_g f$. What are the properties of the Riemannian manifold (M,h) ? (see Chapters 4 and 6).*

Let $X \in T_xM$, $\|X\| = 1$ and let $\gamma : (-a,a) \longrightarrow A \subset M$ be a geodesic such that $\gamma(o) = x$, $\dot{\gamma}(o) = X$. To a continuous function $f\colon A \longrightarrow \mathbb{R}$ we can attach the numbers

$$Cf(x;\ X) = \liminf_{t\to o} \frac{1}{t^2}\,[f(\gamma(t)) + f(\gamma(-t)) - 2f(\gamma(o))],$$

$$Cf(x) = \inf_{\|X\| = 1} Cf(x;\ X).$$

6.3. **Theorem**. *If A is an open totally convex set and $f\colon A \longrightarrow \mathbb{R}$ is of class C^2 at $x \in A$, then*

$$Cf(x;\ X) = (f\circ\gamma)''(o),\quad Cf(x) = \min_{\|X\| = 1} \mathrm{Hess}\, f(X,X).$$

Proof. A consequence of Taylor formulas:

$$(f\circ\gamma)(t) = (f\circ\gamma)(o) + t(f\circ\gamma)'(o) + \frac{t^2}{2}(f\circ\gamma)''(\xi_1),\quad \xi_1 \in (o,t)$$

$$(f\circ\gamma)(-t) = (f\circ\gamma)(o) - t(f\circ\gamma)'(o) + \frac{t^2}{2}(f\circ\gamma)''(\xi_2),\quad \xi_2 \in (-t,o).$$

Remark. Suppose it is to be shown that a continuous real function $f(x)$ satisfies $Cf(x_o) \geq 0$. Then it suffices to produce a C^2 function $\bar{f}(x)$ such that $\bar{f}(x) \leq f(x)$ near x_o and $\bar{f}(x_o) = f(x_o)$, and such that $\mathrm{Hess}\,\bar{f}(x_o) \geq 0$. The main point of this procedure is to sidestep arguments involving continuous functions by working only with differentiable functions (see Chapter 4, §4).

Examples on the sphere S^2

We consider the sphere

$$S^2\colon x^2 + y^2 + z^2 = 1$$

and the poles $P(o,o,-1)$, $Q(o,o,1)$. The set $S^2 - \{P,Q\}$ can be parametrized by

$$\begin{cases} x = \sin\theta\cos\varphi \\ y = \sin\theta\sin\varphi \\ z = \cos\theta \qquad , \ \theta \in (0,\pi),\ \varphi \in [0,2\pi]. \end{cases}$$

The Riemannian metric on S^2-{P,Q} induced by the Riemannian metric on $\mathbb{R}^3$ is given by

$$ds^2 = dx^2 + dy^2 + dz^2 = d\theta^2 + \sin^2\theta\, d\varphi^2.$$

Hence

$$g_{11} = 1,\ g_{12} = 0,\ g_{22} = \sin^2\theta,$$

$$g^{11} = 1,\ g^{12} = 0,\ g^{22} = \frac{1}{\sin^2\theta}$$

and the components of the Riemannian connection are

$$\Gamma^1_{11} = \Gamma^2_{11} = \Gamma^2_{22} = \Gamma^1_{12} = \Gamma^1_{21} = 0,$$

$$\Gamma^1_{22} = -\sin\theta\cos\theta,\ \Gamma^2_{12} = \Gamma^2_{21} = \text{cotan}\theta.$$

The geodesics of S^2-{P,Q} are great circles or semicircles.

We consider a function f: S^2-{P,Q} $\longrightarrow \mathbb{R}$ of class C^2 and denote by f_{ij}, i,j = 1,2, the components of Hess_g f. Then

$$f_{11} = \frac{\partial^2 f}{\partial\theta^2},\ f_{12} = \frac{\partial^2 f}{\partial\theta\partial\varphi} - \frac{\partial f}{\partial\varphi}\text{cotan}\,\theta,\ f_{22} = \frac{\partial^2 f}{\partial\varphi^2} + \frac{\partial f}{\partial\theta}\sin\theta\cos\theta.$$

Let us find solutions for the next system of inequations and equations with partial derivatives

$$f_{11} > 0,\quad f_{12} = 0,\quad f_{22} > 0.$$

The general solution of the equation $f_{12} = 0$ is

$$f(\theta,\varphi) = a(\varphi)\sin\theta + b(\theta),\ \theta \in (0,\pi),\ \varphi \in [0,2\pi],$$

where a and b are functions of class C^2.

Let the semicircle γ: $\varphi = 0$, $\theta \in (0,\pi)$. Let $\{U_i,\ i\in I\}$ be a local finite covering of S^2- $\{P,Q,\gamma\}$ such that every $\bar{U}_i$, $i\in I$ to be compact. For

$$f_{\bar{U}_i}(\theta,\varphi) = [-1 + (1 - 1/n)\exp\frac{\varphi}{n}]\sin\theta$$

and $n \in \mathbb{N}^*$ sufficiently large (depending on $\bar{U}_i$) one verifies

$$f_{11}\big|_{\bar{U}_i} > 0,\quad f_{22}\big|_{\bar{U}_i} > 0.$$

Let $\{h_i,\ i\in I\}$ be a partition of unity subordinate to the above covering. The function

$$f(\theta,\varphi) = \sum_{i\in I} h_i(\theta,\varphi)\, f_{\overline{U}_i}(\theta,\varphi)$$

is a solution of

$$f_{11} > 0,\ f_{12} = 0,\ f_{22} > 0$$

on $S^2-\{P,Q,\gamma\}$. Consequently f is a strictly convex function and $\mathrm{Hess}_g f$ is a Riemannian metric on $S^2-\{P,Q,\gamma\}$. Obviously $S^2-\{P,Q,\gamma\}$ is not a totally convex subset of S^2.

Examples on Poincaré plane

The set $H = \{(x,y) \in \mathbb{R}^2 | y > 0\}$ endowed with the Riemannian metric $g_{ij}(x,y) = \frac{1}{y^2}\delta_{ij}$, $i,j = 1,2$, is called the *Poincaré plane (Hyperbolic plane)*. The Riemannian connection on H has the following components

$$\Gamma^1_{11} = \Gamma^1_{22} = \Gamma^2_{12} = \Gamma^2_{21} = 0,\quad \Gamma^1_{12} = \Gamma^1_{21} = \Gamma^2_{22} = -\frac{1}{y},\quad \Gamma^2_{11} = \frac{1}{y}.$$

If $f\colon H \longrightarrow \mathbb{R}$ is of class C^2, the *hyperbolic Hessian* of f has the components

$$f_{11} = \frac{\partial^2 f}{\partial x^2} - \frac{1}{y}\frac{\partial f}{\partial y},\quad f_{12} = \frac{\partial^2 f}{\partial x\partial y} + \frac{1}{y}\frac{\partial f}{\partial x},\quad f_{22} = \frac{\partial^2 f}{\partial y^2} + \frac{1}{y}\frac{\partial f}{\partial y}.$$

The geodesics of the Poincaré plane are the semilines $C_a\colon x = a$, $y > 0$ and the semicircles $C_{b,r}\colon (x-b)^2 + y^2 = r^2$, $y > 0$. They admit the following natural parametrizations

$$C_{x_o}\colon x = x_o,\ y = y_o e^t,\ t \in (-\infty,\infty)$$

$$C_{b,r}\colon x = b - r\tanh t,\ y = \frac{r}{\cosh t},\ t \in (-\infty,\infty).$$

