Hokkaido Mathematical Journal Vol. 20 (1991) p. 205-211

Theorem of Busemann-Mayer on Finsler metrics

To Professor Noboru Tanaka on his sixtieth birthday

Shoshichi KOBAYASHI*
(Received January 12, 1990)

1. Introduction

Let X be a manifold and TX its tangent bundle. A *pseudo-length function* on X is a real valued nonnegative function F on TX satisfying the condition

(1) $$F(c\xi)=|c|F(\xi) \qquad \text{for } \xi\in TX,\ c\in\mathbf{R}.$$

If $F(\xi)>0$ for every nonzero ξ, then F is called a *length function.*

If $\xi\in T_xX$, we write sometimes (x, ξ) for ξ although x is redundant. Similarly, we write occasionally $F(x, \xi)$ for $F(\xi)$. When we are working in a coordinate neighborhood U with a natural identification $TU\simeq U\times\mathbf{R}^n$, the notation $F(x, \xi)$ is more convenient as well as traditional since ξ may be used to denote an element of $\mathbf{R}^n$ as well as an element of TU.

We say that F is *convex* if it defines a pseudo-norm on each tangent space $T_xX, x\in X$, i e., if

(2) $$F(\xi+\xi')\leq F(\xi)+F(\xi') \qquad \text{for } \xi, \xi'\in T_xX.$$

A convex length function is usually called a *Finsler metric.*

Given a pseudo-length function F, its *indicatrix* Γ_x at $x\in X$ is defined to be

(3) $$\Gamma_x=\{\xi\in T_xX\,;\,F(\xi)\leq 1\}.$$

Then Γ_x is (1) star shaped in the sense that if $\xi\in\Gamma_x$ then $c\xi\in\Gamma_x$ for $|c|\leq 1$ and is (2) nontrivial in every direction in the sense that for every $\xi\in T_xX$ there is a nonzero c such that $c\xi\in\Gamma_x$.

Conversely, given a subset Γ_x in each tangent space T_xX satisfying the two conditions above, we can construct a pseudo-length function F by

(4) $$F(\xi)=\inf\{c>0\,;\,\frac{\xi}{c}\in\Gamma_x\} \qquad \text{for } \xi\in T_xX.$$

* Partially supported by NSF Grant DMS-880131.

Then a pseudo-length function F is convex if and only if its indicatrix Γ_x is a convex set for each $x \in X$. Given a (nonconvex) pseudo-length function F we can associate the largest convex pseudo-length function $\hat{F}$ such that $\hat{F} \leq F$ by considering the pseudo-length function defined by the convex hull $\hat{\Gamma}_x$ of Γ_x. Thus, $\hat{\Gamma}_x$ is, by construction, the indicatrix of $\hat{F}$ at x. It is aslo possible to define $\hat{F}$ as the double dual of F, (see [2]); but this fact will not be used here.

Let c be a piecewise smooth curve represented by $x(t)$, $a \leq t \leq b$. If a pseudo-length function F is upper semi-continuous, then the arc-length $L(c)$ of c is defined by

$$(5) \qquad L(c) = \int_c F = \int_a^b F(x'(t))\,dt,$$

and the pseudo-distance $d(p, q)$ between $p, q \in X$ is defined by

$$(6) \qquad d(p, q) = \inf_c L(c),$$

where the infimum is taken over all piecewise smooth curves c from p to q.

As we shall see later, if F is upper semi-continuous, so is $\hat{F}$. Therefore, using $\hat{F}$ we can similarly define the arc-length $\hat{L}(c)$ and the pseudo-distance $\hat{d}(p, q)$:

$$(7) \qquad \hat{L}(c) = \int_c \hat{F} = \int_a^b \hat{F}(x'(t))\,dt,$$

$$(8) \qquad \hat{d}(p, q) = \inf_c \hat{L}(c).$$

Since $\hat{F} \leq F$, we have $\hat{d}(p, q) \leq d(p, q)$.

The purpose of this paper is to prove the following theorem.

THEOREM. *Let X be a manifold with an upper semi-continuous pseudo-length function F. The pseudo-distance d defined by F coincides with the pseudo-distance $\hat{d}$ defined by $\hat{F}$.*

This theorem has been proved by Busemann and Mayer [1] under the assumption that F is continuous and strictly positive. The motivation for our technical generalization comes from complex analysis, namely the intrinsic infinitesimal pseudo-metric of a complex manifold which may be neither continuous nor strictly positive, (see [2], [4]).

