

ALMA MATER STUDIORUM · UNIVERSITÀ DI
BOLOGNA

FACOLTÀ DI SCIENZE MATEMATICHE, FISICHE E NATURALI
Corso di Laurea in Matematica

**GEODETICHE SUB-RIEMANNIANE
PER IL COMPLETAMENTO
PERCETTIVO DEI CONTORNI**

Tesi di Laurea in Analisi Matematica

Relatore:
Chiar.ma Prof.
GIOVANNA CITTI

Correlatore:
Chiar.mo Prof.
ALESSANDRO SARTI

Presentata da:
BENEDETTA
FRANCESCHIELLO

Sessione II
Anno Accademico 2010/2011

Introduzione

Uno degli scopi della tesi è studiare il fenomeno di completamento percettivo attraverso il quale il sistema visivo interpola le informazioni esistenti costruendo le unità percettive. In particolare ci concentreremo sulla percezione di contorni soggettivi. Quando guardiamo un'immagine, come quella indicata in figura 1,

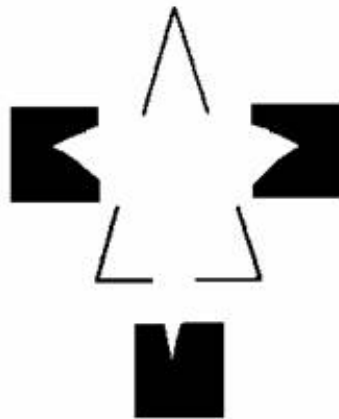


Figura 1: Triangolo di Kanizsa.

il nostro sistema visivo è in grado di percepire non soltanto i contorni effettivamente presenti, ma anche altri contorni, detti contorni soggettivi o illusori. Infatti noi percepiamo molto chiaramente la presenza di una figura bianca ricurva, che copre tre quadrati e un triangolo, benché i bordi di questa figura non siano completamente presenti nell'immagine. Inoltre la regione bianca

ci appare come una superficie in rilievo, più chiara dello sfondo, e si chiama superficie soggettiva perché viene ricostruita dal soggetto che guarda. I suoi bordi si dicono contorni soggettivi.



Figura 2: Un completamento di un'immagine che non ha né una forma nota né una forma geometrica

Ci sono due tipi di completamento che il nostro cervello mette in atto:

- *Il completamento modale*, ovvero quello che avviene quando le parti interpolate dell'immagine sono percepite con la piena modalità della visione e sono fenomenologicamente indistinguibili dallo stimolo reale, come la creazione di contorni illusori.
- *Il completamento amodale*, ovvero quel tipo di processo attraverso cui il cervello ricostruisce i contorni e la configurazione di un oggetto parzialmente occluso.

Questi fenomeni, sono stati diffusamente studiati da Kanizsa in [3] e dagli psicologi della percezione. In particolare è stato osservato che i contorni soggettivi non sono necessariamente rettilinei, e non sono indotti da figure geometriche o figure note come illustrato in figura 2. Sulla base di questi dati sono stati formulati modelli puramente fenomenologici di completamento.

Inizialmente le teorie che hanno modellizzato e studiato le leggi fenomenologiche del completamento percettivo erano espresse in termini di concetti geometrici e minimi del calcolo delle variazioni. In particolare modelli di contorni soggettivi per il completamento delle immagini erano stati ottenuti minimizzando il funzionale delle elastiche di Mumford, ossia

$$\int_{\gamma} (1 + k^2) ds$$

ove k è la curvatura di γ .

Oggi, grazie alle moderne tecniche di neuroimaging, che consentono una descrizione dettagliata della corteccia, è possibile formulare modelli matematici di corteccia visiva, e giustificare su basi neurofisiologiche questi fenomeni percettivi. Presenteremo quindi un modello introdotto da Citti e Sarti in [1], che giustifica il completamento percettivo sulla base della funzionalità della corteccia. La tesi è organizzata in questo modo:

Nel primo capitolo descriveremo l'architettura funzionale della corteccia visiva primaria da un punto di vista neurofisiologico, e ne presenteremo un modello in gruppi di Lie.

Il secondo capitolo sarà invece totalmente dedicato alla Geometria Differenziale, in particolare alle Varietà Riemanniane, e alla geodetiche della struttura, che sono descritte sia come curve ad accelerazione nulla, sia come minimi del funzionale delle lunghezze.

Infine nell'ultimo capitolo mostreremo come è possibile indurre una metrica nella corteccia, le cui geodetiche sono un buon modello per i contorni soggettivi. Il modello verrà poi confrontato con una modifica del classico funzionale delle elastiche.

Indice

Introduzione	i
1 L'architettura funzionale della corteccia visiva	3
1.1 Descrizione Neurofisiologica	3
1.1.1 La Struttura dell'insieme delle cellule semplici	4
1.1.2 La struttura di connettività fra le cellule	5
1.2 Un modello dell'architettura funzionale in gruppi di Lie	6
1.2.1 L'insieme delle cellule semplici modellate come gruppo di rotazione e traslazione	6
1.2.2 Curve integrali e connettività neurale	7
2 Geodetiche in Varietà Riemanniane	11
2.1 Varietà Riemanniane	12
2.2 Geodetiche e Trasporto parallelo	12
2.2.1 Connessioni	13
2.2.2 Derivate covarianti lungo curve	16
2.2.3 Geodetiche	17
2.2.4 Trasporto Parallelo	18
2.2.5 La mappa esponenziale	23
2.3 Geodetiche, lunghezze e distanze	24
2.3.1 Lunghezze di curve	25
2.3.2 Distanza Riemanniana	26
2.3.3 Geodetiche e curve minimizzanti	27

3	Modelli di completamento percettivo	33
3.1	Metrica e geodetiche Riemanniane	33
3.1.1	Una metrica per modellare la corteccia	33
3.1.2	Geodetiche Riemanniane	34
3.2	Geodetiche subriemanniane e contorni soggettivi	37
3.3	Il funzionale delle elastiche modificato	41
3.4	Geodetiche subriemanniane ed elastiche modificate	43
	Bibliografia	45

Elenco delle figure

1	Figura 1.	i
2	Figura 2.	ii
1.1	Figura 3	5
1.2	Figura 4.	9
2.1	Figura 5	14
2.2	Figura 6	19
2.3	Figura 7	21
2.4	Figura 8	24
3.1	Figura 9	39
3.2	Figura 10	40
3.3	Figura 11	41
3.4	Figura 12	42

Capitolo 1

L'architettura funzionale della corteccia visiva

La corteccia visiva primaria, nota anche come V1, è la prima parte del cervello che elabora l'informazione visiva proveniente dalla retina. La sua architettura funzionale è la struttura che presiede e giustifica la sua funzionalità. In questa sezione introdurremo prima una descrizione qualitativa delle strutture base dell'architettura della corteccia visiva primaria, e poi ne forniremo un modello matematico, nel gruppo di rotazione e traslazione.

1.1 Descrizione Neurofisiologica

Dal punto di vista neurofisiologico l'acquisizione del segnale visivo è eseguita nella retina che, dopo un pre-processo, trasmette l'informazione alla corteccia visiva primaria, dove il segnale è nuovamente elaborato. Precisamente ogni cellula riceve input da una regione della corteccia, detta campo recettore e centrata in un intorno di un ben preciso punto (x, y) del campo visivo. Quando il dominio è stimolato da un segnale visivo la cellula risponde generando spikes, e selezionando formalmente l'orientazione dello stimolo retinico. Normalmente il campo recettore è suddiviso in zone "on" nelle quali

le cellule generano spikes rispondendo a un segnale positivo, e in zone “off” caratterizzate da una bassissima risposta delle cellule in presenza di stimolo.

Matematicamente l'intensità di attivazione può essere descritta da una funzione

$$\psi_0 : \mathcal{R} \rightarrow R, \quad (1.1)$$

definita nel piano retinico \mathcal{R} , e chiamata profilo recettore. Ciascuna cellula semplice ha un profilo recettore direzionale che permette la selezione delle orientazioni dello stimolo.

1.1.1 La Struttura dell'insieme delle cellule semplici

L'elaborazione dell'immagine non dipende soltanto dalla configurazione della singola cellula semplice, ma anche dalla struttura funzionale della corteccia, che è organizzata in struttura retinotopica e struttura ipercolonnare.

La Struttura Retinotopica

La struttura retinotopica è una mappa fra retina e corteccia visiva primaria che preserva la topologia retinica. Se identifichiamo la struttura retinica con un piano \mathcal{R} e il livello corticale con M , il segnale corticale può essere dunque descritto con una mappa $q : \mathcal{R} \rightarrow M$, che realizza un isomorfismo. Da qui in poi identificheremo i due piani, e li denoteremo entrambi con M . Tale mappa comporta una semplice deformazione dello stimolo che verrà tralasciato in questo studio.