The *hyperbolic distance* between the points $P_1 = (x_1,y_1)$ and $P_2 = (x_2,y_2)$ is

$$d_H(P_1,P_2) = \begin{cases} \left|\ln \dfrac{y_2}{y_1}\right| & \text{for } x_1 = x_2 \\[2ex] \left|\ln \dfrac{x_1 - b + r}{x_2 - b + r}\ \dfrac{y_2}{y_1}\right| & \text{for } P_1,P_2 \in C_{b,r}. \end{cases}$$

First we want to find solutions for the system

$$f_{11} > 0,\ f_{12} = 0,\ f_{22} > 0.$$

The general solution of the equation $f_{12} = 0$ is

$$f(x,y) = \frac{1}{y}\left(\varphi(x) + \int y\,\psi(y)dy\right),$$

where φ is a function of class C^2 and ψ is a function of class C^1. For

$\varphi(x) = 0$, $\psi(y) = (\alpha + 1)y^{\alpha-1}$, $\alpha < 0$ we obtain $f(x,y) = y^{\alpha}$ which satisfies $f_{11} > 0$, $f_{22} > 0$.

The perpendicular from (x,y) to Oy is the geodesic $x^2 + y^2 = a^2$, $y > 0$. Let us show that $f\colon H \longrightarrow \mathbb{R}$, $f(x,y) = \ln^2 \frac{x + a}{y}$, the square of the distance from $P_1 = (x,y) \in H$ to the vertical geodesic Oy, is a strictly convex function.

Indeed, on the geodesic C_a we find

$$\varphi(t) = \ln^2 \frac{x_o + a}{y_o e^t} = (\ln \alpha - t)^2, \quad \alpha = \frac{x_o + a}{y_o}$$

and hence

$$\varphi'(t) = -2(\ln \alpha - t), \quad \varphi''(t) = 2 > 0, \ \forall\, t \in \mathbb{R}.$$

On the geodesic $C_{b,r}$ we obtain

$$\varphi(t) = \ln^2(\beta \cosh t - \sinh t), \quad \beta = \frac{b + a}{r_1}.$$

The existence condition $\beta \cosh t - \sinh t > 0$, and $\sup\limits_t(\tanh t) = 1$ imply $\beta \geq 1$. We cannot have $\beta \cosh t - \sinh t \leq 1$, because in this case

$$\beta \leq \frac{1 + \sinh t}{\cosh t} = \psi(t), \qquad \psi'(t) = \frac{1 - \sinh t}{\cosh^2 t},$$

t	$-\infty$		t_o		∞
ψ'		+	0	−	
ψ	-1	$\nearrow$	$\sqrt{2}$	$\searrow$	1

and hence $\beta \leq -1$, a contradiction. It remains $\beta \cosh t - \sinh t > 1$ and consequently $\beta \geq \sqrt{2}$. These ensure

$$\varphi'(t) = 2\,\frac{\beta \sinh t - \cosh t}{\beta \cosh t - \sinh t}\, 2\ln(\beta \cosh t - \sinh t),$$

$$\varphi''(t) = 2\left(\frac{\beta \sinh t - \cosh t}{\beta \cosh t - \sinh t}\right)^2 +$$

$$+\, 2\,\frac{\beta^2 - 1}{(\beta \cosh t - \sinh t)^2}\, \ln(\beta \cosh t - \sinh t) > 0, \forall\, t \in \mathbb{R}.$$

Remark. There exist *posynomial functions*

$$f(x,y) = \sum_{i=1}^{m} c_i x^{a_{i1}} y^{a_{i2}}, \ x > 0,\ y > 0,\ c_i > 0,\ a_{ij} \in \mathbb{R},$$

which are convex on the Poincaré plane.

Linear affine functions

Let (M,g) be a complete n-dimensional Riemannian manifold. We may assume without loss of generality that we use only C^∞ real functions f on M. Recall that such a function is called *linear affine* if $\mathrm{Hess}_g f = 0$.

In other words, f is *linear affine* if and only if the vector field grad f is parallel. If grad f is parallel and nowhere zero, then the hypersurfaces Σ_c: f(x) = c are totally geodesic.

6.4. **Theorem** [20]. *(M,g_{ij}) admits a family of totally geodesic hypersurfaces if and only if $g_{ij}(x^1,\dots,x^n)$ reduces to $Q_{\alpha\beta}(y^1,\dots,y^{n-1})$, $\alpha,\beta = 1,\dots,n-1$ and $Q_{nn}(y^1,\dots,y^n)$.*

6.5. **Choquet theorem.** (Private communication, 1981) *An n-dimensional Riemannian manifold (M,g) is the Riemannian product of an (n-p+1)-dimensional Riemannian manifold and the Euclidean space $\mathbb{R}^{p-1}$ (locally at least) if and only if the vector space of all linear affine functions on M has dimension p.*

Proof. Let (U,x^i) be a coordinate neighborhhod and Γ^h_{ij}, i,j,h = = 1,...,n, be the components of the Riemannian connection determined by the components g_{ij} of the Riemannian metric g. Assume that

$$f_o = \text{const} \neq 0,\ f_1,\dots,f_{p-1}$$

are (linearly independent) linear affine functions on M, i.e.,

$$\frac{\partial^2 f_\alpha}{\partial x^i \partial x^j} - \Gamma^h_{ij}\frac{\partial f_\alpha}{\partial x^h} = 0,\ \alpha = 0,1,\dots,p-1 .$$

This hypothesis is equivalent to the fact that

$$\text{grad } f_1,\dots,\text{grad } f_{p-1}$$

are p-1 nonzero parallel vector fields. By changing the coordinates [20] it is proved that the metric $g = g_{ij}dx^i \otimes dx^j$ can be written in the form

$$g = g_{\alpha'\beta'}dx^{\alpha'} \otimes dx^{\beta'} + dx^{(n-p+2)'} \otimes dx^{(n-p+2)'} + \dots + dx^{n'} \otimes dx^{n'},$$

$$\alpha',\beta' = 1,\dots,n-p+1,$$

and hence (M,g) is the Riemannian product between

$$(M_1, g_1 = g_{\alpha'\beta'}dx^{\alpha'} \otimes dx^{\beta'}),$$

$$(\mathbb{R}^{p-1}, g_2 = dx^{(n-p+2)'} \otimes dx^{(n-p+2)'} + \dots + dx^{n'} \otimes dx^{n'}).$$

Conversely, suppose that (M,g) is the specified Riemannian product. The components $\Gamma^{h'}_{i'j'}$, $i',j',h' = 1,\dots,n$ of the Riemannian connection determined by $g = g_1 + g_2$ are all zero excepting (maybe) $\Gamma^{\alpha'}_{\beta'\gamma'} = \bar{\Gamma}^{\alpha'}_{\beta'\gamma'}$, $\alpha',\beta',\gamma' = 1,\dots,n-p+1$, where $\bar{\Gamma}^{\alpha'}_{\beta'\gamma'}$ are the Christoffel symbols of g_1. In this case the coordinate functions

$$x^{(n-p+2)'}, \dots, x^{n'}$$

prove to be linear affine. Taking into account that a nonzero constant is also a linear affine function, it follows that the vector space of all linear affine functions on M has dimension p.

