§5. The First variation of Arc Length and Energy.

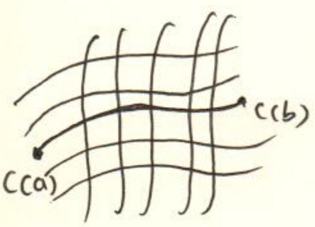
(In retrospection of our calculations in Chap II.)

In Chap II, we derive the geodesic equations as the Euler-Lagrange equations of the Length and Energy functionals via local coordinates computations. Now, with the convenient notion of (Levi-Civita) connections, we can carry out an easier (intrinsic) computation.

Recall for any smooth curve $c:[a,b]\to M$, we have

$$L(c) := \int_a^b \sqrt{\langle \dot{c}(t), \dot{c}(t)\rangle}\, dt$$
$$= \int_a^b \sqrt{\langle dc(\tfrac{\partial}{\partial t}), dc(\tfrac{\partial}{\partial t})\rangle}\, dt$$
$$E(c) = \frac{1}{2}\int_a^b \langle dc(\tfrac{\partial}{\partial t}), dc(\tfrac{\partial}{\partial t})\rangle dt.$$

<u>Definition</u> 7. Let $c:[a,b]\to M$ be a smooth curve, $\varepsilon>0$.
A <u>variation</u> of c is a ~~differentiable~~ smooth map
$$F: [a,b]\times(-\varepsilon,\varepsilon)\to M$$
with $F(t,0)=c(t)$, $\forall t\in[a,b]$.



The variation is called <u>proper</u> if the endpoints stay fixed, i.e. $F(a,s)=c(a)$, $F(b,s)=c(b)$, $\forall s\in(-\varepsilon,\varepsilon)$. □

For simplicity, we will denote
$$\frac{\partial F}{\partial s} = dF\left(\frac{\partial}{\partial s}\right),\quad \frac{\partial F}{\partial t} = dF\left(\frac{\partial}{\partial t}\right).$$
We also denote $c_s(t) = F(t,s)$.

Definition 8: We call

$$V(t) = \frac{\partial F}{\partial s}(t,0) = \frac{\partial F}{\partial s}(c(t))$$

the variational field of f along c. (It is a vector field along c)

Theorem 2 (The First Variation Formula) Let $F(t,s)$ be a variation of a smooth curve c.

Let us write $L(s) := L(c_s)$, $E(s) := E(c_s)$ for short.

Then

$$\frac{d}{ds}\Big|_{s=0} L(s) := L'(0) = \int_a^b \frac{1}{|\dot{c}(t)|}\left(\frac{d}{dt}\langle V(t), \dot{c}(t)\rangle - \langle V(t), \underbrace{\tilde{\nabla}_{\frac{\partial}{\partial t}} \dot{c}(t)}_{\frac{D\dot{c}}{dt}}\rangle\right) dt$$

This is oftenly written as $\nabla_{\dot{c}}\dot{c}$

$$\frac{d}{ds}\Big|_{s=0} E(s) := E'(0) = \langle V(b), \dot{c}(b)\rangle - \langle V(a), \dot{c}(a)\rangle - \int_a^b \langle V(t), \tilde{\nabla}_{\frac{\partial}{\partial t}} \dot{c}(t)\rangle dt.$$

Proof:

$$\frac{d}{ds} L(s) = \int_a^b \frac{d}{ds}\left\langle \frac{\partial F}{\partial t}(t,s), \frac{\partial F}{\partial t}(t,s)\right\rangle^{\frac{1}{2}} dt$$

$$= \int_a^b \frac{1}{2\left\langle \frac{\partial F}{\partial t}(t,s), \frac{\partial F}{\partial t}(t,s)\right\rangle^{\frac{1}{2}}} \cdot \frac{d}{ds}\left\langle \frac{\partial F}{\partial t}(t,s), \frac{\partial F}{\partial t}(t,s)\right\rangle dt$$

$$\overset{(*1)\text{ on p. 90}}{=} \int_a^b \frac{1}{\left\langle \frac{\partial F}{\partial t}(t,s), \frac{\partial F}{\partial t}(t,s)\right\rangle^{\frac{1}{2}}} \cdot \left\langle \underbrace{\tilde{\nabla}_{\frac{\partial}{\partial s}} \frac{\partial F}{\partial t}(t,s)}_{\frac{D}{\partial s}\frac{\partial F}{\partial t}}, \frac{\partial F}{\partial t}(t,s)\right\rangle dt$$