2. Proof of the theorem.

The following lemma goes back to Carathéodory (see, for example,

[2], [5 ; p. 15]).

LEMMA 1. *Let V be a real vector space, and Γ a subset containing the origin $0 \in V$. Then an element $v \in V$ is in the convex hull $\hat{\Gamma}$ of Γ if and only if it is contained in a finite dimensional simplex having its vertices in Γ and having 0 as one of its vertices.*

Using Lemma 1 we prove the following

LEMMA 2. *Given $\eta \in \hat{\Gamma}_x$ and $\varepsilon > 0$, there exist linearly independent $\xi_1, \cdots, \xi_m \in \Gamma_x$ such that*

$$\eta = \xi_1 + \cdots + \xi_m \text{ and } \hat{F}(\eta) + \varepsilon > F(\xi_1) + \cdots + F(\xi_m).$$

If $\hat{F}(\eta) > 0$, there exist linearly independent $\xi_1, \cdots, \xi_m \in \Gamma_x$ such that

$$\eta = \xi_1 + \cdots + \xi_m \text{ and } \hat{F}(\eta) = F(\xi_1) + \cdots + F(\xi_m).$$

PROOF. For any positive real number s we set $s\Gamma_x = \{s\xi ; \xi \in \Gamma_x\}$ and $s\hat{\Gamma}_x = \{s\xi ; \xi \in \hat{\Gamma}_x\}$.

Let $r = \hat{F}(\eta)$. Then $\eta \in (r+\varepsilon)\hat{\Gamma}_x$ for any $\varepsilon > 0$. (If $r > 0$, then $\eta \in r\hat{\Gamma}_x$.) By Lemma 1, there exist linearly independent $\eta_1, \cdots, \eta_m$ in $(r+\varepsilon)\Gamma_x$ (in $r\Gamma_x$ if $r > 0$) such that

$$\eta = \sum t_i \eta_i \qquad \text{with } t_i > 0,\ \sum t_i \leq 1.$$

Then

$$\sum F(t_i \eta_i) = \sum t_i F(\eta_i) < (r+\varepsilon) \sum t_i \leq (r+\varepsilon).$$

By setting $\xi_i = t_i \eta_i$, we obtain the desired inequality.

If $r > 0$, then we can drop ε and obtain the inequality $\sum F(t_i \eta_i) \leq r$. Hence,

$$\sum F(\xi_i) \leq \hat{F}(\eta).$$

The reverse inequality follows from $\hat{F}(\xi_i) \leq F(\xi_i)$ and the triangular inequality satisfied by $\hat{F}$. Q. E. D.

The first application of Lemma 2 is the following

LEMMA 3. *If F is upper semi-continuous, so is $\hat{F}$.*

PROOF. Let $\eta_0 \in T_{x_0}X$, and $\varepsilon > 0$. Multiplying η_0 by a suitable nonzero constant, we may assume that $\hat{F}(\eta_0) \leq 1$. By Lemma 2, given $\varepsilon > 0$ there exist $\xi_1, \cdots, \xi_m \in \mathrm{T}_{x_0}X$ with $F(\xi_i) \leq 1$ such that

$$\eta_0 = \xi_1 + \cdots + \xi_m \text{ and } \hat{F}(\eta_0) + \varepsilon > F(\xi_1) + \cdots + F(\xi_m).$$

Let V_i be a neighborhood of ξ_i in TX such that

$$F(\xi_i') < F(\xi_i) + \frac{1}{m}\varepsilon \qquad \text{for } \xi_i' \in V_i.$$

Let W be the neighborhood of η_0 in TX defined by $W = V_1 + \cdots + V_m$. Then for any $\eta' = \xi_1' + \cdots + \xi_m' \in W$ with $\xi_i' \in V_i$ we have

$$F(\eta_0) + 2\varepsilon > \sum(F(\xi_i) + \frac{1}{m}\varepsilon) > \sum F(\xi_i') \geq$$
$$\sum \hat{F}(\xi_i') \geq F(\sum \xi_i') = \hat{F}(\eta').$$

Q. E. D.

LEMMA 4. *If F is an upper semi-continuous pseudo-length function on X, it is the limit of a monotone decreasing sequence of continuous length functions H_k. Furthermore, if $\hat{H}_k$ denotes the continuous convex length function associated with H_k, then $\hat{F}$ is the limit of a monotone decreasing sequence $\{\hat{H}_k\}$.*