6.6. **Theorem**. *Let (M,g) be an n-dimensional Riemannian manifold and $f:M \longrightarrow \mathbb{R}$ be a function of class C^{∞}. If f has no critical points, then there exists a linear symmetric connection $\bar{\Gamma}^h_{ij}$, $i,j,h = 1,\dots,n$ on M such that f is linear affine with respect to $\bar{\Gamma}$, i.e., $\mathrm{Hess}_{\bar{\Gamma}}\, f = 0$.*

Proof. Being given f, we consider the algebraic system

$$\frac{\partial^2 f}{\partial x^i \partial x^j} - \bar{\Gamma}^h_{ij} \frac{\partial f}{\partial x^h} = 0, \quad i,j,h = 1,\dots,n,$$

with the unknowns $\bar{\Gamma}^h_{ij}$. In order to obtain a solution which is necessarily a linear connection we look for

$$\bar{\Gamma}^h_{ij} = \Gamma^h_{ij} + T^h_{ij},$$

where Γ^h_{ij} is the Riemannian connection determined by g_{ij}, and T^h_{ij} is an arbitrary tensor field of type (1,2), symmetric with respect to i and j. Then the preceding system is transcribed in the tensorial form

$$T^h_{ij} f_h = f_{ij}, \quad f_h = \frac{\partial f}{\partial x^h},$$

where f_{ij} are the components of the $\mathrm{Hess}_g\, f$, and T^h_{ij} are the new unknowns. The general solution of this system is

$$T^h_{ij} = \frac{f_{ij} f^h}{\|\mathrm{grad}\, f\|^2} + S^h_{ij}, \quad f^h = g^{hj} \frac{\partial f}{\partial x^j}, \quad \|\mathrm{grad}\, f\|^2 = g^{ij} f_i f_j,$$

where S^h_{ij} is a tensor field of type (1,2), symmetric with respect to i and j, and which satisfies

$$S^h_{ij} f_h = 0.$$

Consequently

$$\bar{\Gamma}^h_{ij} = \Gamma^h_{ij} + \frac{f_{ij}\, f^h}{\|\mathrm{grad}\ f\|^2} + S^h_{ij}$$

is the most general linear symmetric connection on M which makes f a linear affine function.

§7. Convex programs on Riemannian manifolds

Many of the properties of convex programs on Euclidean space carry over to the case of a Riemannian manifold (or more generally to the case of a manifold with a torsion-free linear connection) because they are independent of the particular Riemannian metric (torsion-free linear connection) which yields convexity. All that we have to do is to find suitable forms and proofs.

Let M be a finite-dimensional differentiable manifold, let A be a subset of M and $f\colon A \longrightarrow \mathbb{R}$ be a real function.

7.1. **Definition**. A program of type $\min\limits_{x\in A} f(x)$ is called *convex* if there exists a Riemannian metric g on M such that the Riemannian manifold (M,g) is complete, the set A is totally convex in (M,g) and f is a convex function.

Having in mind this definition, in the sequel we suppose that (M,g) is a complete Riemannian manifold.

7.2. **Theorem**. *Let $x_0 \in A \subset M$ be a local minimum point of $f\colon A \longrightarrow \mathbb{R}$. If A is star-shaped at x_0 and f is convex at x_0, then x_0 is a global minimum point.*

Proof. By hypothesis there exists $\varepsilon > 0$ such that

$$f(x_0) \le f(x),\ \forall\ x \in B(x_0,\varepsilon)\cap A.$$

Suppose $z \in A$ and $f(z) < f(x_0)$. Consider $\gamma_{x_0z}(t)$, $t \in (0,1)$. Since A is star-shaped at x_0 and f is convex at x_0, it follows

$$f(\gamma_{x_0z}(t)) \le (1 - t)f(x_0) + tf(z) < f(x_0).$$

But

$$x = \gamma_{x_0z}(t) \in B(x_0,\varepsilon) \cap A$$

for some $t \in (0,1)$ and hence $f(x) \ge f(x_0)$. This is a contradiction, and hence

$$f(z) \ge f(x_0)\ ,\ \forall\ z \in A.$$

This means that any local minimum point of a convex function is also a global minimum point. Since each global minimum point is a local minimum point, it follows that for a convex function the set of local minimum points and the set of global minimum points coincides. Let x_1 and x_2 be two minimum points of the convex function f. Since both are global minimum points, we deduce $f(x_1) \le f(x_2)$, $f(x_2) \le f(x_1)$ and hence $f(x_1) = f(x_2)$; so f has the same minimum value on A.

7.3. **Theorem**. *The set of minimum points for a convex program* $\min_{x \in A} f(x)$ *is totally convex.*

Proof. If f has no minimum value on A, then the set of minimum points is empty and hence totally convex. If f has a minimum value m on A, then the set of minimum points is the intersection of two totally convex sets

$$A \cap \{ z \in A,\ f(z) \le m\}$$

and hence it is totally convex.

7.4. **Corollary**. *If the set of minimum points of a convex program has at least two distinct points, then it has an infinity and the function that we minimize is not strictly convex.*

Proof. If x and y are minimum points, then according to the preceding theorem each point of a geodesic arc which joins x and y is a minimum point of f. Moreover, under these conditions,

$$f(\gamma_{xy}(t)) = (1-t)f(x) + tf(y) = f(x) = f(y),\ \forall\ t \in [0,1].$$

7.5. **Theorem**. *Let* $A \subset M$ *be an open totally convex set and* $f:A \longrightarrow \mathbb{R}$ *be a convex function of class* C^1. *If x is a critical point of f, then x is a minimum point of A (It is sufficient that f be differentiable only at the point x).*

Proof. The convexity implies

$$f(x) + df(\dot{\gamma}_{xy})(x) \le f(y).$$

Since x is a critical point of f, i.e., $df(x) = 0$, we find

$$f(x) \le f(y),\ \forall\ y \in A.$$

The converse of this theorem is not true.

7.6. **Theorem**. *Let* $A \subset M$ *be a subset with nonvoid interior. If a convex function* $f: A \longrightarrow \mathbb{R}$ *has a global maximum point in the interior of A, then f is constant.*

Proof. Suppose that x' is a global maximum point for f over A, i.e., $f(y) \le f(x')$, $\forall\ y \in A$, and that x' is an interior point of A. Choose $x \in A$ so that x' be an interior point of a geodesic which joins x with y. This means that $\exists\ s \in (0,1)$ such that $\gamma_{xy}(s) = x'$. It follows

$$f(x') \le (1-s)f(x) + sf(y).$$

Since x' is a global maximum point we have $f(x) \le f(x')$. If $f(y) < f(x')$, then

$$f(x') < (1-s)f(x') + sf(x') = f(x'),$$

which is a contradiction. It remains $f(y) = f(x')$, $\forall\ y \in A$.

Remarks. 1) If the convex function f is not constant, then all global maximum points lie on the boundary of the totally convex set A. Moreover, no such point can be a critical point (critical points are global minimum points).

2) If a point in a set A is both a global minimum point and a global maximum point for $f: A \longrightarrow \mathbb{R}$, then f is constant.

Open problem. Geometric programming is a set of techniques developed for solving non-linear optimization problems.These techniques have found applications in a surprisingly wide variety of fields: gas transmission compressor design, marketing, design of welded beam structures, chemical engineering processes, cofferdam and pressure vessel design, ship operation and design, modular design, nuclear systems, etc.[See R.Duffin, E.Peterson, C.Zener, Geometric programming; theory and applications, John Wiley & Sons, New York, 1967; C.Beightler, D.Phillips, Applied geometric programming, John Wiley & Sons, 1976]. In this context, we suggest the following problems:

1) *Determine significant geometric programs which are convex on the Poincare manifold*

$$\left(\{(x,y) \in \mathbb{R}^n,\ x \in \mathbb{R}^{n-1},\ y > 0\},\ g_{ij}(x,y) = \frac{1}{y^2}\,\delta_{ij}\right).$$

Hint. Any geometric program which refers to *sygnomials* of the form

$$f(x,y)=y^{-1}\sum_{i=1}^{m} c_i x^{a_{i1}} + \sum_{i=1}^{m} \frac{c_i}{a_{i2}+2}\, y^{a_{i2}+1}\ ,\ x > 0,\ y > 0,\ c_i \in \mathbb{R}\ ,$$

is a convex program on the Poincare plane (see §6).