La Struttura Ipercolonnare

La struttura ipercolonnare organizza le cellule corticali in colonne corrispondenti al parametro di orientazione: ad ogni posizione sulla retina è associata un insieme di cellule, detto ipercolonna, costituito da tutte le cellule sensibili a tutte le possibili orientazioni e centrate nel punto fissato. Formalmente sopra ad ogni punto (x, y) della retina, è presente una intera

ipercolonna di cellule, ciascuna individuata da una orientazione θ . Il suo profilo recettore si otterrà dal profilo recettore ψ_0 della cellula centrata in 0 mediante una traslazione $T_{x,y}$ di un vettore (x, y) e rotazione R_θ di un angolo θ :

$$\psi_{x,y,\theta} = \psi_0 \circ T_{x,y} \circ R_\theta. \quad (1.2)$$

Identificheremo pertanto le cellule semplici mediante i punti (x, y, θ) dove (x, y) denota la posizione bidimensionale sulla retina e $\theta \in [0, \pi]$ indica la direzione cui è sensibile la cellula. La struttura può essere schematicamente descritta dalla figura 1:

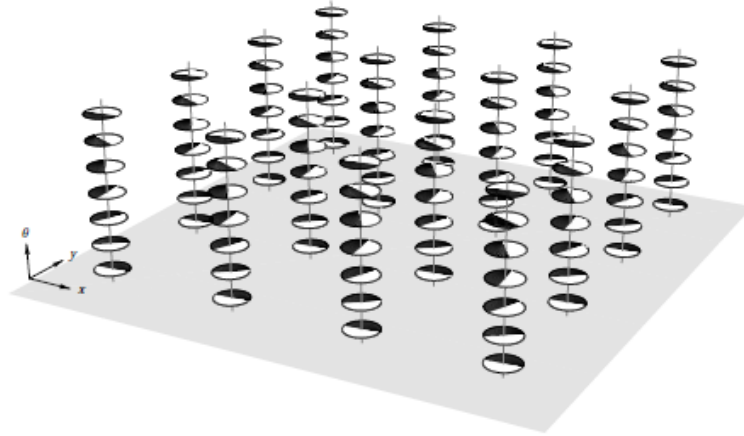


Figura 1.1: Organizzazione di V1 in Ipercolonne. Ogni cerchio è una *odd single cell*, l'angolo degli assi è l'angolo di ricezione. Ogni profilo recettore è ottenuto da un nucleo madre attraverso la traslazione del nucleo di vettore (x_1, y_1) e la rotazione su sè stesso di angolo θ .

1.1.2 La struttura di connettività fra le cellule

La trasmissione del segnale fra le cellule avviene lungo gli assoni delle stesse, che sono lunghe strutture nervose che collegano fra loro le cellule. La fitta rete di questi assoni prende il nome di connettività, e si divide in connettività intracorticale e connettività a lungo raggio, o orizzontale.

1. *Connettività intracorticale*: Il circuito intracorticale opera la selezione dell'orientazione del massimo output dell'ipercolonna in risposta allo stimolo visivo, eliminando tutti gli altri output. Tale meccanismo è chiamato *Eliminazione non massimale o Selezione dell'Orientamento*.
2. *La connettività a lungo raggio* assicura la connessione fra ipercolonne. La connessione orizzontale lega cellule che hanno la stessa orientazione e appartenenti a diverse ipercolonne. Recenti studi hanno rilevato una propagazione del segnale neuronale tramite la connessione corticale. Tale struttura facilita il processo di integrazione, alla base della formazione di contorni regolari e illusori e delle superfici soggettive.

Ovviamente l'architettura funzionale della corteccia visiva è molto più complicata degli schemi delineati, ma ci limitiamo a descrivere quelle strutture che intervengono nel modello di completamento che vogliamo presentare.

1.2 Un modello dell'architettura funzionale in gruppi di Lie

1.2.1 L'insieme delle cellule semplici modellate come gruppo di rotazione e traslazione

Come abbiamo osservato ogni profilo recettore di una cellula semplice può essere espresso nella forma (1.2), ed identificato mediante una rotazione e una traslazione. Il gruppo delle rototraslazioni, denotato come $SE(2)$ è detto gruppo tridimensionale dei movimenti rigidi del piano. Abbiamo denotato con T_{x_1, y_1} la traslazione di vettore (x_1, y_1) e R_θ la rotazione di angolo θ , di matrice:

$$\mathcal{R}_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

Pertanto un elemento generico di $SE(2)$ è della forma $T_{x_1, y_1} \circ R_\theta$ e applicato ad un elemento generico (x, y) :

$$(T_{x_1, y_1} \circ R_\theta) \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} + R_\theta \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$$

Pertanto tutti i profili recettori possono essere visti come

$$\phi_{x_1, y_1, \theta} \begin{pmatrix} x \\ y \end{pmatrix} = \phi_0 \circ (T_{x_1, y_1} \circ R_\theta) \begin{pmatrix} x \\ y \end{pmatrix}.$$

L'insieme dei parametri $g_1 = (x_1, y_1, \theta)$ forma un gruppo con l'operazione indotta dalla composizione $(T_{x_1, y_1} \circ R_\theta) \circ (T_{x_2, y_2} \circ R_{\theta_2})$:

$$g_1 \circ g_2 = (x_1, y_1, \theta) +_R (x_2, y_2, \theta_2) = \left(\left(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + R_\theta \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} \right)^T, \theta + \theta_2 \right)$$

L'operazione $+_R$ verifica gli assiomi dell'operazione di gruppo. Ecco pertanto come l'operazione $+_R$ in $\mathbb{R}^2 \times \mathcal{S}^1$ generi il gruppo delle rototraslazioni, SE(2). Le simmetrie che abbiamo descritto non solo danno l'invarianza per rototraslazioni nella rappresentazione di immagini sulla retina ma ci dicono anche che i segnali inviati sono indipendenti dalla loro posizione o orientamento nello spazio fenomenologico.

1.2.2 Curve integrali e connettività neurale

La struttura di gruppo descritta fino a qui identifica solo l'insieme delle cellule semplici e l'insieme delle loro simmetrie. Se vogliamo studiare l'elaborazione operata dalle cellule e la trasmissione del segnale visivo lungo la connettività neurale, dobbiamo studiare le curve naturali della struttura.

In presenza di uno stimolo visivo la corteccia utilizza, come già detto, il premeccanismo di soppressione non massimale per selezionare ad ogni punto (x, y) la direzione θ dello stimolo al punto (x, y) . Se lo stimolo è rappresentato da una curva $\tilde{\gamma}$ non è restrittivo supporre che $\tilde{\gamma}$ sia parametrizzata d'arco, ovvero $\tilde{\gamma}$ sia della forma

$$\tilde{\gamma}(t) = (x(t), y(t)) \tag{1.3}$$

dove $(\dot{x}(t), \dot{y}(t)) = (\cos(\theta(t)), \sin(\theta(t)))$. Il meccanismo di selettività dell'orientazione seleziona il valore $\theta(t)$, liftando formalmente la curva retinica $\tilde{\gamma}(t)$ ad una curva corticale in $\mathbb{R}^2 \times \mathcal{S}^1$:

$$\gamma(t) = (x(t), y(t), \theta(t)).$$

La curva γ soddisfa:

$$\gamma'(t) = \begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \end{pmatrix} = \begin{pmatrix} \cos\theta \\ \sin\theta \\ \dot{\theta} \end{pmatrix} = X_1 + kX_2 \quad (1.4)$$

se denotiamo $k = \dot{\theta}$ e

$$X_1 = \begin{pmatrix} \cos\theta \\ \sin\theta \\ 0 \end{pmatrix}, X_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad (1.5)$$

Tutte le curve liftate nella corteccia sono quindi curve integrali dei campi X_1 e X_2 , pur essendo immerse in una struttura 3D. In particolare non esistono "curve naturali" per la struttura con derivata nella direzione

$$X_3 = \begin{pmatrix} -\sin\theta \\ \cos\theta \\ 0 \end{pmatrix}, \quad (1.6)$$

che è ortogonale a X_1 e X_2 (La rappresentazione dei vettori segue nella figura successiva).

Se supponiamo che vi sia un minimo errore sperimentale, nell'operazione di lifting possiamo introdurre uno strumento geometrico che consenta di pesare diversamente le lunghezze dei vettori velocità in direzione X_1 , X_2 e in direzione X_3 . Assegnando una lunghezza molto elevata alle curve in direzione X_3 , queste tenderanno a sparire in un processo di minimizzazione. Abbiamo quindi ora bisogno della nozione di metrica Riemanniana.

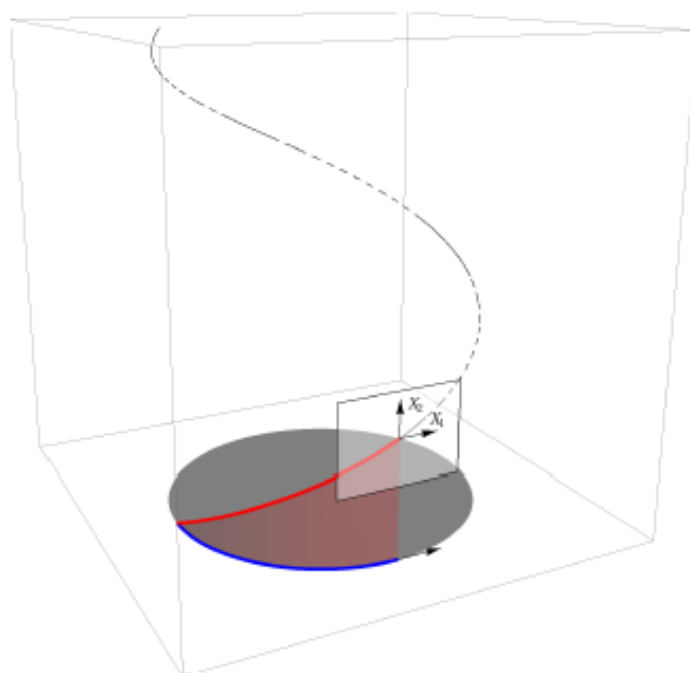


Figura 1.2: Lifting di una curva

Capitolo 2

Geodetiche in Varietà Riemanniane

Il modello che abbiamo in parte descritto nel capitolo precedente, proposto da G.Citti e A.Sarti in [1], modella la corteccia visiva come un Gruppo di Lie. La connettività fra le cellule può essere descritta mediante curve integrali. Per individuare le traiettorie ottimali che possono descrivere i contorni soggettivi, dobbiamo interpretare il gruppo di Lie come varietà e introdurre strumenti di geometria differenziale. Questo permetterà di interpretare i contorni soggettivi come le curve 'naturali' della struttura perché hanno accelerazione nulla.

Per fare questo richiamiamo dapprima la nozione di varietà Riemanniana e poi introdurremo il concetto di geodetica, ovvero generalizzeremo alle varietà Riemanniane il concetto di linee rette dello spazio euclideo. Per farlo ci sono due strade: la prima è quella di portare al contesto delle varietà Riemanniane il fatto che l'accelerazione delle rette risulta identicamente nulla. La seconda è quella di definire le geodetiche come curve che minimizzano la distanza tra due punti. Nelle sezioni 2 e 3 introdurremo le due nozioni e mostreremo che sono equivalenti.