$$\overset{(*2)\text{ on p. 92.}}{=} \int_a^b \frac{1}{\left\langle \frac{\partial F}{\partial t}(t,s), \frac{\partial F}{\partial t}(t,s)\right\rangle^{\frac{1}{2}}} \left\langle \tilde{\nabla}_{\frac{\partial}{\partial t}} \frac{\partial F}{\partial s}(t,s), \frac{\partial F}{\partial t}(t,s)\right\rangle dt.$$

$$\overset{(*1)\text{ on p. 90}}{=} \int_a^b \frac{1}{\left\langle \frac{\partial F}{\partial t}(t,s), \frac{\partial F}{\partial t}(t,s)\right\rangle^{\frac{1}{2}}} \left(\frac{d}{dt}\left\langle \frac{\partial F}{\partial s}(t,s), \frac{\partial F}{\partial t}(t,s)\right\rangle - \left\langle \frac{\partial F}{\partial s}(t,s), \tilde{\nabla}_{\frac{\partial}{\partial t}} \frac{\partial F}{\partial t}(t,s)\right\rangle\right) dt.$$

$\Rightarrow \frac{d}{ds}\Big|_{s=0} L(s) = \int_a^b \frac{1}{|\dot c(t)|}\left(\frac{d}{dt}\langle V(t), \dot c(t)\rangle - \langle V(t), \tilde\nabla_{\frac{\partial}{\partial t}} \dot c(t)\rangle\right) dt.$

Similarly, we obtain the formula for $E'(0)$. □

Observe that when c is parametrized proportionally to arclength, i.e. $|\dot c(t)| \equiv$ const, the variations of ~~length~~ L and E leads to the same critical point. (We observed this fact using Hölder ineq. in Chap. II).

A smooth curve $c : [a,b] \to M$ is a critical point of the energy E for all proper variations. iff $\tilde\nabla_{\frac{\partial}{\partial t}} \dot c(t) = 0$. (i.e., c is a geodesic).

(Note, by property of parallel transport $\tilde\nabla_{\frac{\partial}{\partial t}} \dot c(t) = 0 \Rightarrow |\dot c(t)| \equiv$ const $\Rightarrow$ c is parametrized proportionally to arc length).

More generally, we consider piecewise smooth curve $c : [a,b] \to M$. That is, we have a subdivision

$$a = t_0 < t_1 < t_2 < \cdots < t_k < t_{k+1} = b$$

s.t. c is smooth on each interval $[t_i, t_{i+1}]$.

Correspondingly, we consider "piecewise "smooth" variations" of c, which are continuous functions

$$F : [a,b] \times (-\varepsilon, \varepsilon) \to M$$

such that F is smooth on each $[t_i, t_{i+1}] \times (-\varepsilon, \varepsilon)$. and $\frac{\partial F}{\partial s}$ is ~~even~~ well defined even at t_i's.

Then, as a direct consequence of Theorem 1, we have

Corollary 1. Let c be a piecewise smooth curve and F be a corresponding piecewise smooth variation. Then

$$E'(0) = \frac{d}{ds}\Big|_{s=0} E(c_s) = \langle V(b), \dot c(b)\rangle - \langle V(a), \dot c(a)\rangle$$
$$- \sum_{i=1}^{k} \langle V(t_i), \dot\gamma(t_i^+) - \dot\gamma(t_i^-)\rangle$$
$$- \int_a^b \langle V(t), \tilde\nabla_{\frac{\partial}{\partial t}} \dot c(t)\rangle dt.$$

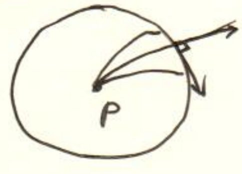
It turns out the first variation formulae are also very useful for non-proper variations. We discuss here Gauss' lemma.

Recall in a normal neighborhood U_p of a point $p \in M$, we can introduce a polar coordinate such that

$$g = dr \otimes dr + g_{\varphi\varphi}(r,\varphi)\, d\varphi \otimes d\varphi.$$

Here the fact $g_{r\varphi} \equiv 0$ on the whole U_p is also called Gauss' lemma.

Lemma (Gauss' Lemma). In U_p, the geodesics through p are perpendicular to the hypersurfaces



$$\{\exp_p(v) : \underset{r}{\|v\| = \text{const}} < \delta\}$$

(Precisely, let ~~$\rho(t) = tv$~~ $v \in T_pM$, $\rho(r) = rv$ is a ray through $0 \in T_pM$.

Let $w \in T_{\rho(r)}(T_pM)$ is perpendicular to $\rho'(r)$.