2) *Find Riemannian metrics which change relevant geometric programs into convex programs.*

§8. Duality in convex programming

Let (M,g) be a complete finite-dimensional Riemannian manifold. In the sequel we refer to the *convex program*

$$\min_{x\in A} f(x), \tag{9}$$

where $f: M \longrightarrow \mathbb{R}$ is the *convex objective function*, and the totally convex subset A is described by the systems of inequalities $\psi_\imath(x) \geq 0$, $\imath = 1,\dots,r$, where $\psi_\imath : M \longrightarrow \mathbb{R}$ are concave functions. A is called the *set of admissible solutions*. The program is called *consistent* if $A \neq \emptyset$ and *superconsistent* if $\operatorname{int} A \neq \emptyset$, i.e., there exists $y \in M$ such that $\psi_\imath(y) > 0$, $\imath = 1,\dots,r$. For $x_o \in A$ we denote by $I(x_o)$ the set of indices $\imath$ having the property that the inequalities which describe A are active at x_o, i.e.,

$$I(x_o) = \{ \imath \mid \psi_\imath(x_o) = 0\} .$$

8.1. **Lemma**. *If the convex program (9) is superconsistent and the function $\psi_\imath$ are of class C^1, then the vectors $\operatorname{grad} \psi_\imath$, $\imath \in I(x_o)$ are positively linearly independent.*

Proof. By hypothesis there exist $y \in M$ such that

$$\psi_\imath(y) > 0, \quad \imath = 1,\dots,r,$$

and the convexity of the functions $-\psi_\imath$ implies

$$0 < \psi_\imath(y) = \psi_\imath(y) - \psi_\imath(x_o) \leq d\psi_\imath(\dot\gamma_{x_oy}(o)), \quad \imath \in I(x_o),$$

where $\gamma_{x_oy}(t)$, $t \in [0,1]$ is a geodesic from x_o to y.

Suppose that there exist $v_\imath \geq 0$, $\imath \in I(x_o)$,not all zero, such that

$$\sum_{\imath\in I(x_o)} v_\imath \, d\psi_\imath(x_o) = 0.$$

It follows

$$\sum_{\imath\in I(x_o)} v_\imath \, d\psi_\imath(\dot\gamma_{x_oy}(o)) = 0$$

and the positivity of each term of the sum implies

$$v_\imath \, d\psi_\imath(\dot\gamma_{x_oy}(o)) = 0.$$

This relation, and the fact that the second factor is strictly positive give $v_\imath = 0$ for all $\imath \in I(x_o)$, which contradicts the hypothesis. It remains that

$$\operatorname{grad} \psi_\imath(x_o), \quad \imath \in I(x_o)$$

are positively linearly independent.

The convex program (9) is called *primal problem*. The function defined by

$$L(x,v) = f(x) - \sum_{\imath=1}^{r} v^\imath \psi_\imath(x), \quad x \in M, \quad v^\imath \geq 0, \quad \imath = 1,\dots,r$$

is called the *Lagrange function* attached to the primal problem. The program

(10) $$\max L(x,v)$$

with constraints

$$x \in A, \; v = (v^1,\dots,v^r) \in \mathbb{R}^r_+ ,$$

$$\operatorname{grad} f(x) = \sum_{\imath=1}^{r} v^\imath \operatorname{grad} \psi_\imath(x)$$

is called *dual problem*.

8.2. **Duality theorem**. *Suppose that the convex program (9) is superconsistent and the functions f and $\psi_\imath$ are of class C^1. If x_o is the optimal solution of the primal problem (9), then there exists $v_o \in \mathbb{R}^r_+$ such that (x_o, v_o) is the optimal solution of dual problem (10) and $f(x_o) = L(x_o,v_o)$.*

Proof. Let x_o be the solution of the primal problem. Lemma 8.1 shows that the vectors $\operatorname{grad} \psi_\imath(x_o)$, $\imath \in I(x_o)$ are positively linearly independent. Therefore, Fritz John Theorem implies the existence of numbers v_o for which

1) $v_o^\imath \geq 0$, $\imath = 1,\dots,r$; $v_o^\imath = 0$ for $\imath \notin I(x_o)$, i.e., $v_o^\imath \psi_\imath(x_o) = 0$,

2) $\operatorname{grad} f(x_o) = \sum_{\imath=1}^{r} v_o^\imath \operatorname{grad} \psi_\imath(x_o)$.

So (x_o, v_o) is an admissible solution of the dual problem (10). The definition of the set $I(x_o)$ shows that for any $v \in \mathbb{R}^r_+$, the relations

$$L(x_o,v) = f(x_o) - \sum_{\imath=1}^{r} v^\imath \psi_\imath(x_o) = f(x_o) - \sum_{\imath \notin I(x_o)} v^\imath \psi_\imath(x_o) \leq$$

$$\leq f(x_o) = L(x_o,v_o) = f(x_o) - \sum_{\imath=1}^{r} v_o^\imath \psi_\imath(x_o)$$

are satisfied.

Let us now suppose that (x,v) is any admissible solution of the dual problem (10). Since the function $x \longrightarrow L(x,v)$ is convex, the point x, which satisfies the condition

$$\operatorname{grad} f(x) - \sum_{l=1}^{r} v^l \operatorname{grad} \psi_l(x_o) = 0,$$

is a global minimum point. Hence

$$L(x,v) \le L(x_o,v) \le L(x_o,v_o) \ , \tag{11}$$

for the arbitrary admissible solution (x,v) of the dual problem. It follows that (x_o,v_o) is the optimal solution of the dual problem.

§9. Kuhn - Tucker theorem on Riemannian manifolds

The solution of a convex programming problem is completely characterized by the saddle point theorem which has initially been stated on $\mathbb{R}^n$ by Kuhn and Tucker [56].

9.1. **Kuhn-Tucker theorem**. *Suppose the convex program (9) is superconsistent and the functions* f *and* ψ_l *are of class* C^1. *A point* $x_o \in A \subset M$ *is the optimal solution of the primal problem (9) if and only if there exists* $v_o = (v_o^1, \ldots, v_o^r)$ *such that*

1) $v_o^l \ge 0,\ l = 1,\ldots,r\ ;\quad v_o^l\, \psi_l(x_o) = 0,\ l = 1,\ldots,r$

2) $L(x_o,v) \le L(x_o,v_o) \le L(x,v_o),\ \forall\, x \in M,\ v \in \mathbb{R}^r_+$.