2.1 Varietà Riemanniane

Definizione 2.1. Sia V uno spazio vettoriale. Un *prodotto scalare* su V è una forma bilineare simmetrica definita positiva, cioè un'applicazione $g : V \times V \rightarrow \mathbb{R}$ tale che:

$$(i) \quad g(u_1 + u_2, v) = g(u_1, v) + g(u_2, v) \quad \forall u_1, u_2, v \in V;$$

$$(ii) \quad g(\lambda u, v) = \lambda g(u, v) \quad \forall u, v \in V, \forall \lambda \in \mathbb{R};$$

$$(iii) \quad g(u, v) = g(v, u) \quad \forall u, v \in V.$$

$$(iv) \quad g(u, u) \geq 0 \quad \forall u \in V, \text{ con } g(u, u) = 0 \Leftrightarrow u = 0.$$

Definizione 2.2. Una *metrica Riemanniana* g su una varietà differenziabile M è una funzione che ad ogni punto p di M associa un prodotto scalare g_p , definito sullo spazio tangente $T_p M$, che dipende differenziabilmente da p . In altri termini, per ogni coppia di campi vettoriali X e Y , l'applicazione $p \rightarrow g_p(X_p, Y_p)$ è differenziabile.

Definizione 2.3. Una *Varietà Riemanniana* è una coppia (M, g) , dove M è una varietà differenziabile e g una metrica Riemanniana su M .

2.2 Geodetiche e Trasporto parallelo

Per introdurre la definizione di geodetica come curva a derivata seconda nulla, è necessario introdurre la definizione di derivata del vettore tangente alla curva. Se volessimo infatti definire $\ddot{\gamma}$ differenziando semplicemente $\dot{\gamma}$ rispetto a t , si avrebbe un rapporto incrementale che contiene la differenza tra i vettori $\dot{\gamma}(t)$ e $\dot{\gamma}(t_0)$. Ma chiaramente non è possibile calcolare direttamente questa differenza, perché tali vettori appartengono rispettivamente a $T_{\gamma}(t)$ e $T_{\gamma}(t_0)$, spazi vettoriali isomorfi ma diversi. Pertanto per definire l'accelerazione di una curva su una varietà stretta è necessario introdurre la nozione di derivata di un campo vettoriale $\dot{\gamma}$ lungo la curva. Per fare ciò

è necessario però introdurre connessioni, derivate covarianti e trasporto parallelo, ovvero nozioni di geometria cruciali che ci permetteranno di generalizzare le derivate direzionali dello spazio Euclideo e di definire le derivate direzionali di campi vettoriali.

2.2.1 Connessioni

Dapprima daremo la definizione di fibrato e sezione di fibrato, per poter definire la connessione per differenziare sezioni di fibrati. In un secondo momento adatteremo la definizione al caso dei campi vettoriali lungo curve.

Definizione 2.4. Un *fibrato* è una particolare funzione $\pi : E \rightarrow B$ che si comporta localmente come la proiezione di un prodotto su un *fattore*. È per meglio dire una funzione suriettiva continua fra spazi topologici che è localmente un prodotto. Più precisamente, ogni punto $x \in B$ possiede un intorno aperto U tale che la controimmagine $\pi^{-1}(U)$ è omeomorfa ad un prodotto $U \times F$ (ove F è un fissato spazio topologico) e la π letta su questo prodotto è la proiezione sul primo fattore.

Tale definizione è equivalente a dire che il seguente diagramma commuta:

$$\begin{array}{ccc}
 \pi^{-1}(U) & \xrightarrow{\phi} & U \times F \\
 \downarrow \pi & & \swarrow \text{proj}_1 \\
 U & &
 \end{array}$$

Lo spazio B è detto base, F è chiamata *fibra*, E è lo *spazio totale* e π la *proiezione*.

Esempio 2.1. Il nastro di Moebius è un esempio di fibrato vettoriale. La base B consiste nella circonferenza, la fibra F è un segmento. Dato $x \in B$, U è un piccolo arco (intorno di x sulla circonferenza) e $\pi^{-1}(U)$ è omeomorfo a un quadrato $U \times F$. Globalmente questo non è vero.

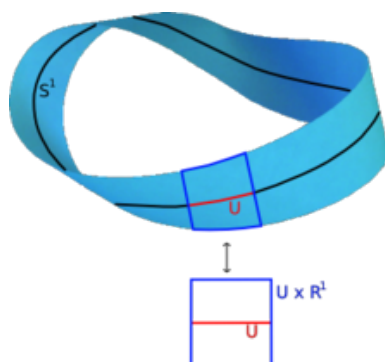


Figura 2.1: Il nastro di Moebius è un esempio di fibrato vettoriale

Ad ogni modo l'idea del fibrato è quella di identificare localmente uno spazio come prodotto. Si dice *fibrato vettoriale* il fibrato la cui fibra F è uno spazio vettoriale. Di fatto un *fibrato vettoriale* è una costruzione che associa ad ogni punto di una varietà topologica (o differenziabile) uno spazio vettoriale.

Definizione 2.5. Sia $\pi : E \rightarrow M$ un fibrato vettoriale su M . Si dice *sezione di E* la mappa $\phi : M \rightarrow E$ tale che $\pi \circ \phi = Id_M$.

Osservazione 1. Si osservi che in topologia differenziale il fibrato tangente $T(M)$ ad una varietà differenziabile M è l'insieme formato dall'*unione disgiunta* di tutti gli spazi tangenti ai punti di M . Questo insieme è dotato di una struttura di varietà differenziabile ed è generalmente visualizzato come fibrato vettoriale, tramite l'applicazione:

$$\pi : T(M) \rightarrow M$$

la cui controimmagine $\pi^{-1}(x) \forall x \in M$ è proprio lo spazio tangente alla varietà nel punto x , $T_x(M)$.

Introdotte queste nozioni siamo pronti per definire la connessione:

Definizione 2.6. Sia $\pi : E \rightarrow M$ un fibrato vettoriale su una varietà M , e sia $\mathcal{E}(M)$ lo spazio delle sezioni lisce di E , e $\mathcal{T}(M)$ l'insieme delle sezioni lisce

di $T(M)$ (ossia lo spazio dei *campi vettoriali* lisci su M). Una *connessione* su E è una mappa

$$\begin{aligned}\nabla : \mathcal{T}(M) \times \mathcal{E}(M) &\longrightarrow \mathcal{E}(M) \\ (X, Y) &\longmapsto \nabla_X Y\end{aligned}$$

tale che:

(i) $\nabla_X Y$ è $C^\infty(M)$ -lineare in X :

$$\nabla_{fX_1 + gX_2} Y = f\nabla_{X_1} Y + g\nabla_{X_2} Y$$

ove f, g sono funzioni.

(ii) $\nabla_X Y$ è \mathbb{R} -lineare in Y :

$$\nabla_X (aY_1 + bY_2) = a\nabla_X Y_1 + b\nabla_X Y_2$$

ove $a, b \in \mathbb{R}$

(iii) ∇ soddisfa le proprietà del prodotto:

$$\nabla_X (fY) = f\nabla_X Y + (Xf)Y$$

ove f è ancora una funzione.

$\nabla_X Y$ è chiamata *derivata covariante di Y nella direzione di X* . Una connessione su $T(M)$ viene detta *connessione lineare* o *connessione su M* .

È in effetti proprio quest'ultimo il caso al quale siamo interessati, ovvero intuitivamente all'idea di avere la possibilità di connettere spazi tangenti vicini. La definizione di *derivata covariante* ci interessa pertanto nel caso in cui tale connessione sia lineare (che sia cioè una connessione su $T(M)$). Più formalmente:

Definizione 2.7. Definiamo *connessione lineare* su M la mappa:

$$\begin{aligned}\nabla : \mathcal{T}(M) \times \mathcal{T}(M) &\rightarrow \mathcal{T}(M) \\ (X, Y) &\longmapsto \nabla_X Y\end{aligned}$$

che soddisfa le proprietà sopra elencate.

È possibile dare l'espressione in componenti di una connessione lineare. Sia E_i , $i \in I$ un riferimento locale per $T(M)$ su un aperto U (solitamente identifichiamo $E_i = \frac{\partial}{\partial x_i}$, ma qui trattiamo il caso generale). Se fissiamo i e j , allora il vettore $\nabla_{E_i} E_j$ sarà esprimibile in termini degli elementi della base con opportuni coefficienti che indicheremo con Γ_{ij}^k :

$$\nabla_{E_i} E_j = \Gamma_{ij}^k E_k$$

Tali funzioni Γ_{ij}^k su U sono chiamate *simboli di Christoffel* di ∇ rispetto al riferimento E_i , e determinano completamente la connessione ∇ su U .

2.2.2 Derivate covarianti lungo curve

Le curve su una varietà M sono definite come delle applicazioni lisce $\gamma : I \rightarrow M$ con $I \subset \mathbb{R}$. Nei casi in cui l'intervallo I è chiuso o semichiuso, si può estendere γ a una curva liscia definita su un intervallo aperto che contenga I , lavorare con questa nuova curva e poi restringersi all'intervallo iniziale. Si può quindi assumere per comodità che γ sia definita su un intervallo aperto.

Definizione 2.8. Un *campo vettoriale lungo una curva* $\gamma : I \rightarrow M$ è una mappa liscia $V : I \rightarrow T(M)$ tale che $V(t) \in T_{\gamma(t)}M$ per ogni $t \in I$. Lo spazio dei campi vettoriali lungo γ verrà indicato con $\mathcal{T}(M)$.

Introduciamo ora un lemma, che non dimostreremo, che ci permette di definire la derivata direzionale di un campo vettoriale lungo una curva.