PROOF. Let $SX \subset TX$ be the tangent unit sphere bundle defined by a Riemannian metric g of X. Since $F_{|SX}$ is an upper semi-continuous non-negative function, it is a limit of a monotone decreasing sequence of continuous positive functions H_k on SX (see, for example [3; p. 43]). Since $F(-\xi) = F(\xi)$, we can choose H_k in such a way that $H_k(-\xi) = H_k(\xi)$. We extend H_k to TX by setting.

$$H_k(c\xi) = |c| H_k(\xi) \qquad \text{for } \xi \in SX,\ c \in \mathbf{R}.$$

Then $\{H_k\}$ is a monotone decreasing sequence of continuous length functions, and $F(\xi) = \lim H_k(\xi)$ for $\xi \in TX$. Since $F \leq H_{k+1} \leq H_k$, we have $\hat{F} \leq \hat{H}_{k+1} \leq \hat{H}_k$. Given a nonzero $\xi \in TX$, choose a convex length function G such that $G(\xi) = 1$. Since $\lim H_k(\xi) = F(\xi)$, given $\varepsilon > 0$ there is an integer k_0 such that

$$H_k(\xi) < F(\xi) + \varepsilon = F(\xi) + \varepsilon G(\xi) \qquad \text{for } k > k_0.$$

Hence,

$$\hat{H}_k(\xi) < \hat{F}(\xi) + \varepsilon G(\xi) = \hat{F}(\xi) + \varepsilon \qquad \text{for } k > k_0.$$

Thus, $\lim \hat{H}_k(\xi) = \hat{F}(\xi)$. Q. E. D.

Let $p, q \in X$, and let c be a piecewise smooth curve from p to q reperesented by $x(t)$, $a \leq t \leq b$. Since $\hat{F} \leq F$, we have $\hat{L}(c) \leq L(c)$. The problem is to show that given $\varepsilon > 0$ there is another curve $\tilde{c}$ from p to q such that

$$(9)\qquad L(\tilde{c})<\hat{L}(c)+3\varepsilon.$$

By subdividing c if necessary, we may assume that c is contained in a single coordinate neighborhood U. For the sake of convenience we fix an arbitrarily chosen Riemannian metric g on X. It is convenient to choose g in such a way that in U it is the Euclidean metric defined by the local coordinate system. Without loss of generality we may assume that the velocity $x'(t)$ is of unit length with respect to g.

Let H_k be as in Lemma 4 and put

$$(10)\qquad \hat{L}_k(c)=\int_c \hat{H}_k.$$

Then by the Lebesgue convergence theorem, given $\varepsilon>0$ there is an integer k_0 such that

$$(11)\qquad \hat{L}_k(c)<\hat{L}(c)+\varepsilon \qquad \text{for } k>k_0.$$

Let $\pi=(a=t_0<t_1<\cdots<t_r=b)$ be a subdivision of the interval $[a, b]$, and let $|\pi|=\max\{t_1-t_0,\cdots,t_r-t_{r-1}\}$. We set

$$\Delta t_i=t_i-t_{i-1}.$$

Using the local coordinate system in U, we define

$$\Delta x_i=x(t_i)-x(t_{i-1})\in\mathbf{R}^n.$$

Under the identification $T_{x(t_{i-1})}\simeq\mathbf{R}^n$ by the coordinate system, $\Delta x_i/\Delta t_i$ is approximately equal to $x'(t_{i-1})$ so that $|\Delta x_i/\Delta t_i|$ is approximately equal to 1.

For each π and k, we set

$$(12)\qquad \hat{S}_{k,\pi}=\sum_{i=1}^{r}\hat{H}_k(x(t_{i-1}), \Delta x_i)=\sum_{i=1}^{r}\hat{H}_k\left(x(t_{i-1}), \frac{\Delta x_i}{\Delta t_i}\right)\Delta t_i.$$

(As we explained in the preceding section, we write the base point $x(t_{i-1})$ in (12) explicitly since we are using the local coordinate system).

Then, given $\varepsilon>0$ there is $\delta_1>0$ such that

$$(13)\qquad \hat{S}_{k,\pi}<\hat{L}_k(c)+\varepsilon \qquad \text{if } |\pi|<\delta_1.$$

For a fixed $k>k_0$, there is $\delta_2>0$ such that for every $t, a\leq t\leq b$,

$$(14)\qquad H_k(y, \xi)<H_k(x(t), \xi)+\frac{\varepsilon|\xi|}{b-a} \qquad \text{for } |y-x(t)|<\delta_2,\ \xi\in\mathbf{R}^n,$$

where $|y-x(t)|$ denotes the Euclidean distance from $x(t)\in c$ to $y\in X$ with respect to the local coordinate system.