Proof. Let us suppose $x_o \in A$ is the optimal solution of the primal problem (9). The Duality Theorem shows that there exists $v_o \ge 0$ such that (x_o,v_o) verifies the system

$$v^l \ge 0,\ l = 1,\ldots,r\ ;\quad v^l\, \psi_l(x) = 0,\ l = 1,\ldots,r$$

$$\operatorname{grad} f(x) - \sum_{l=1}^{r} v^l \operatorname{grad} \psi_l(x) = 0.$$

Hence (x_o,v_o) verifies the conditions 1). The relations 2) follow from (11) and from the fact that, $x \longrightarrow L(x,v_o)$ being convex, x_o is a minimum point (critical point), i.e.,

$$L(x_o,x_o) \le L(x,v_o) = f(x) - \sum_{l=1}^{r} v_o^l\, \psi_l(x),\ \forall\ x \in M.$$

Let us now suppose that $(x,v) \in M \times \mathbb{R}^r$ verifies the Kuhn-Tucker conditions 1) - 2). The implication

$$\forall\ v \ge 0 \Rightarrow L(x_o,v) - L(x_o,v_o) \le 0$$

is the same as

$$\forall\ v \ge 0 \Rightarrow \sum_{l=1}^{r} (v_o^l - v^l)\psi_l(x_o) \le 0.$$

Taking succesively

$$v^l = v_o^l + 1,\ v^m = v_o^m\ ,\ m \ne l,$$

we find

$$\psi_l(x_o) \ge 0,\ l = 1,\dots,r$$

and hence $x_o \in A$.

On the other hand, the implication

$$\forall\ x \in M \Rightarrow L(x_o,v_o) \le L(x\ ,v_o)$$

is equivalent to

$$\forall\ x \in M \Rightarrow f(x_o) - f(x) - \sum_{l=1}^{r} v_o^l\ (\psi_l(x_o) - \psi_l(x\ )) \le 0.$$

Considering $v_o^l\ \psi_l(x_o) = 0$, we infer

$$f(x_o) - f(x) \le - \sum_{l=1}^{r} v_o^l\ \psi_l(x).$$

If $x \in A$, i.e., $\psi_l(x) \ge 0$, then $f(x_o) - f(x) \le 0$. So x_o is the optimal solution of the primal problem (9).

Remark. In [58] is given an example which shows that the hypothesis of superconsistency cannot be eliminated.

In order to obtain another variant for the dual problem, we focus again on the primal problem (9). Fix $v^l \ge 0$, $l = 1,\dots,r$ and denote

$$\varphi(v) = \inf_{x\in A} \left[f(x) - \sum_{l=1}^{r} v^l\ \psi_l(x)\right].$$

The function $v \longrightarrow \varphi(v)$, $v \in \mathbb{R}_+^r$ is well defined, it is concave and can take the value $-\infty$.

9.2. **Theorem**. *For any* $v \in \mathbb{R}_+^r$ *and any* $x \in A$, *the relation*

$\varphi(v) \le f(x)$ is satisfied. If the conditions of Kuhn-Tucker theorem are valid, then

$$\max_{v\in\mathbb{R}^r_+} \varphi(v) = \min_{x\in A} f(x).$$

Proof. For $x \in A$, $v \in \mathbb{R}^r_+$ we have

$$\varphi(v) \le f(x) - \sum_{l=1}^{r} v^l \psi_l(x) \le f(x).$$

Let us now suppose that the conditions of Theorem 9.1 are satisfied, i.e., there exists $v_o \in \mathbb{R}^r_+$ such that the conditions 1) - 2) of the theorem are satisfied. These relations imply

$$\varphi(v_o) = f(x_o) - \sum_{l=1}^{r} v_o^l \psi_l(x_o) = f(x_o) .$$

Since $\varphi(v) \le f(x_o) = \varphi(v_o)$, it follows that v_o is a maximum point of φ on $\mathbb{R}^r_+$ and

$$\max_{v\in\mathbb{R}^r_+} \varphi(v) = \varphi(v_o) = f(x_o) = \min_{x\in A} f(x).$$

The problem of the maximization of the function φ on $\mathbb{R}^r_+$ is sometimes called the *dual problem* of the convex programming and v is called the *vector of dual variables*.

Remark. This context allows the following paraphrase: in the conditions of the Kuhn-Tucker Theorem the maximum value of the objective function in the dual problem is the same as the minimum value of the objective function in the primal problem; the Lagrange multipliers of the primal problem represent at the same time the solution of the dual problem.

§10. Quasiconvex functions on Riemannian manifolds

Let (M,g) be a complete finite-dimensinal Riemannian manifold, let $A \subset M$ be a totally convex set, and $\gamma_{xy} : [0,1] \longrightarrow M$ be a geodesic which joins the points x and y, i.e., $\gamma_{xy}(0) = x$, $\gamma_{xy}(1) = y$.

10.1. **Definition**. The function $f: A \longrightarrow \mathbb{R}$ is called *quasiconvex* if

$$f(\gamma_{xy}(t)) \le \max \{f(x), f(y)\}$$

whenever $x, y \in A$ and $t \in (0,1)$.

When the preceding inequality is strict, for $x \neq y$ and $t \in (0,1)$ the function f is said to be *strongly quasiconvex*.

The function $f:A \longrightarrow \mathbb{R}$ is called *quasiconcave* if $-f$ is quasi-convex. Since

$$\max \{-f(x), -f(y)\} = - \min\{f(x), f(y)\}$$

it follows that f is quasiconcave if and only if

$$f(\gamma_{xy}(t)) \geq \min \{f(x), f(y)\}.$$

Any convex function is quasiconvex. The sublevel sets

$$A^c = \{x \in A \mid f(x) \leq c\}$$

attached to the convex function f are totally convex. The converse of these propositions are not true. For example the function $f:(0,\infty) \longrightarrow \mathbb{R}$ $f(u) = \ln u$ is quasiconvex, but not convex. The next theorem shows that total convexity of the sublevel sets of a function is a necessary and sufficient condition for the function to be quasiconvex. Finally, note that, in contrast to Chapter 3, Theorem 3.6 stating the continuity of convex functions on open sets, quasiconvex functions are not necessarily continuous, as the example

$$f:\mathbb{R} \longrightarrow \mathbb{R},\ f(u) = \begin{cases} u & \text{for } u > 0 \\ u - 1 & \text{for } u \leq 0 \end{cases}$$

shows.

10.2. **Theorem**. *The function $f:A \longrightarrow \mathbb{R}$ is quasiconvex if and only if $A^c = f^{-1}(-\infty, c]$ is totally convex for each $c \in \mathbb{R}$.*

Proof. Let $f:A \longrightarrow \mathbb{R}$ be quasiconvex, $c \in \mathbb{R}$ and $x,y \in A^c$. Then $f(x) \leq c$, $f(y) \leq c$. Let $t \in (0,1)$. As

$$f(\gamma_{xy}(t)) \leq \max \{f(x), f(y)\} \leq c,$$

it follows $\gamma_{xy}(t) \in A^c$ and hence A^c is totally convex.

Suppose that each A^c is totally convex, i.e., $\gamma_{xy}(t) \in A^c$, $\forall\, x,y \in A^c$, $t \in (0,1)$ and $c = \max \{f(x), f(y)\}$. We find

$$f(\gamma_{xy}(t)) \leq c = \max \{f(x), f(y)\}.$$

Hence f is quasiconvex.

10.3. **Corollary**. *The balls $B(x_o,r)$ are totally convex if and only if the functions $f(x) = d(x_o,x)$ are quasiconvex for all x_o.*

Proof. $B(x_o,r) = f^{-1}(-\infty,r]$.

10.4. **Theorem**. *Let $f:A \longrightarrow \mathbb{R}$ be a quasiconvex function.*

1) Every local minimum is a global minimum or f is a constant in a *neighborhood of the local minimum.*

2) The subset of global minimum points is totally convex.