Lemma 2.2.1. *Sia ∇ una connessione lineare su M . Per ogni curva $\gamma : I \rightarrow M$, ∇ determina un unico operatore:*

$$\mathcal{D}_t : \mathcal{T}(\gamma) \rightarrow \mathcal{T}(\gamma)$$

tale che:

(i) \mathcal{D}_t è lineare su \mathbb{R} :

$$\mathcal{D}_t(aV + bW) = a\mathcal{D}_tV + b\mathcal{D}_tW \text{ per ogni } a, b \in \mathbb{R}$$

(ii) \mathcal{D}_t soddisfa le proprietà del prodotto:

$$\mathcal{D}_t(fV) = \dot{f}V + f\mathcal{D}_tV \text{ per } f \in \mathcal{C}^\infty(I)$$

(iii) Se V è estendibile, allora per ogni estensione \tilde{V} di V ,

$$\mathcal{D}_tV(t) = \nabla_{\dot{\gamma}(t)}\tilde{V}$$

Per ogni $V \in \mathcal{T}(\gamma)$, \mathcal{D}_tV è detta derivata covariante di V lungo γ .

Osservazione 2. Si può dare una scrittura in coordinate della derivata covariante utilizzando i simboli di Christoffel attraverso la formula:

$$\mathcal{D}_tV(t_0) = (\dot{V}^k(t_0) + V^j(t_0)\dot{\gamma}^i(t_0)\Gamma_{ij}^k(\gamma(t_0))) \quad (2.1)$$

2.2.3 Geodetiche

Ora che abbiamo tutti gli strumenti a disposizione, siamo pronti per definire le geodetiche, ovvero le curve che estendono il concetto di retta ad una varietà Riemanniana.

Definizione 2.9. Sia ∇ una connessione lineare su una varietà M , e sia $\gamma : I \rightarrow M$ una curva su M . L'accelerazione di γ è il campo vettoriale $\mathcal{D}_t\dot{\gamma}$ lungo γ . Si dice che γ è una *geodetica* per ∇ se la sua accelerazione è nulla, cioè se

$$\mathcal{D}_t\dot{\gamma} \equiv 0$$

Osservazione 3. Sia (U, ϕ) una carta locale e siano (x^i) delle coordinate su U . Si ha che γ è una geodetica se e soltanto se le sue componenti $\gamma(t) = (x^1(t), \dots, x^n(t))$ soddisfano l'equazione delle geodetiche:

$$\ddot{x}^k(t) + \dot{x}^i(t)\dot{x}^j(t)\Gamma_{ij}^k(x(t)) = 0$$

Si tratta di un sistema di equazioni differenziali ordinarie del secondo ordine, che può essere trasformato in un sistema del primo ordine introducendo delle

variabili ausiliarie v^1, \dots, v^n tali che $v^i = \dot{x}^i$. Effettuando la sostituzione si ottiene:

$$\begin{cases} \dot{v}^k(t) &= -v^i(t)v^j(t)\Gamma_{ij}^k(x(t)) \\ \dot{x}^i(t) &= v^i(t) \end{cases} \quad (2.2)$$

Il *teorema di esistenza e unicità delle geodetiche* ci garantisce appunto l'esistenza e l'unicità di tali curve data una varietà Riemanniana munita di una connessione lineare. Inoltre è proprio da questo teorema che possiamo dedurre l'esistenza di *geodetiche massimali*, definite come segue:

Definizione 2.10. Siano ∇ una connessione su una varietà M , $p \in M$ e $V \in T_pM$. Si dice *geodetica massimale* l'unica geodetica $\gamma : I \rightarrow M$ tale che $\gamma(0) = p$ e $\dot{\gamma}(0) = V$ che non può essere estesa a un intervallo più grande di I . Una geodetica massimale è anche chiamata *geodetica con punto iniziale p e velocità iniziale V* e si indica con γ_V .

Siamo ora pronti per introdurre il concetto di *trasporto parallelo*, che ci permetterà di comprendere come trasportare vettori lungo curve sulle varietà, in modo che restino paralleli rispetto alla connessione.

2.2.4 Trasporto Parallelo

Definizione 2.11. Sia M una varietà munita di una connessione lineare ∇ . Un campo vettoriale V lungo una curva γ è detto *parallelo lungo γ* rispetto a ∇ se $\mathcal{D}_t V \equiv 0$. V è parallelo se è parallelo lungo ogni curva.

Esempio 2.2. Cerchiamo di visualizzare la cosa: la connessione è un oggetto matematico che connette gli spazi tangenti appartenenti a punti diversi di una varietà differenziabile. Tale connessione tra i due spazi tangenti è effettuata sulla base di una curva che li collega. Intuitivamente, la connessione definisce un modo di far scivolare lo spazio tangente lungo la curva. È tale operazione di scivolamento ad essere chiamata trasporto parallelo.

Osservazione 4. A partire dalla definizione di geodetica data deduciamo che una geodetica è caratterizzata come una curva il cui campo vettoriale velocità è parallelo lungo la curva.

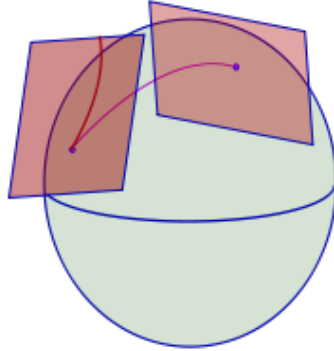


Figura 2.2: Connessione e trasporto parallelo: Una connessione sulla sfera permette di “far scivolare” il piano tangente ad un punto lungo una curva.

Osservazione 5. Un fatto cruciale riguardante i campi vettoriali paralleli è che dato un vettore tangente a un punto di una curva, questo può essere esteso in modo unico a un campo vettoriale parallelo lungo una curva.

Definizione 2.12. Sia $\gamma : I \rightarrow M$ una curva tale che t_0 e $t_1 \in I$, allora il *trasporto parallelo* è una funzione

$$\begin{aligned} P_{t_0 t_1} : T_{\gamma(t_0)}M &\longrightarrow T_{\gamma(t_1)}M \\ V_0 &\longmapsto V(t_1) \end{aligned}$$

dove V è l'estensione parallela di V_0 lungo γ . (ossia V è l'unico campo vettoriale parallelo lungo γ tale che $V(t_0) = V_0$).

Ragioniamo ora su come ottenere la derivata covariante lungo γ a partire dal trasporto parallelo; questo ci serve per farci capire in che senso la connessione *connette* spazi tangenti vicini. Anche stavolta, concentrandoci sul risultato, tralascieremo la dimostrazione.

Proposizione 2.2.2. Sia ∇ una connessione su M e $\gamma : I \rightarrow M$ una curva su M con $t_0 \in I$. Allora:

$$\mathcal{D}_t V(t_0) = \lim_{t \rightarrow t_0} \frac{P_{t_0 t}^{-1} V(t) - V(t_0)}{t - t_0}$$

Riconduciamoci ora al caso delle geodetiche in una varietà Riemanniana M munita di metrica g . Nel caso di una varietà differenziabile è possibile definire diverse connessioni. Nel caso di una varietà Riemanniana lavoreremo invece soltanto con la connessione di *Levi-Civita*, le cui geodetiche risultano compatibili con le proprietà della metrica. Tale connessione infatti è l'unica senza torsione che preserva la metrica, o, formalmente:

Definizione 2.13. Sia (M, g) una varietà Riemanniana. Una connessione ∇ è di *Levi-Civita* se:

- (i) ∇ è *simmetrica*, ossia la sua torsione τ è nulla ($\tau \equiv 0$), ove definiamo come *torsione* la mappa $\tau : \mathcal{T} \times \mathcal{T} \rightarrow \mathcal{T}$, e l'immagine $\tau(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y]$.
- (ii) ∇ preserva la metrica (o è *compatibile* con la metrica). In altre parole: $\nabla_X(g(Y, Z)) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z)$ o, equivalentemente $\nabla_X g = 0$.

Possiamo esprimere le stesse proprietà utilizzando i simboli di Christoffel, e avremo pertanto che la simmetria e la compatibilità si tradurranno come segue:

- (i) $\Gamma_{ij}^k = \Gamma_{ji}^k$
- (ii) la derivata covariante del tensore metrico è nulla: $\nabla_k g_{ij} = 0$

Osservazione 6. La *torsione* è un tensore che misura il grado di torsione degli spazi tangenti lungo una geodetica.

Lemma 2.2.3 (Lemma Fondamentale della geometria Riemanniana). *Sia (M, g) una varietà Riemanniana. Esiste un'unica connessione lineare ∇ su M che sia simmetrica e compatibile con g . Tale connessione è chiamata connessione Riemanniana o connessione di Levi-Civita di g .*

Dimostrazione. Si inizia con la prova dell'unicità supponendo che ∇ sia una connessione simmetrica e compatibile con g . Siano $X, Y, Z \in \mathcal{T}(M)$ dei

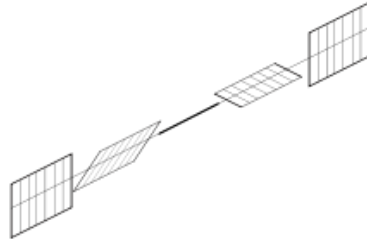


Figura 2.3: Torsione dei piani tangenti lungo una geodetica

campi vettoriali; per la compatibilità con g di ∇ si ha:

$$\begin{aligned} X \langle Y, Z \rangle &= \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X Z \rangle \\ Y \langle Z, X \rangle &= \langle \nabla_Y Z, X \rangle + \langle Z, \nabla_Y X \rangle \\ Z \langle X, Y \rangle &= \langle \nabla_Z X, Y \rangle + \langle X, \nabla_Z Y \rangle \end{aligned}$$

Grazie alla simmetria di ∇ si ha ad esempio:

$$\langle Y, \nabla_X Z \rangle = \langle Y, \nabla_Z X \rangle + \langle Y, [X, Z] \rangle$$

e sostituendo l'ultimo termine di ciascuna delle uguaglianze precedenti si ha:

$$\begin{aligned} X \langle Y, Z \rangle &= \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_Z X \rangle + \langle Y, [X, Y] \rangle \\ Y \langle Z, X \rangle &= \langle \nabla_Y Z, X \rangle + \langle Z, \nabla_X Y \rangle + \langle Z, [Y, X] \rangle \\ Z \langle X, Y \rangle &= \langle \nabla_Z X, Y \rangle + \langle X, \nabla_Y Z \rangle + \langle X, [Z, Y] \rangle \end{aligned}$$