Let $\delta=\min\{\delta_1, \delta_2\}$, and fix a subdivision π of $[a, b]$ such that $|\pi|<\delta$. Since $|\Delta x_i/\Delta t_i|$ is approximately equal to 1 and since $\Delta t_i<\delta$, we have $|\Delta x_i|<\delta$.

We fix i and $k>k_0$. Since $\hat{H}_k(\Delta x_i)>0$, by Lemma 1 there exist linearly independent $\xi_1,\cdots,\xi_m\in T_{x(t_{i-1})}X$ such that

$$(15)\qquad \Delta x_i=\xi_1+\cdots+\xi_m$$

and

$$(16)\qquad \hat{H}_k(x(t_{i-1}), \Delta x_i)=H_k(x(t_{i-1}), \xi_1)+\cdots+H_k(x(t_{i-1}), \xi_m).$$

This says that the length of the line segment from the origin to Δx_i in $T_{x(t_{i-1})}X$ measured by the length function $\hat{H}_k$ is equal to the length of the polygonal path from the origin to Δx_i via vertices $\xi_1, \xi_1+\xi_2, \xi_1+\xi_2+\xi_3,\cdots, \xi_1+\xi_2+\cdots+\xi_{r-1}$ measured by the length function H_k.

Using the local coordinate system we identify a neighborhood of the origin in $T_{x(t_{i-1})}X$ with a neighborhood of $x(t_{i-1})$ in X. Let $\tilde{c}_i$ be the polygonal path in X corresponding to the polygonal path in $T_{x(t_{i-1})}X$ described above. Then $\tilde{c}_i$ goes from $x(t_{i-1})$ to $x(t_i)$. For each i, we replace the portion of c from $x(t_{i-1})$ to $x(t_i)$ by $\tilde{c}_i$. Let $\tilde{c}$ be the resulting path from $p=x(a)$ to $q=x(b)$. We shall show that $\tilde{c}$ has the desired property.

We estimate the length $L(\tilde{c}_i)$ of $\tilde{c}_i$ measured by H_k. Since $|\Delta x_i|<\delta$ and since $\hat{H}_k(\xi_j)\leq H_k(\xi_j)$, it follows from (15) and (16) that $\tilde{c}_i$ is contained in the δ-neighborhood of $x(t_{i-1})$. (By "δ-neighborhood" we mean the Euclidean neighborhood $\{y\in X; |y-x(t_{i-1})|<\delta\}$.) Therefore, by (14)

$$(17)\qquad H_k(y, \xi)<H_k(x(t_{i-1}), \xi)+\frac{\varepsilon|\xi|}{b-a}\qquad \text{for } \xi\in\mathbf{R}^n$$

at every point y of $\tilde{c}_i$. Integrating (17) along $\tilde{c}_i$ and using (16), we obtain

$$(18)\qquad \int_{\tilde{c}_i}H_k<\hat{H}_k(x(t_{i-1}), \Delta x_i)+\frac{\varepsilon|\Delta x_i|}{b-a}.$$

Since $\Sigma|\Delta x_i|$ is approximately equal to $\Sigma\Delta t_i=b-a$, summing over i we obtain

$$(19)\qquad \int_{\tilde{c}}H_k<\hat{S}_{k,\pi}+\varepsilon.$$

Combining (11), (13) and (19) we obtain

(20) $$\int_{\bar{c}} H_k < \hat{L}(c) + 3\varepsilon.$$

Since $\int_{\bar{c}} F \leq \int_{\bar{c}} H_k$, we obtain the desired inequality (9), thus completing the proof of the theorem.

Bibliography

[1] H. BUSEMANN and W. MAYER, On the foundation of calculus of variations, Trans. Amer. Math. Soc. 49 (1941), 173-198.
[2] S. KOBAYASHI, A new invariantinfinitesimal metric, Intern'l Journ. Math. 1 (1990), 83-90.
[3] E. J. MCSHANE, *Integration*, Princeton Univ. Press, 1944.
[4] H. L. ROYDEN, Remarks on the Kobayashi metrics, Springer-Verlag, Lecture Notes in Math. 185 (1971), 125-137.
[5] F. A. VALENTINE, *Convex Sets*, R. E. Krieger Publ. Co., 1976.

Department of Mathematics
University of California, Berkeley
Berkeley, CA 94720 USA