3) f is constant on every closed geodesic in A.

4) If f has a global maximum point in the interior of A, then f is constant.

Proof. 1) Let x_o be a local minimum point of f. This means that there exists $\varepsilon > 0$ such that

$$f(x_o) \le f(x), \ \forall\, x \in B(x_o, \varepsilon) \cap A.$$

Suppose there exists $z \in A$ with $z \ne x_o$ and $f(z) \le f(x_o)$. Consider $f(\gamma_{x_o z}(t))$, $t \in (0,1)$. As f is quasiconvex we have $f(\gamma_{x_o z}(t)) \le f(x_o)$.

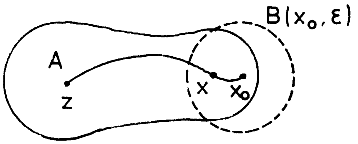

Fig.7

But there exist values of t (Fig.7) such that $x = \gamma_{x_o z}(t) \in B(x_o, \varepsilon) \cap A$. Hence $f(x) \ge f(x_o)$, i.e., $f(x) = f(x_o)$ or $f(x) > f(x_o)$. The first relation means $f(x) = \text{const}$ in a neighborhood of x_o . The second relation gives a contradiction which leads to $f(x_o) < f(z)$, $\forall\, z \in A$, i.e., x_o is a global minimum point.

2) If f has no minimum value on A, then the set of minimum points is empty and hence totally convex. If f has the minimum m on A, then the set of minimum points is the intersection of two totally convex sets, $A \cap A^m$, and hence it is totally convex.

3) By the preceding theorem A^c is totally convex. It follows that for every closed geodesic we have $f(\gamma(s)) \le f(\gamma(t))$, $\forall\, s,t \in \mathbb{R}$. Hence $f(\gamma(t)) = \text{const}$.

4) Suppose $x' \in \operatorname{int} A$ and $f(x) \le f(x')$, $\forall\, x \in A$. Choose $x \in A$ such that $x' = \gamma_{xy}(s)$, $s \in (0,1)$. It follows $f(x') \le \max\{f(x), f(y)\}$. If $f(x) \ge f(y)$, then $f(x) \le f(x') \le f(x)$ and hence $f(x) = \text{const}$; if $f(y) >$ $> f(x)$, then $f(x) \le f(x') < f(y)$, which is a contradiction.

10.5. **Corollaries**. 1) *If the minimum points set of a quasiconvex* *program has at least two points, then it has an infinity and the function that we minimize is not strongly quasiconvex.*

2) *If M has much closed geodesics, then the quasiconvex functions on M reduce to constants.*

3) *If M is a compact manifold, then the quasiconvex functions on M reduce to constants.*

Nontrivial examples of quasiconvex functions which are not convex

Let H be the Poincare plane. The functions $f_1, f_2 : H \longrightarrow \mathbb{R}$, $f_1(x,y) = x$, $f_2(x,y) = (x - a)^2 + y^2$ are quasiconvex. Indeed, the level sets of f_1 respectively f_2 are geodesics of H and a geodesic separates H in two convex subsets (some of them are sublevel sets).

The function $h: H \longrightarrow \mathbb{R}$, $h(x,y) = - y^2 - \ln y$ is quasiconvex because the sublevel sets are described by $y \geq y_0$ (convex sets).

10.6. **Theorem**. *Let $f:M \longrightarrow \mathbb{R}$ be a quasiconvex function. For each value c of f, the inclusion $M^c \subset M$ induces a homomorphism of the fundamental group $\pi_1(M^c)$ onto $\pi_1(M)$.*

Proof. Each element of $\pi_1(M,x)$, $x \in M^c$, can be represented by a geodesic loop γ at x. But by Theorem 10.2, γ lies in M^c.

10.7. **Definition**. The function $f:A \longrightarrow \mathbb{R}$ is called *strictly quasiconvex* if

$$f(\gamma_{xy}(t)) < \max\,\{f(x),\ f(y)\},$$

$\forall\ x,y \in A$, $\forall\ t \in (0,1)$ and $f(x) \neq f(y)$.

Any strongly quasiconvex function is strictly quasiconvex (the converse is not true). Any local minimum of a strictly quasiconvex function is a global minimum (not necessarily unique).

10.8. **Theorem**. *If $\varphi:\mathbb{R} \longrightarrow \mathbb{R}$ is an increasing function and $f:A \longrightarrow \mathbb{R}$ is quasiconvex, then $\varphi \circ f$ is a quasiconvex function on A.*

Proof. Every increasing function φ satisfies

$$\varphi(\max\,\{a,b\}) = \max\,\{\varphi(a),\ \varphi(b)\}.$$

We have

$$f(\gamma_{xy}(t)) \leq \max\,\{f(x),\ f(y)\}$$

and

$$\varphi \circ f(\gamma_{xy}(t)) \leq \varphi(\max\,\{f(x),\ f(y)\}) = \max\,\{\varphi \circ f(x),\ \varphi \circ f(y)\}.$$

Let ∇ be the Riemannian connection on M and $F:\ M \longrightarrow M$ be a diffeomorphism. Denote by $F_*\nabla$ the image of ∇ by F. If γ is a geodesic of (M,∇), then $F\circ\gamma$ is a geodesic of $(M, F_*\nabla)$.

10.9. **Theorem**. *If $f:A \longrightarrow \mathbb{R}$ is a quasiconvex function and F is a diffeomorphism, then $f\circ F^{-1}$ is a quasiconvex function on $F(A)$.*

Proof. Let $x,y \in A$ and $\gamma_{xy}(t)$ be a geodesic arc joining x to y. The set $F(A)$ is totally convex and the geodesic $F\circ\gamma_{xy}$ joins the points $F(x)$ and $F(y)$. We have

$$(f\circ F^{-1})(F(\gamma_{xy}(t)) = f(\gamma_{xy}(t)) \le \max\{f(x), f(y)\} =$$
$$= \max\{(f\circ F^{-1})(F(x)), (f\circ F^{-1})(F(y))\},$$

i.e., $f\circ F^{-1}$ is quasiconvex on $F(A)$.

§11. Distance from a point to a closed totally convex set

In this paragraph we develop the ideas in Chapter 2, §6.

Let (M,g) be a complete finite-dimensional Riemannian manifold with sectional curvature $K \le 0$. A submanifold A of M is called *totally geodesic* at a point $x_o \in A$ if every geodesic $\gamma(t)$ which is tangent to A at x_o is contained in A for small values of t. If A is totally geodesic at every point of A, then A is called a *totally geodesic submanifold.*

We denote by NA the *normal bundle* of A.

11.1. **Lemma**. *A submanifold A of M is closed and totally convex if and only if A is totally geodesic and the exponential map $\exp: NA \longrightarrow M$ is a diffeomorphism.*

Proof. Let A be a closed totally convex submanifold of M and $x,y \in A$. Suppose $\alpha: [0,1] \longrightarrow A$ is the unique shortest geodesic in A joining x and y and $\beta: [0,1] \longrightarrow M$ is a minimal geodesic in M from x to y. Since A is totally convex, β lies in A. The relation $L(\beta) \le L(\alpha)$ and the uniqueness of α imply $\beta = \alpha$. Consequently α is a geodesic of M and hence A is totally geodesic.

The total convexity of A implies that the inclusion $i : A \longrightarrow M$ induces a homomorphism i_* of $\pi_1(A)$ onto $\pi_1(M)$. Since A is closed and totally geodesic, the Hadamard-Hermann Theorem [42] asserts that $\exp: NA \longrightarrow M$ is a diffeomorphism (thus i_* is an isomorphism).