Sommando le prime due equazioni e sottraendo la terza si ottiene

$$X \langle Y, Z \rangle + Y \langle Z, X \rangle - Z \langle X, Y \rangle = 2 \langle \nabla_X Y, Z \rangle + \langle Y, [X, Z] \rangle + \langle Z, [Y, X] \rangle - \langle X, [Z, Y] \rangle,$$

Da cui, risolvendo rispetto a $\langle \nabla_X Y, Z \rangle$

$$\langle \nabla_X Y, Z \rangle =$$

$$\frac{1}{2} (X \langle Y, Z \rangle + Y \langle Z, X \rangle - Z \langle X, Y \rangle - \langle Y, [X, Z] \rangle - \langle Z, [Y, X] \rangle + \langle X, [Z, Y] \rangle) \quad (2.3)$$

Siano ora ∇^1 e ∇^2 due connessioni simmetriche e compatibili con g . Poiché il termine in 2.3 non dipende dalla connessione, si ha che

$$\langle \nabla_X^1 Y - \nabla_X^2 Y, Z \rangle = 0$$

per ogni X, Y, Z . Ma questo può accadere solo se $\nabla_X^1 Y = \nabla_X^2 Y$ per ogni X e Y , cioè solo se $\nabla^1 = \nabla^2$. Per quanto riguarda l'esistenza, basta dimostrare che una connessione Riemanniana esiste in ogni carta, l'unicità assicura che le connessioni così costruite su ciascuna carta si equivalgono sulle intersezioni. Sia $(U, (x^i))$ una carta locale e si applichi la (3.2) ai campi vettoriali $\partial_i (= \frac{\partial}{\partial x^i})$, le cui parentesi di Lie sono nulle. Si ottiene

$$\langle \nabla_{\partial_i} \partial_j, \partial_l \rangle = \frac{1}{2} (\partial_i \langle \partial_j, \partial_l \rangle + \partial_j \langle \partial_l, \partial_i \rangle - \partial_l \langle \partial_i, \partial_j \rangle)$$

Sostituendo nell'equazione precedente le definizioni dei coefficienti della metrica $g_{ij} = \langle \partial_i, \partial_j \rangle$ e dei simboli di Christoffel $\nabla_{\partial_i} \partial_j = \Gamma_{ij}^m \partial_m$ si ha:

$$\Gamma_{ij}^m g_{ml} = \frac{1}{2} (\partial_i g_{jl} + \partial_j g_{il} - \partial_l g_{ij})$$

Si moltiplica infine per la matrice inversa g^{lk} e, poichè $g_{ml} g^{lk} = \delta_m^k$, si ha:

$$\Gamma_{ij}^k = \frac{1}{2} g^{lk} (\partial_i g_{jl} + \partial_j g_{il} - \partial_l g_{ij}) \quad (2.4)$$

Questa formula definisce una connessione su ogni carta, e siccome dalla precedente equazione risulta evidente che $\Gamma_{ij}^k = \Gamma_{ji}^k$, la connessione è simmetrica. Resta da provare soltanto la compatibilità con la metrica:

$$\begin{aligned} \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X Z \rangle &= \\ \frac{1}{2} (X \langle Y, Z \rangle + Y \langle Z, X \rangle - Z \langle X, Y \rangle - \langle Y, [X, Z] \rangle - \langle Z, [Y, X] \rangle + \langle X, [Z, Y] \rangle) + \\ \frac{1}{2} (X \langle Y, Z \rangle + Z \langle Y, X \rangle - Y \langle Z, X \rangle - \langle Z, [X, Y] \rangle - \langle Y, [Z, X] \rangle + \langle X, [Y, Z] \rangle) &= \\ X \langle Y, Z \rangle. \end{aligned}$$

□

Osservazione 7. All'interno della dimostrazione di tale teorema arriviamo alla costruzione della formula (2.4) che ci permette di definire una connessione su ogni carta e di calcolare i simboli di Christoffel della connessione

Riemanniana relativi alla carta scelta. Tale costruzione avviene a partire dalla matrice g_{ik} e le geodetiche rispetto a questa connessione sono chiamate *geodetiche Riemanniane* o semplicemente *geodetiche*.

Definizione 2.14. Sia γ una curva su una varietà Riemanniana, definiamo la *velocità* di γ al tempo t la lunghezza del vettore velocità $|\dot{\gamma}(t)|$. Si dice che γ ha *velocità costante* se $|\dot{\gamma}(t)|$ non dipende da t .

Osservazione 8. Tutte le *geodetiche Riemanniane* sono *curve a velocità costante*, e le isometrie preservano le geodetiche.

2.2.5 La mappa esponenziale

Per quanto visto quando abbiamo trattato le geodetiche su una varietà differenziabile, sappiamo che $\forall p \in M$ esiste un'unica geodetica massimale passante per p con vettore velocità $V \in T_p M$. Possiamo definire pertanto una *mappa dal fibrato tangente* all'insieme delle geodetiche di M , che chiamiamo *mappa esponenziale*, che ci permetterà di associare ad un intorno dello 0 sul piano tangente un intorno del punto p nella varietà. Questa relazione, dipenderà solo dalle geodetiche, e quindi dalla metrica.

Definizione 2.15. Sia ∇ una connessione lineare su M . Il dominio della *mappa esponenziale* è l'insieme:

$\mathcal{E} := \{V \in T(M) \mid \gamma_V \text{ è definita in un intervallo che contiene } [0, 1]\}$. La mappa esponenziale $\exp : \mathcal{E} \rightarrow M$ è definita da

$$\exp(V) = \gamma_V(1)$$

Inoltre per ogni $p \in M$ si pone $\mathcal{E}_p := \mathcal{E} \cap T_p M$ e si chiama *mappa esponenziale ristretta* \exp_p la restrizione a \mathcal{E}_p di \exp .

Proposizione 2.2.4. *La mappa esponenziale gode delle seguenti proprietà:*

- *Ciascun \mathcal{E}_p è stellato rispetto all'origine*
- *Per ogni $V \in T(M)$, la geodetica γ_V è data da: $\gamma_V(t) = \exp(tV) \forall t \in \mathbb{R}$ per cui uno dei due membri è definito.*

- *La mappa esponenziale è liscia.*

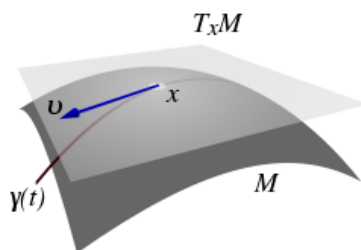


Figura 2.4: La mappa esponenziale associa ad ogni vettore v dello spazio tangente il secondo estremo $\gamma(1)$ dell'unica geodetica $\gamma(t)$ passante per il punto e tangente a v

2.3 Geodetiche, lunghezze e distanze

Lo scopo di questa parte è dimostrare che tutte le curve di minima lunghezza sono geodetiche e che, almeno localmente, tutte le geodetiche sono curve di minima lunghezza. La simmetria della connessione Riemanniana gioca un ruolo fondamentale nelle dimostrazioni di questi risultati, che ci permetteranno di dare un volto alle curve. Per richiamare l'idea che ci accompagna in questo capitolo lo scopo è sempre, dato il funzionale delle elastiche, ricondurci a un problema di ricerca delle geodetiche data una metrica Riemanniana g . Sia d'ora in poi M una varietà differenziabile di dimensione n dotata di una metrica Riemanniana g ; tutte le derivate covarianti e le geodetiche sono da considerarsi rispetto alla connessione Riemanniana di g . Vogliamo innanzitutto dimostrare che una varietà è uno spazio metrico introducendo una distanza definita in termini della lunghezza di una curva.

2.3.1 Lunghezze di curve

Definizione 2.16. sia $\gamma : [a, b] \rightarrow M$ una curva liscia, allora la lunghezza di γ è definita come

$$\mathcal{L}(\gamma) := \int_a^b \|\dot{\gamma}(t)\| dt$$

Osservazione 9. Sappiamo che la lunghezza di una curva è indipendente dalla parametrizzazione scelta, ove per definizione si dice *riparametrizzazione di γ* una curva della forma $\tilde{\gamma} = \gamma \circ \phi$ ove $\phi : [c, d] \rightarrow [a, b]$ è una funzione liscia con inversa liscia.

Definizione 2.17. Una curva *regolare* è una curva liscia $\gamma : I \rightarrow M$ tale che $\dot{\gamma}(t) \neq 0$. Osserviamo che intuitivamente tale condizione esclude curve che abbiano cuspidi o punti angolosi. Notiamo che poiché le geodetiche hanno velocità costante, automaticamente sono classificabili come curve regolari.

Definizione 2.18. Una curva $\gamma : [a, b] \rightarrow M$ è detta curva regolare a tratti se esiste una suddivisione $a = a_0 < a_1 < \dots < a_k = b$ di $[a, b]$ tale che $\gamma|_{[a_{i-1}, a_i]}$ è una curva regolare $\forall i = 1, \dots, k$

Osserviamo che tutte le distanze su varietà Riemanniane verranno misurate su curve come queste, che verranno pertanto chiamate “*curve ammissibili*”. I limiti sinistro e destro del vettore velocità in ciascun a_i sono ben definiti, diversi da zero e non necessariamente uguali. Saranno indicati con:

$$\dot{\gamma}(a_i^-) := \lim_{t \rightarrow a_i^-} \dot{\gamma}(t)$$

$$\dot{\gamma}(a_i^+) := \lim_{t \rightarrow a_i^+} \dot{\gamma}(t)$$

È possibile estendere la definizione di riparametrizzazione partendo dalla definizione della lunghezza di γ come somma delle lunghezze dei tratti $\gamma|_{[a_{i-1}, a_i]}$. In tal caso $\phi : [c, d] \rightarrow [a, b]$ è un omeomorfismo la cui restrizione ai sottointervalli $[c_{i-1}, c_i]$ è una funzione liscia con inversa liscia.