Then $\langle (d\exp_p)(\rho(r))(w), (d\exp_p)(\rho(r))(\rho'(r))\rangle = 0$.) (*)

Proof: Let $v(s) : (-\varepsilon, \varepsilon) \to T_pM$ be ~~the~~ a curve with $\underset{=\rho(r)}{v(0) = rv}$, $\dot v(0) = w$. and $\|v(s)\| = r$.

Then we have a variation

$$F(t,s) = \exp_p(t\, v(s)), \quad t \in [0,r], \ s \in (-\varepsilon, \varepsilon).$$

with $F(t,0) = \exp_p(tv) = c(t)$

Notice that ~~$E(F(t,s))$~~

$$E(c_s) = \frac{1}{2}\int_0^r \langle v(s), v(s)\rangle\, dt = \frac{1}{2} r^3 \equiv \text{const}.$$

Theorem 2 $\Rightarrow$

$$0 = E'(0) = \langle V(r), \dot c(r)\rangle - \langle V(0), \dot c(0)\rangle - \int_0^r \langle V(t), \tilde\nabla_{\frac{\partial}{\partial t}} \dot c(t)\rangle dt.$$

Since $\tilde\nabla_{\frac{\partial}{\partial t}} \dot c(t) = 0$, and $V(0) = \frac{\partial F}{\partial s}\Big|_{t=0, s=0} = 0$, we conclude $\langle V(r), \dot c(r)\rangle = 0 \quad (*)$.

Recall $V(r) = \frac{\partial F}{\partial s}\Big|_{s=0, t=r} = \frac{\partial}{\partial s}\Big|_{s=0} \exp_p(r v(s))$

$$= (d\exp_p)(\rho(r))(w).$$

$$\dot c(r) = (d\exp_p)(\rho(r))(\rho'(r)).$$

We see $(*)$ implies $(\spadesuit)$ $\square$

Lecture 10, 2017.03.28

§5. Covariant differentiation, Hessian, and Laplacian.

Recall the covariant derivative of a (r,s)-tensor A ~~along X~~ satisfies
$$\nabla_{fX+gY} A = f\nabla_X A + g\nabla_Y A.$$
That is $\nabla_X A$ is linear over C^∞ fcts for the argument X.
"tensorial"

Therefore we can define a $(r,s+1)$-tensor ∇A for each (r,s)-tensor A.
as
$$\nabla A(\omega^1,\dots,\omega^r, X_1,\dots,X_s, X)$$
$$:= \nabla_X A(\omega^1,\dots,\omega^r, X_1,\dots,X_s), \quad \forall\ \omega^i \in \Gamma(T^*M), X_j, X \in \Gamma(TM)$$
We call ∇A the covariant differentiation of A.

Particularly, consider a $(0,0)$-tensor, i.e., a function f.
The covariant differentiation ∇f of f is given as
$$\forall X \in \Gamma(TM):\ \nabla f(X) := \nabla_X f = X(f) = df(X)$$
$$\Rightarrow \quad \nabla f = df \text{ is a } (0,1)\text{-tensor.}$$
exterior derivative

We can then discuss iteratively
$$\nabla^2 f := \nabla(\nabla f), \quad \nabla^3 f = \overbrace{\nabla(\nabla(\nabla f))}^{\nabla(\nabla^2 f)}, \ \dots$$
generally, $\nabla^2 A$, $\nabla^3 A$, ...

$$\boxed{\text{Warning: } \nabla^2 A(\dots, X, Y) \neq \nabla_Y \nabla_X A(\dots) \ !!}$$

For $X, Y \in \Gamma(TM)$, we have
$$\nabla^2 f(X,Y) = \nabla(\nabla f)(X,Y) = \nabla_Y(\nabla f)(X)$$
$$= Y(\nabla f(X)) - (\nabla f)(\nabla_Y X)$$
$$= Y(Xf) - \nabla_Y X(f).$$

Proposition 11. $\nabla^2 f(X,Y) - \nabla^2 f(Y,X) = T(X,Y)f.$

Proof: $\nabla^2 f(X,Y) - \nabla^2 f(Y,X)$

$$= YXf - (\nabla_Y X)f - XYf + (\nabla_X Y)f$$

$$= [Y,X]f - (\nabla_Y X - \nabla_X Y)f$$

$$= T(X,Y)f. \qquad \square$$

That is, when the connection ∇ is torsion-free, we have $\nabla^2 f(X,Y) = \nabla^2 f(Y,X)$, $\forall X,Y \in \Gamma(TM)$.

i.e. $\nabla^2 f$ is a symmetric $(0,2)$-tensor field.

We call $\nabla^2 f$ the $\boxed{\text{Hessian}}$ of f.