Assume that A is a totally geodesic submanifold of M and $\exp : NA \longrightarrow M$ is a diffeomorphism. Then A is closed in M. Let $\gamma : [0,1] \longrightarrow M$ be a geodesic in M joining points $x,y \in A$. Since exp is a diffeomorphism, γ is fixed-endpoint homotopic to a curve in A from x to y, hence to a geodesic τ of A joining x to y. But A is totally geodesic, so τ is a geodesic of M. Since $K \le 0$, we have $\gamma = \tau$, hence γ lies in A.

Our aim now is to extend the preceding lemma to arbitrary closed totally convex sets. For this we recall that a subset S of the topological space M is called a *retract* of M if there exists a continuous map $r: M \longrightarrow S$, called a *retraction*, such that $r|_S = id_S$. A subset S of the topological space M is called a *deformation retract* of M if there is a retraction $r: M \longrightarrow S$ and a homotopy $H: M \times [0,1] \longrightarrow M$ such that

$$H(x,o) = x, \ H(x,1) = r(x)$$
$$H(z,t) = z, \ z \in S, \ t \in [0,1].$$

If M is a *curve-connected topological space* and S is a deformation retract of M, then $\pi_1(M,z)$ is isomorphic to $\pi_1(S,z)$, $z \in S$.

Let A be a totally convex subset of the Riemannian manifold (M,g) and $x_o \in A$. The vector $X \in T_{x_o}M$ is tangent to A provided there exists a curve $\alpha: [0,\varepsilon] \longrightarrow A$ such that $\alpha(o) = x_o$, $\alpha'(o) = X$. If $x \in M$, a *perpendicular* from x to A is a geodesic $\gamma: [0,1] \longrightarrow M$ with the properties

$$\gamma(o) = x, \ \gamma(1) \in A, \ g(\dot\gamma(1), X) \ge 0, \ \forall\, X \in T_{\gamma(1)}M.$$

If $x \in A$, then $\gamma(t) = x$ is the unique perpendicular from x to A. In fact, if γ is a geodesic loop at $x \in A$, then $\gamma(t)$ lies in A, so $\gamma(1-t)$ is initially in A, whence $0 \le g(\dot\gamma(1), -\dot\gamma(1))$ and hence γ is constant.

Since (M,g) is a complete Riemannian manifold with $K \le 0$, the sum of the angles of a geodesic triangle is at most π [41].

11.2. **Lemma**. *If M is simply connected and A is a closed totally convex subset of M, then:*

1) *for each $x \in M$ there exists a unique perpendicular γ_x from x to A,*

2) *the perpendicular γ_x is the shortest geodesic from x to A,*

3) *the function*

$$\rho: M \longrightarrow A, \ \rho(x) = \gamma_x(1)$$

is a continuous retraction,

4) *the retraction ρ satisfies*

$$d(\rho(x),\ \rho(y)) \le d(x,y).$$

Proof. 1) - 2). Let $x \in M - A$. Since A is closed, there is a point y of A nearest to x. Let γ_x be the geodesic from x to y. The geodesic γ_x is perpendicular to A since γ_x is the shortest geodesic from x to A. Let α be a perpendicular from x to A and assume $\alpha \ne \gamma_x$. The hypotheses on M show that $\alpha(1) \ne \gamma_x(1)$. Let τ be the geodesic from $\gamma_x(1)$ to $\alpha(1)$. By total convexity, τ lies in A. From the definition of the perpendicular it follows: the angle between τ and γ_x at $\gamma_x(1)$ is at least $\pi/2$, the angle between τ and $\alpha(1)$ is at least $\pi/2$. Thus the sum of the angles of the geodesic triangle $\tau\alpha\gamma_x$ exceeds π. This contradiction shows that $\alpha = \gamma_x$.

3) the continuity of ρ follows from 2).

4) Let $x,y \in M$ and

$$r:\ [0,1] \times [0,1] \longrightarrow M$$

be the rectangle such that $r(\cdot,v)$ is the geodesic from $\gamma_x(v)$ to $\gamma_y(v)$. Let $L(v)$ be the length of $r(\cdot,v)$. Since $\left\| \frac{\partial r}{\partial u} \right\|$ is constant with respect to u, we have

$$L^2(v) = \left(\int_o^1 \left\| \frac{\partial r}{\partial u}(u,v) \right\| du \right)^2 = \int_o^1 \left\| \frac{\partial r}{\partial u}(u,v) \right\|^2 du.$$

From the first variation formula we obtain

$$L^{2'}(1) = 2(-g(\gamma_y'(1),\ -\tau'(1)) - g(\gamma_x'(1),\ \tau'(o))) \le 0,$$

where $\tau = r(\cdot,1)$. Since $K \le 0$, and γ_x, γ_y are geodesics, the second variation formula gives

$$L^{2''}(v) = 2\int_o^1 \left(\left\| \nabla_u \frac{\partial r}{\partial u} \right\|^2 - R\left(\frac{\partial r}{\partial u}, \frac{\partial r}{\partial v}, \frac{\partial r}{\partial u}, \frac{\partial r}{\partial v} \right) \right) du \ge 0.$$

Hence

$$L(1) \le L(o), \text{ i.e., } d(\rho(x),\ \rho(y)) \le d(x,y).$$

Remark. The set A can be a geodesic.

We now eliminate the hypothesis of simple connectivity from the preceding lemma.

11.3. **Theorem**. *Let $A \subset M$ be a closed totally convex set.*

1) *For each point $x \in M$ there is a unique perpendicular γ_x from x to A.*

2) *The perpendicular γ_x is the shortest geodesic from x to A.*

3) *The function*

$$H:\ M \times [0,1] \longrightarrow A,\ H(x,t) = \gamma_x(t)$$

is a continuous deformation of M onto A.

4) *The continuous retraction*

$$\rho\colon M \longrightarrow A,\ \rho(x) = \gamma_x(1)$$

satisfies

$$d(\rho(x),\ \rho(y)) \le d(x,y).$$

Proof. Let $\pi\colon M_1 \longrightarrow M$ be the simply connected Riemannian covering of M. Clearly $A_1 = \pi^{-1}(A)$ is closed and totally convex and we apply the preceding lemma in which we have established a retraction

$$\rho_1\colon M_1 \longrightarrow A_1 .$$

1)-2). Let $x \in M_1$ and τ_x be the geodesic joining x to $\rho_1(x)$. The retraction ρ_1 commutes with every deck transformation δ, since $\delta(A_1) = A_1$. It follows

$$\delta\tau_x = \tau_{\delta(x)}\ ,\ \forall\ x \in M_1 .$$

Hence, for $y \in M$, the geodesics τ_x ,$x \in \pi^{-1}(y)$, all project to the same geodesic σ_y in M which is the unique shortest geodesic from y to A. Lifting to M_1 , it follows that σ_y is the unique perpendicular from y to A.

3) Define

$$h\colon M_1 \times [0,1] \longrightarrow M_1\ ,\ h(x,t) = \tau_x(t)\ .$$

Since ρ_1 is continuous, so is h. But

$$\pi\tau_x = \sigma_{\pi(x)} \Rightarrow \pi h = H(\pi \times 1)$$

and hence H is continuous.

4) By lifting a minimal geodesic from $y \in M$ to $z \in M$, we obtain y_1 , $z_1 \in M_1$ such that $d(y,z) = d(y_1,z_1)$. Then

$$d(y_1,z_1) \ge d(\rho_1(y_1),\ \rho_1(z_1)) \ge d(\rho(y),\ \rho(z)),$$

since both ρ_1 and π are distance-nonincreasing.