Definizione 2.19. La funzione lunghezza d’arco di una curva ammissibile $\gamma : [a, b] \rightarrow M$ è la funzione $s : [a, b] \rightarrow \mathbb{R}$ definita da:

$$s(t) := \mathcal{L}(\gamma|_{[a, t]}) = \int_a^t \|\dot{\gamma}(u)\| du$$

Per il teorema fondamentale del calcolo integrale s è liscia dove lo è γ , e $\dot{s}(t)$ è uguale alla velocità $\|\dot{\gamma}(u)\|$ di γ .

Osservazione 10. Tra tutte le possibili parametrizzazioni di una curva, quelle a velocità unitaria sono particolarmente importanti e ogni curva ammissibile ne ammette una.

Osservazione 11. Per ogni curva ammissibile $\gamma : [a, b] \rightarrow M$ di lunghezza $\mathcal{L}(\gamma) = l$ esiste un'unica riparametrizzazione $\tilde{\gamma} : [0, l] \rightarrow M$ che conserva l'orientazione tale che $\tilde{\gamma}$ ha velocità unitaria.

Osservazione 12. Abbiamo inoltre che data una curva γ con velocità unitaria, la cui parametrizzazione è definita su $[0, l]$, la funzione lunghezza d'arco di γ è $s(t) = t$, infatti

$$s(t) = \int_0^t \|\dot{\gamma}(u)\| du = t.$$

In tal caso si dice che la curva è *parametrizzata d'arco*.

Definizione 2.20. Siano $\gamma : [a, b] \rightarrow M$ una curva ammissibile e $f \in C^\infty([a, b])$. Si definisce l'integrale di f rispetto alla lunghezza d'arco come:

$$\int_\gamma f ds := \int_a^b f(t) \|\dot{\gamma}(t)\| dt$$

Anche tale nozione è indipendente dalla parametrizzazione.

Definizione 2.21. Una funzione continua $V : [a, b] \rightarrow T(M)$ tale che $V_t \in T_{\gamma(t)}M \forall t$ è detta campo vettoriale lungo γ liscio a tratti se esiste una suddivisione, possibilmente più fine, $a = \tilde{a}_0 < \tilde{a}_1 < \dots < \tilde{a}_m = b$ tale che V sia liscio su ciascun sottointervallo $[\tilde{a}_{i-1}, \tilde{a}_i]$. Dato un qualunque vettore $V_a \in T_{\gamma(a)}M$, esiste un unico campo vettoriale liscio a tratti parallelo a V_a lungo γ . È sufficiente trasportare parallelamente V_a lungo il primo tratto di γ fino a $\gamma(a_1)$, e così via.

2.3.2 Distanza Riemanniana

Essendo le geodetiche curve che minimizzano la distanza fra due punti, è necessario introdurre la nozione di distanza su una varietà Riemanniana:

Definizione 2.22. Sia M una varietà Riemanniana connessa. Per ogni coppia di punti p e $q \in M$ si definisce la *distanza Riemanniana* $d(p, q)$ come:

$$d(p, q) := \inf\{\mathcal{L}(\gamma) \mid \gamma \text{ è una curva regolare a tratti con } \gamma(a) = p \text{ e } \gamma(b) = q\}$$

Per definizione di connessione, proprietà equivalente alla connessione per archi negli spazi topologici sufficientemente regolari, abbiamo la buona definizione dell'esistenza di un cammino continuo $c : [a, b] \rightarrow M$. Per definizione di compattezza, sappiamo che c'è una suddivisione di $[a, b]$ tale che $c([a_{i-1}, a_i])$ è contenuto in una singola carta $\forall i$. Pertanto sostituendo ogni segmento con un cammino liscio otteniamo una curva ammissibile da p a q , da cui la ben definizione della distanza.

Il seguente lemma, di cui tralascieremo la dimostrazione, è cruciale per introdurci in uno spazio metrico, ambiente nel quale vogliamo lavorare al modello.

Lemma 2.3.1. *Con la distanza d appena definita, ogni varietà Riemanniana connessa è uno spazio metrico la cui topologia indotta è proprio la topologia della varietà.*

2.3.3 Geodetiche e curve minimizzanti

In questa sezione vogliamo arrivare a dimostrare, per concludere l'approfondimento teorico al quale dobbiamo far fronte per costruire il modello, che tutte le curve di minima lunghezza sono geodetiche e che, almeno localmente, tutte le geodetiche sono curve di minima lunghezza. Per fare ciò passeremo attraverso l'introduzione del concetto di *famiglia ammissibile* e di alcuni teoremi riguardanti l'argomento. Iniziamo con il definire il concetto di *curva ammissibile minimizzante*:

Definizione 2.23. Una curva ammissibile $\gamma : [a, b] \rightarrow M$ su una varietà Riemanniana si dice minimizzante se, data una qualsiasi altra curva ammissibile $\tilde{\gamma}$ con gli stessi estremi si ha che $\mathcal{L}(\gamma) = \mathcal{L}(\tilde{\gamma})$, ovvero se e soltanto se $d(\gamma(a), \gamma(b)) = \mathcal{L}(\gamma)$.

In modo da dimostrare che tutte le curve minimizzanti sono geodetiche, consideriamo \mathcal{L} come un funzionale sull'insieme delle curve ammissibili su M . Di conseguenza cercare le curve minimizzanti è equivalente a cercare il minimo di questo funzionale. Per i nostro scopi, notiamo che se γ è una curva minimizzante e Γ_s è una famiglia di curve ammissibili con stessi estremi tali che $\mathcal{L}(\Gamma_s)$ è una mappa differenziabile di s e $\Gamma_0 = \gamma$, allora la derivata rispetto a s di $\mathcal{L}(\Gamma_s)$ deve essere nulla in $s = 0$, poiché lì troviamo un minimo del funzionale.

Diamo ora la definizione di *famiglia ammissibile*:

Definizione 2.24. Una famiglia ammissibile di curve è una mappa $\Gamma : (-\varepsilon, \varepsilon) \times [a, b] \rightarrow M$ che sia liscia su ogni rettangolo della forma $(-\varepsilon, \varepsilon) \times [a_{i-1}, a_i]$ per qualche suddivisione $a = a_0 < \dots < a_k = b$, e tale che $\Gamma_s(t) := \Gamma(s, t)$ sia una curva ammissibile per ogni $s \in (-\varepsilon, \varepsilon)$. Se Γ è una famiglia ammissibile, un *campo vettoriale lungo* Γ è una mappa continua $V : (-\varepsilon, \varepsilon) \times [a, b] \rightarrow T(M)$ tale che $V(s, t) \in T_{\Gamma(s,t)}M$ per ogni (s, t) e tale che $V|_{(-\varepsilon, \varepsilon) \times [\tilde{a}_{i-1}, \tilde{a}_i]}$ sia liscia per una suddivisione $a = \tilde{a}_0 < \dots < \tilde{a}_m = b$.

Osserviamo che ogni famiglia ammissibile Γ definisce due gruppi di curve:

- *le curve principali* $\Gamma_s(t) = \Gamma(s, t)$ definite su $[a, b]$ ponendo $s = \text{costante}$ (fissando s).
- *le curve trasverse* $\Gamma^t(s) = \Gamma(s, t)$ definite su $(-\varepsilon, \varepsilon)$ fissando t .

Si ha inoltre che le curve trasverse sono lisce su $(-\varepsilon, \varepsilon) \forall t$, mentre le curve principali in generale sono solo regolari a tratti.

Quando Γ è liscio, i vettori tangenti a queste due famiglie di curve sono esempi di campi vettoriali lungo Γ . Li indichiamo con:

$$\partial_t \Gamma(s, t) := \frac{d}{dt} \Gamma_s(t) \quad \text{e} \quad \partial_s \Gamma(s, t) := \frac{d}{ds} \Gamma^{(t)}(s)$$

$\partial_s \Gamma$ è continuo su tutto il rettangolo $(-\varepsilon, \varepsilon) \times [a, b]$: da un lato il suo valore sul segmento $(-\varepsilon, \varepsilon) \times a_i$ dipende solo dai valori di Γ su quel segmento in quanto la derivata è fatta solo rispetto a s ; d'altro canto è continuo su ogni rettangolo

$(-\varepsilon, \varepsilon) \times [a_{i-1}, a_i]$ e $(-\varepsilon, \varepsilon) \times [a_i, a_{i+1}]$ quindi i limiti destro e sinistro per $t = a_i$ devono essere uguali. Quindi $\partial_s \Gamma$ è sempre un campo vettoriale lungo Γ . Dato un campo vettoriale V lungo Γ , possiamo calcolare la derivata covariante di V sia lungo le curve principali che lungo quelle traverse. I campi vettoriali risultanti lungo Γ si indicano rispettivamente con $\mathcal{D}_t V$ e $\mathcal{D}_s V$. Un ingrediente chiave nella dimostrazione che le curve minimizzanti sono geodetiche è la simmetria della connessione Riemanniana. Il seguente lemma è cruciale per le prove a seguire:

Lemma 2.3.2 (Lemma di simmetria). *Sia $\Gamma : (-\varepsilon, \varepsilon) \times [a, b] \rightarrow M$ una famiglia ammissibile di curve su una varietà Riemanniana. Su ogni rettangolo $(-\varepsilon, \varepsilon) \times [a_{i-1}, a_i]$ su cui Γ è liscia:*

$$\mathcal{D}_s \partial_t \Gamma = \mathcal{D}_t \partial_s \Gamma$$

Dimostrazione. Facendo il conto nelle coordinate locali (x^i) intorno a ogni punto $\Gamma(s_0, t_0)$ e scrivendo le componenti di Γ come $\Gamma(s, t) = (x^1(s, t), \dots, x^n(s, t))$ si ha:

$$\partial_t \Gamma = \frac{\partial x^k}{\partial t} \partial_k \quad \text{e} \quad \partial_s \Gamma = \frac{\partial x^k}{\partial s} \partial_k$$