Example: On $\mathbb{R}^n$, given the canonical connection, we have

$$\nabla^2 f(X,Y) = (Y^1, \dots, Y^n)\left(\frac{\partial^2 f}{\partial x^i \partial x^j}\right)\begin{pmatrix} X^1 \\ \vdots \\ X^n \end{pmatrix}$$

where $X = X^i \frac{\partial}{\partial x^i}$, $Y = Y^j \frac{\partial}{\partial x^j}$. $\square$

Since $\nabla^2 f\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right) = \frac{\partial}{\partial x^i}\left(\frac{\partial}{\partial x^j} f\right) = \frac{\partial^2}{\partial x^i \partial x^j}$.

The trace of the Hessian is the Laplacian

For any $X \in \Gamma(TM)$, we have a linear map

$$\nabla X : \Gamma(TM) \to \Gamma(TM)$$
$$Y \mapsto \nabla_Y X.$$

At a point p, $\nabla X : T_pM \to T_pM$ is a linear transformation between two vector space. Hence, it make sense to talk about the trace of ∇X at each p, which gives us a function on M.

Lemma $\forall X \in \Gamma(TM)$, we have $\operatorname{div}(X) = \operatorname{tr}(X)$.

Recall from Chap 1, $\operatorname{div} X = \frac{1}{\sqrt{G}} \frac{\partial}{\partial x^i}(X^i \sqrt{G})$, $G = \det(g_{ij})$ in local coordinate.

<u>Proof</u>: We only need to prove it at one point $p \in M$.

Pick a coordinate neighborhood $p \in U$, (U,x), we have

$$\nabla X = (\nabla_{\frac{\partial}{\partial x^i}} X) dx^i$$

$$= \frac{\partial X^k}{\partial x^i} \frac{\partial}{\partial x^k} \otimes dx^i + X^j \Gamma_{ij}^k \frac{\partial}{\partial x^k} \otimes dx^i$$

Therefore

$$tr(\nabla X) = \frac{\partial X^i}{\partial x^i} + X^j \Gamma_{ij}^i$$

<u>Proposition</u>: Let ∇ be the Levi-Civita connection on (M,g).

Then $\Gamma_{ji}^j = \frac{1}{\sqrt{G}} \frac{\partial}{\partial x^i} \sqrt{G}$. (see Appendix)

$$= \frac{\partial X^i}{\partial x^i} + X^i \Gamma_{ji}^j = \frac{\partial X^i}{\partial x^i} + X^i \frac{1}{\sqrt{G}} \frac{\partial}{\partial x^i} \sqrt{G}.$$

$$= \frac{1}{\sqrt{G}} \frac{\partial}{\partial x^i} \left(X^i \sqrt{G} \right) = \operatorname{div}(X).$$

$\square$

~~Now for an (0,2) tensor S:~~

Recall since g is non-degenerate and bilinear on T_pM, we have isomorphisms between TM and T^*M:

$$\flat : TM \to T^*M$$
$$X \mapsto \flat(X). \qquad \flat(X)(Y) = g(X,Y).$$

and

$$\sharp : T^*M \to TM$$
$$\omega \mapsto \sharp(\omega) \qquad g(\sharp(\omega), Y) = \omega(Y).$$

In local coordinate

$$\flat(X^i \partial_i) = g_{ij} X^i dx^j$$
$$\sharp(\omega_i dx^i) = g^{ij} \omega_i \frac{\partial}{\partial x^j}.$$

$\flat$, $\sharp$: musical isomorphisms.

Now we define the trace of S as the trace of the linear map $X \longmapsto \# S(X, \cdot)$.

Note $\underline{g(\# S(X,\cdot), Y) = S(X,Y)}$.

In local coordinate, $S = S_{ij}\, dx^i dx^j$

$$\# S(X,\cdot) = \# (S_{ij} X^i dx^j)$$

$$= S_{ij} X^i g^{jk} \frac{\partial}{\partial x^k}.$$

Hence $\operatorname{tr}(X \longmapsto \# S(X,\cdot)) = g^{ij} S_{ij} = g^{ij} S(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j})$.

$:= \operatorname{tr} S$

Let us come back to $\operatorname{Hess} f$.