11.4. **Corollary**. *Let A be a closed totally convex subset of M. If $\gamma\colon [0,1] \longrightarrow M$ is a closed geodesic, then γ is freely homotopic to a closed geodesic in A. If γ is not entirely in A, then all points of γ are at the same distance from A and the deformation takes place along a flat (finite length) cylinder with γ perpendicular to the meridians. Thus if $K < 0$, then A contains every closed geodesic of M.*

Proof. Denote $y = \gamma(o) = \gamma(1)$ and τ_y the perpendicular from y to A. To the rectangle used in the proof of Lemma 11.2 (viewed as being in the simply connected covering of M) corresponds a rectangle in M whose longitudinal curves are the geodesic loops in the same free homotopy class as γ, based at the points of τ_y. Since the curve γ is smooth, we have

$$L^{2\prime}(o) = 2(g(\tau'_y(o), \gamma'(1)) - g(\tau'_y(o), \gamma'(o))) = 0,$$

$$L^{2\prime}(o) \le L^{2\prime}_{\ll}(v) \le L^{2\prime}(1) \le 0.$$

So L is constant, the second variation vanishes, and the deformation takes place along a flat cylinder. The longitudinal and transverse vector fields are again orthogonal, so the loops comprising the rectangle are smooth.

11.5. **Corollary**. *If M has $K < 0$ and admits a convex function f without minimum, then there are no closed geodesics in M.*

Proof. Since f has no minimum, we have

$$\bigcap_{a \in f(M)} M^a = \emptyset .$$

By Corollary 11.4 any closed geodesic of M is included in this intersection.

Remarks. 1) This result fails for $K \le 0$; the (flat) circular cylinder M: $x^2 + y^2 = 1$ in $\mathbb{R}^3$ admits such a function (for example $f: M \longrightarrow \mathbb{R}$, $f(x,y,z) = e^z$) and has many closed geodesics.

2) Let $K < 0$. If M admits a convex function having a minimum, then the minimum set contains all the closed geodesics of M.

§12. DISTANCE BETWEEN TWO CLOSED TOTALLY CONVEX SETS

In this paragraph we develop the ideas in Chapter 2, §7.

The Riemannian manifold (M,g) is again supposed to be complete, and with the sectional curvature $K \le 0$. The geodesics are parametrized by arclength.

We begin with a useful remark on the distance between two geodesics.

12.1. **Theorem**. *Let (M,g) be a simply connected manifold with sectional curvature $K \le 0$ and let $\alpha(t)$, $\beta(t)$, $t \in \mathbb{R}$, be two geodesics. We allow $\beta(t) = const$, whereas $\|\alpha'(t)\| = 1$. The distance function $f(t) = d(\alpha(t), \beta(t))$ has the prrperties:*

1) *the equation $f(t) = 0$ has at most one solution,*

2) *for all t with $f(t) \neq 0$, f is C^∞ and convex.*

Proof. Let $\gamma_t(s)$, $s \in [0,1]$ be the unique geodesic joining $\alpha(t)$ to $\beta(t)$. Generally, $\|\gamma_t'(s)\| = 1$ and the mapping

$$F: [0,1] \times \mathbb{R} \longrightarrow M, \ F(s,t) = \gamma_t(s)$$

is of class C^∞. It satisfies

$$F(o,t) = \alpha(t), \ F(1,t) = \beta(t),$$

$$\frac{\partial F}{\partial s}(s,t) = \gamma_t'(s), \quad \left\| \frac{\partial F}{\partial s}(s,t) \right\| = d(\alpha(t), \beta(t)),$$

$$\frac{\partial F}{\partial t}(s,t) = Y_t(s) = \text{a Jacobi field along } \gamma_t(s),$$

$$Y_t(o) = \dot\alpha(t), \ Y_t(1) = \dot\beta(t).$$

Consider the function $h(t) = \frac{1}{2} f^2(t)$. We observe that

$$h(t) = \frac{1}{2}\, g\left(\frac{\partial F}{\partial s}(s,t), \frac{\partial F}{\partial s}(s,t)\right).$$

Hence

$$h'(t) = g\left(\frac{\nabla}{\partial t}\frac{\partial F}{\partial s}, \frac{\partial F}{\partial s}\right)(s,t) = g\left(\frac{\nabla}{\partial s}\frac{\partial F}{\partial t}, \frac{\partial F}{\partial s}\right)(s,t)$$

and

$$h''(t) = g\left(\frac{\nabla}{\partial s}\frac{\partial F}{\partial t}, \frac{\nabla}{\partial s}\frac{\partial F}{\partial t}\right)(s,t) + g\left(R\left(\frac{\partial F}{\partial t}, \frac{\partial F}{\partial s}\right)\frac{\partial F}{\partial t}, \frac{\partial F}{\partial s}\right)(s,t) +$$

$$+ \frac{d}{ds}\, g\left(\frac{\nabla^2 F}{\partial t^2}, \frac{\partial F}{\partial s}\right)(s,t).$$

The least term vanishes for $s = 0$ and $s = 1$. So we find

$$h''(t) = g\left(\frac{\nabla}{\partial s} Y_t(s), \frac{\nabla}{\partial s} Y_t(s)\right) - R(Y_t(s), \gamma_t'(s), Y_t(s), \gamma_t'(s)) \geq 0,$$

i.e., h is a convex function.

Considering $\sqrt{2h(t)} = f(t) = d(\alpha(t), \beta(t))$ and assuming $f(t) \neq 0$, we have

$$f'(t) = \frac{h'(t)}{f(t)}, \quad f''(t) = \frac{h''(t)f^2(t) - (h')^2(t)}{f^3(t)}.$$

So we get

$$f''(t) = \frac{1}{f^3(t)} \left\{ g\left(\frac{\nabla}{\partial s} Y_t(s), \frac{\nabla}{\partial s} Y_t(s)\right) g(\gamma_t'(s), \gamma_t'(s)) - \right.$$

$$\left. - g\left(\frac{\nabla}{\partial s} Y_t(s), \gamma_t'(s)\right)^2 - R(Y_t(s),\gamma_t'(s),Y_t(s),\gamma_t'(s))\, g(\gamma_t'(s),\gamma_t'(s))\right\} \geq 0,$$

by the Cauchy-Schwarz inequality and the negativity of sectional curvature.

Remarks. 1) If we have $f''(t) = 0$, then $\frac{\nabla}{\partial s} Y_t(s)$ must be collinear to $\gamma_t'(s) \neq 0$ and $K = 0$ on the 2-plane spanned by $\gamma_t'(s)$ and $Y_t(s)$.

2) A Riemannian manifold of curvature $K \leq 0$ has neither conjugate nor focal points, because this is true for the Euclidean space of curvature $K = 0$.

12.2. **Theorem**. *Let A,B be two closed totally convex sets in the Riemannian manifold (M,g). Let $x \in A$, $y \in B$ and $f(x,y) = d^2(x,y)$, where $d(x,y)$ is the distance from x to y.*

1) The set $A\times B$ is a totally convex set in $M \times M$;

2) $f: A\times B \longrightarrow \mathbb{R}$, $f(x,y) = d^2(x,y)$ is a convex function of class C^∞;

3) if it exists, the minimum of f is reached for a point $(x,y) \in A\times B$ such that the geodesic joining x to y is a common perpendicular of the sets A and B.