Usando la scrittura in coordinate della derivata covariante lungo una curva data dalla formula (2.1) si ottiene:

$$\begin{aligned} \mathcal{D}_s \partial_t \Gamma &= \left(\frac{\partial^2 x^k}{\partial s \partial t} + \frac{\partial x^i}{\partial t} \frac{\partial x^j}{\partial s} \Gamma_{ij}^k \right) \partial_k \\ \mathcal{D}_t \partial_s \Gamma &= \left(\frac{\partial^2 x^k}{\partial t \partial s} + \frac{\partial x^i}{\partial s} \frac{\partial x^j}{\partial t} \Gamma_{ji}^k \right) \partial_k \end{aligned}$$

Invertendo i ruoli di i e j nella seconda uguaglianza e usando la condizione di simmetria $\Gamma_{ji}^k = \Gamma_{ij}^k$ si osserva che le due espressioni sono uguali. \square

Definizione 2.25. Sia $\gamma : [a, b] \rightarrow M$ una curva ammissibile, una *variazione di γ* è una famiglia ammissibile Γ tale che $\Gamma_0(t) = \gamma(t)$ per ogni $t \in [a, b]$. Questa è chiamata *variazione propria* o *variazione con estremi fissati* se si

ha anche $\Gamma_s(a) = \gamma(a)$ e $\Gamma_s(b) = \gamma(b)$ per ogni s . Sia Γ una variazione di γ , il *campo variazionale* di Γ è il campo vettoriale $V(t) = \partial_s \Gamma(0, t)$ lungo γ . Un campo vettoriale V lungo γ si dice *proprio* se $V(a) = V(b) = 0$. Un campo vettoriale di una variazione propria è proprio.

Lemma 2.3.3. *Siano γ una curva ammissibile e V un campo di vettori lungo γ , allora esiste una variazione Γ di γ che ha V come campo variazionale. Inoltre se V è proprio si può trovare Γ propria.*

Dimostrazione. Sia $\Gamma(s, t) = \exp(sV(t))$. Per la compattezza di $[a, b]$ esiste $\varepsilon > 0$ tale che Γ è definito su $(-\varepsilon, \varepsilon) \times [a, b]$. Chiaramente Γ è continua su tutto il dominio e liscia su $(-\varepsilon, \varepsilon) \times [a_{i-1}, a_i]$ per ogni sottointervallo $[a_{i-1}, a_i]$ su cui V è liscia. Per le proprietà della mappa esponenziale il campo variazionale di Γ è V . Inoltre se $V(a) = V(b) = 0$, si ha che $\Gamma(s, a) \equiv \gamma(a)$ e $\Gamma(s, b) \equiv \gamma(b)$, quindi Γ è propria. \square

Possiamo ora calcolare un'espressione per la derivata del funzionale delle lunghezze lungo una variazione propria; tradizionalmente, tale derivata di un funzionale che ha per dominio un insieme di funzioni è detta *variazione prima*.

Proposizione 2.3.4. *Sia $\gamma : [a, b] \rightarrow M$ una curva ammissibile con velocità unitaria, Γ una variazione propria di γ , e V il suo campo variazionale. Allora:*

$$\frac{d}{ds} \mathcal{L}(\Gamma_s) \Big|_{s=0} = - \int_a^b \langle V, \mathcal{D}_t \dot{\gamma} \rangle dt - \sum_{i=1}^{k-1} \langle V(a_i), \Delta_i \dot{\gamma} \rangle \quad (2.5)$$

dove $\Delta_i \dot{\gamma} = \dot{\gamma}(a_i^+) - \dot{\gamma}(a_i^-)$ è il salto del vettore tangente $\dot{\gamma}$ in a_i .

Dimostrazione. Si scriva

$$T(s, t) = \partial_s \Gamma(s, t) \quad S(s, t) = \partial_t \Gamma(s, t)$$

Su ogni sottointervallo $[a_{i-1}, a_i]$ su cui Γ è liscio, poichè il dominio di integrazione è compatto, si può differenziare sotto il segno di integrale e si

ottiene:

$$\begin{aligned}
\frac{d}{ds}\mathcal{L}(\Gamma_s|_{[a_{i-1}, a_i]}) &= \int_{a_{i-1}}^{a_i} \frac{\partial}{\partial s} \langle T, T \rangle^{1/2} dt \\
&= \int_{a_{i-1}}^{a_i} \frac{\partial}{\partial s} \langle T, T \rangle^{-1/2} 2 \langle \mathcal{D}_s T, T \rangle dt \\
&= \int_{a_{i-1}}^{a_i} \frac{1}{|T|} \langle \mathcal{D}_t S, T \rangle dt
\end{aligned}$$

dove si è usato il lemma di Simmetria nell'ultimo passaggio. Fissando $s = 0$ e notando che $S(0, t) = V(t)$ e $T(0, t) = \dot{\gamma}(t)$ (che ha lunghezza unitaria), si ha:

$$\begin{aligned}
\frac{d}{ds}\mathcal{L}(\Gamma_s|_{[a_{i-1}, a_i]}) &= \int_{a_{i-1}}^{a_i} \langle \mathcal{D}_t V, \dot{\gamma} \rangle dt \\
&= \int_{a_{i-1}}^{a_i} \left(\frac{d}{dt} \langle V, \dot{\gamma} \rangle - \langle V, \mathcal{D}_t \dot{\gamma} \rangle \right) dt \\
&= \langle V(a_i), \dot{\gamma}(a_i^-) \rangle - \langle V(a_{i-1}), \dot{\gamma}(a_{i-1}^+) \rangle - \int_{a_{i-1}}^{a_i} \langle V, \mathcal{D}_t \dot{\gamma} \rangle dt
\end{aligned}$$

Infine, sommando in i e notando che $V(a_0) = V(a_k) = 0$ in quanto Γ è una variazione propria, si ottiene la tesi. \square

Poiché ogni curva ammissibile ha una parametrizzazione a velocità unitaria e la lunghezza è indipendente dalla parametrizzazione, è utile ma non è restrittivo richiedere, come nella proposizione precedente, che γ abbia velocità unitaria.

Teorema 2.3.5. *Ogni curva minimizzante è una geodetica quando è data una parametrizzazione a velocità unitaria.*

Dimostrazione. Sia $\gamma : [a, b] \rightarrow M$ minimizzante e con velocità unitaria, e sia $a = a_0 < \dots < a_k = b$ una suddivisione tale che γ sia liscia su $[a_{i-1}, a_i]$. Sia Γ una qualunque variazione propria di γ , siccome γ è minimizzante $d\mathcal{L}(\Gamma_s)/ds(0) = 0$. Ogni campo vettoriale V proprio lungo γ è il campo variazionale di una variazione propria, quindi il termine a destra della 2.5 è nullo per V .

Si vuole dimostrare che $\mathcal{D}_t\dot{\gamma} = 0$ su ogni sottointervallo $[a_{i-1}, a_i]$, così si avrebbe che γ è una geodetica all'interno di ciascuno di essi. Si considera una funzione $\phi \in \mathcal{C}^\infty$ tale che $\phi > 0$ su (a_{i-1}, a_i) e $\phi = 0$ altrove. Allora la 2.5 con $V = \phi\mathcal{D}_t\dot{\gamma}$ diventa

$$0 = - \int_{a_{i-1}}^{a_i} \phi |\mathcal{D}_t\dot{\gamma}|^2 dt.$$

Poiché l'integrando è non negativo si ha $\mathcal{D}_t\dot{\gamma} = 0$ su ogni sottointervallo. Si vuole ora provare che $\Delta_i\dot{\gamma} = 0$, cioè che γ non ha spigoli. Per ogni $0 \leq i \leq k$ si considera un campo vettoriale lungo γ tale che $V(a_i) = \Delta_i\dot{\gamma}$ e $V(a_j) = 0$ per $j \neq i$. Allora la formula precedentemente numerata si riduce a $-|\Delta_i\dot{\gamma}|^2 = 0$. Dunque $\dot{\gamma}$ è continua; per l'unicità delle geodetiche tangenti a una certa direzione si ha che $\gamma|_{[a_i, a_{i+1}]}$ è la continuazione di $\gamma|_{[a_{i-1}, a_i]}$ e quindi γ è liscia ed è una geodetica dappertutto. \square

In realtà con questo teorema abbiamo fatto un passo importante, ossia per provare che γ era una geodetica non abbiamo usato il fatto che fosse una curva minimizzante, ma solo il fatto che fosse un punto critico di \mathcal{L} , cioè che:

$$\frac{d\mathcal{L}(\Gamma_s)}{ds}(0) = 0$$

per ogni variazione propria di Γ_s di γ . Pertanto il teorema precedente implica il seguente corollario:

Corollario 2.3.6. *Una curva ammissibile con velocità unitaria è un punto critico di \mathcal{L} se e solo se è una geodetica.*

Dimostrazione. Per provare che γ è una geodetica sapendo che è un punto critico di \mathcal{L} , si segue la dimostrazione del teorema precedente. Viceversa, se γ è una geodetica, il primo addendo della formula che precedentemente non abbiamo numerato si annulla perchè $\mathcal{D}_t\dot{\gamma} = 0$, il secondo perchè $\dot{\gamma}$ non ha salti. \square

Questi ultimi teoremi ci hanno permesso di dimostrare l'equivalenza fra il calcolo delle geodetiche su una varietà Riemanniana munita di una metrica g e il calcolo del minimo del funzionale delle lunghezze.