<u>Lemma</u>: $\forall X, Y \in \Gamma(TM)$,

$$\operatorname{Hess} f(X,Y) = g(\nabla_X \operatorname{grad} f, Y).$$

<u>Proof</u>: $RHS = \nabla_X (g(\operatorname{grad} f, Y)) - g(\operatorname{grad} f, \nabla_X Y)$

$$= \nabla_X (Yf) - (\nabla_X Y) f$$

$$= X(Yf) - (\nabla_X Y) f = \operatorname{Hess} f(X,Y). \quad \square$$

Hence we have

$$\operatorname{tr} \operatorname{Hess} f = \operatorname{tr}(X \mapsto \nabla_X \operatorname{grad} f)$$

$$= \operatorname{tr}(\nabla \operatorname{grad} f)$$

$$= \operatorname{div}(\operatorname{grad} f)$$

$$= \Delta f. \quad \text{Laplace – Beltrami operator}$$

Recall from Chap 1. that

$$\Delta f = \frac{1}{\sqrt{G}} \frac{\partial}{\partial x^k} \left(g^{ik} \sqrt{G} \frac{\partial f}{\partial x^i} \right). \quad \square$$

Appendix: The techniqual Lemma.

Prop: Let ∇ be the Levi-Civita connection on (M,g).
Then $\Gamma^j_{ji} = \frac{1}{\sqrt{G}} \frac{\partial}{\partial x^i}(\sqrt{G})$, $G := \det(g_{ij})$.

Proof: Recall

$$\Gamma^j_{ji} = \frac{1}{2} g^{jk}(g_{jk,i} + g_{ki,j} - g_{ji,k})$$

$$= \frac{1}{2} g^{jk} \cdot \frac{\partial}{\partial x^i}(g_{jk})$$

$$= \frac{1}{2}\mathrm{tr}\left[\left(g^{rs}\right)_{r,s} \cdot \frac{\partial}{\partial x^i}\left(g_{jk}\right)_{j,k}\right]$$

matrix multiplication.

Note moreover, $\left(g^{rs}\right)_{r,s}$ is the inverse matrix of $\left(g_{jk}\right)_{j,k}$.

We need the following result:

Lemma. Let $A = A(t)$ be a family of nonsingular matrices that depends ~~on~~ smoothly on t, then

$$\mathrm{tr}\left(A^{-1} \frac{d}{dt} A\right) = \frac{d}{dt} \ln \det A.$$

Sketch of proof:
Observe the Lemma is obvious, when A is 1×1. ~~And~~
For a diagonal matrix A.

$$\mathrm{tr}\left[\begin{pmatrix} A_1^{-1}(t) & & \\ & \ddots & \\ & & A_n^{-1}(t)\end{pmatrix}\begin{pmatrix} A_1'(t) & & \\ & \ddots & \\ & & A_n'(t)\end{pmatrix}\right]$$

$$= \mathrm{tr}\begin{pmatrix} A_1^{-1}(t)A_1'(t) & & \\ & \ddots & \\ & & A_n^{-1}(t)A_n'(t)\end{pmatrix} = \sum_{i=1}^n A_i^{-1}(t)\, A_i'(t)$$

$$\Vert$$

$$\frac{d}{dt} \ln\det\begin{pmatrix} A_1(t) & & \\ & \ddots & \\ & & A_n(t)\end{pmatrix} = \frac{d}{dt} \ln \prod_{i=1}^n A_i(t) = \frac{d}{dt}\sum_{i=1}^n \ln A_i(t)$$

Recall both trace and det are invariants under similar transformation. For diagonalizable A, we have $A = P^{-1}DP$.

$$\det A = \det D$$
$$tr\left(A^{-1}\frac{d}{dt}A\right) = tr\left(D^{-1}\frac{d}{dt}D\right).$$

$$\Big(tr\left(P^{-1}D^{-1}P\,(P^{-1}DP)'\right) = tr\big(P^{-1}D^{-1}P(P^{-1})'DP + P^{-1}D^{-1}PP^{-1}D'P + P^{-1}D^{-1}PP^{-1}DP'\big)$$
$$= tr\big(\underbrace{P(P^{-1})' + P^{-1}P'}_{(P P^{-1})'=0}\big) + tr\left(D^{-1}D'\right) = tr\left(D^{-1}D'\right)\Big).$$

Hence Lemma is true for diagonalizable matrices.
By standard perturbation trick, one can prove Lemma in its full generality. □

Let us continue:

$$\Gamma^j_{ji} = \frac{1}{2} tr\left[\left(g^{rs}\right)_{r,s}\frac{\partial}{\partial x^i}\left(g_{jk}\right)_{j,k}\right]$$
$$= \frac{1}{2}\frac{\partial}{\partial x^i}\ln\det\left(g_{jk}\right)_{j,k} = \frac{1}{2}\frac{\partial}{\partial x^i}\left(\ln G\right)$$
$$= \frac{\partial}{\partial x^i}\left(\ln\sqrt{G}\right) = \frac{1}{\sqrt{G}}\frac{\partial}{\partial x^i}\left(\sqrt{G}\right).$$
□