Capitolo 3

Modelli di completamento percettivo

3.1 Metrica e geodetiche Riemanniane

3.1.1 Una metrica per modellare la corteccia

Con le nozioni di geometria Riemanniana introdotte, un buon modello di metrica che consenta come richiesto alla fine del primo capitolo di pesare diversamente le lunghezze dei vettori velocità lungo le direzioni X_1 , X_2 e X_3 , introdotti in (1.5) e (1.6) sarà una metrica definita sul tangente a $\mathbb{R}^2 \times \mathcal{S}^1$ in (x, y, θ) che ad ogni vettore

$$v = \alpha_1 X_1 + \alpha_2 X_2 + \alpha_3 X_3$$

assegna la lunghezza:

$$\|v\| = \sqrt{\alpha_1^2 + \alpha_2^2 + \frac{\alpha_3^2}{\varepsilon^2}}.$$

Questo significa assegnare una metrica, che ha la matrice seguente, rispetto alla base del tangente X_1, X_2, X_3

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{\varepsilon^2} \end{pmatrix}$$

In coordinate Euclidee la metrica diviene:

$$\begin{pmatrix} \cos(\theta) & 0 & -\sin(\theta) \\ \sin(\theta) & 0 & \cos(\theta) \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{\varepsilon^2} \end{pmatrix} \begin{pmatrix} \cos(\theta) & \sin(\theta) & 0 \\ 0 & 0 & 1 \\ -\sin(\theta) & \cos(\theta) & 0 \end{pmatrix} =$$

$$\begin{pmatrix} \cos^2 \theta + \frac{\sin^2 \theta}{\varepsilon^2} & (1 - \frac{1}{\varepsilon^2}) \sin \theta \cos \theta & 0 \\ (1 - \frac{1}{\varepsilon^2}) \sin \theta \cos \theta & \sin^2 \theta + \frac{\cos^2 \theta}{\varepsilon^2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

La matrice associata all'inversa della metrica risulterà allora

$$\begin{pmatrix} \cos^2 \theta + \varepsilon^2 \sin^2 \theta & (1 - \varepsilon^2) \sin \theta \cos \theta & 0 \\ (1 - \varepsilon^2) \sin \theta \cos \theta & \sin^2 \theta + \varepsilon^2 \cos^2 \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Osserviamo che è stato applicato quindi un peso ε lungo la direzione X_3 .

3.1.2 Geodetiche Riemanniane

Scelta una metrica per la corteccia, modelliamo i contorni soggettivi come le geodetiche della metrica. Con il programma *Mathematica* a partire dalla metrica si possono calcolare i simboli di Christoffel e le equazioni delle geodetiche, risolvendo per i primi la formula:

$$\Gamma_{\mu\nu}^{\lambda} = \frac{1}{2} g^{\lambda\sigma} (\partial_{\mu} g_{\sigma\nu} + \partial_{\nu} g_{\sigma\mu} - \partial_{\sigma} g_{\mu\nu})$$

che discende direttamente dalla dimostrazione del Lemma 2.2.3. Pertanto i simboli di Christoffel ottenuti sono i seguenti:

$$\begin{aligned}
\Gamma_{13}^1 &= -\frac{(-1 + \varepsilon^4) \sin 2\theta}{4\varepsilon^2} \\
\Gamma_{13}^2 &= \frac{(-1 + \varepsilon^2)(1 - \varepsilon^2 + (1 + \varepsilon^2) \cos 2\theta)}{4\varepsilon^2} \\
\Gamma_{23}^1 &= \frac{(-1 + \varepsilon^2)(-1 + \varepsilon^2 + (1 + \varepsilon^2) \cos 2\theta)}{4\varepsilon^2} \\
\Gamma_{23}^2 &= \frac{(-1 + \varepsilon^4) \sin 2\theta}{4\varepsilon^2} \\
\Gamma_{31}^1 &= \frac{(-1 + \varepsilon^2) \cos \theta \sin \theta}{\varepsilon^2} \\
\Gamma_{32}^1 &= -\frac{(-1 + \varepsilon^2) \cos 2\theta \sin \theta}{2\varepsilon^2} \\
\Gamma_{32}^2 &= \left(-1 + \frac{1}{\varepsilon^2}\right) \cos \theta \sin \theta
\end{aligned}$$

e si osservi che poiché la connessione affine è simmetrica rispetto alla permutazione degli ultimi due indici, è stato sufficiente indicare solo le componenti indipendenti.

Una volta calcolati i simboli di Christoffel le componenti delle equazioni delle geodetiche sono calcolate utilizzando la formula:

$$\frac{du^\alpha}{dt} = -\Gamma_{\beta\gamma}^\alpha u^\beta u^\gamma$$

ove le u^α sono le componenti della velocità. Pertanto nel nostro caso i risultati ottenuti sono i seguenti:

$$\begin{aligned}
\frac{d}{dt}u^1 &= \frac{((-1 + \varepsilon^4) \sin 2\theta u^1 + ((-1 + \varepsilon^2)^2 - (-1 + \varepsilon^4) \cos 2\theta)u^2)u^3}{2\varepsilon^2} \quad (3.1) \\
\frac{d}{dt}u^2 &= -\frac{(-1 + \varepsilon^2)((-1 + \varepsilon^2)u^1 + (1 + \varepsilon^2) \cos 2\theta u^1 + (1 + \varepsilon^2) \sin 2\theta)u^2)u^3}{2\varepsilon^2} \\
\frac{d}{dt}u^3 &= \frac{(-1 + \varepsilon^2)(\cos 2\theta u^1 u^2 - \cos \theta \sin \theta((u^1)^2 - (u^2)^2))}{\varepsilon^2}
\end{aligned}$$

che messe a sistema con:

$$\begin{cases} \frac{d}{dt}x = u^1 \\ \frac{d}{dt}y = u^2 \\ \frac{d}{dt}\theta = u^3 \end{cases}$$

ci danno esattamente il sistema di equazioni differenziali del secondo ordine visto nella (2.2) che, risolto, ci dà le geodetiche ricercate.

Possiamo scrivere queste equazioni in termini di variabili più naturali. Un sistema di riferimento naturale sul piano tangente ed ortonormale è X_1 , X_2 , ϵX_3 . Pertanto introduciamo le coordinate nelle direzioni dei campi:

$$\dot{\gamma}_{X_1} = \dot{x} \cos(\theta) + \dot{y} \sin \theta \quad \dot{\gamma}_{X_2} = \dot{\theta} \quad \dot{\gamma}_{X_3} = \frac{1}{\epsilon^2}(-\dot{x} \sin \theta + \dot{y} \cos \theta).$$

Sostituendo al secondo membro si ha:

$$\begin{cases} \frac{d}{dt} u^1 = \cos(\theta) \gamma_{X_3} u^3 - u^2 u^3 + \epsilon^2 \sin(\theta) \gamma_{X_1} u^3 \\ \frac{d}{dt} u^2 = \sin(\theta) \gamma_{X_3} u^3 + u^1 u^3 - \epsilon^2 \cos(\theta) \gamma_{X_1} u^3 \\ \frac{d}{dt} u^3 = (\epsilon^2 - 1) \gamma_{X_1} \gamma_{X_3} \end{cases}$$

Volendo esprimere anche le prime due in termini di γ_{X_1} , possiamo moltiplicare la prima per $\cos(\theta)$ e la seconda per $\sin(\theta)$:

$$\cos(\theta) \frac{d}{dt} u^1 + \sin(\theta) \frac{d}{dt} u^2 = (1 - \epsilon^2) u^3 \gamma_{X_3}$$

moltiplicando la prima per $-\sin(\theta)$ e la seconda per $\cos(\theta)$:

$$-\sin(\theta) \frac{d}{dt} u^1 + \cos(\theta) \frac{d}{dt} u^2 = (1 - \epsilon^2) u^3 \gamma_{X_1}$$

Da qui si ottiene immediatamente:

$$\gamma_{\dot{X}_1} = u^3 \gamma_{X_3}$$

$$\gamma_{\dot{X}_3} = -u^3 \gamma_{X_1}$$

che sostituite nelle equazioni precedenti ci danno le equazioni:

$$\begin{aligned} \frac{d}{dt}(-\cos \theta \gamma_{X_1} + \sin \theta \gamma_{X_3}) &= 0 \\ \frac{d}{dt}(-\sin \theta \gamma_{X_1} - \cos \theta \gamma_{X_3}) &= 0 \\ -\frac{d}{dt}(\gamma_{X_2}) + (\epsilon^2 - 1) \gamma_{X_1} \gamma_{X_3} &= 0 \end{aligned} \tag{3.2}$$

Dalle equazioni precedenti segue che esistono costanti ξ_1 , ξ_2 tali che

$$\cos(\theta) \gamma_{X_1} - \sin(\theta) \gamma_{X_3} = \xi_1$$

$$\sin(\theta)\gamma_{X_1} + \cos(\theta)\gamma_{X_3} = \xi_2$$

Quindi

$$\gamma_{X_1} = \xi_1 \cos(\theta) + \xi_2 \sin(\theta) \quad (3.3)$$

$$\gamma_{X_3} = -\xi_1 \sin(\theta) + \xi_2 \cos(\theta)$$

E poiché $\gamma_{X_2} = \theta'$, allora dall'ultima delle equazioni (3.2) si ottiene

$$\theta'' = (1 - \epsilon^2)(\xi_1 \cos(\theta) + \xi_2 \sin(\theta))(-\xi_1 \sin(\theta) + \xi_2 \cos(\theta)). \quad (3.4)$$

Abbiamo quindi disaccoppiato le variabili, e questa equazione si può adesso risolvere facilmente per determinare θ . Nota la funzione θ dalle espressioni di γ_{X_1} , γ_{X_3} , si determinano equazioni per \dot{x} e \dot{y} , che mettiamo a sistema con la (3.4):

$$\begin{cases} \dot{x} = \cos(\theta)(\xi_1 \cos(\theta) + \xi_2 \sin(\theta)) - \sin(\theta)\epsilon^2(-\xi_1 \sin(\theta) + \xi_2 \cos(\theta)) \\ \dot{y} = \sin(\theta)(\xi_1 \cos(\theta) + \xi_2 \sin(\theta)) + \cos(\theta)\epsilon^2(-\xi_1 \sin(\theta) + \xi_2 \cos(\theta)) \\ \theta'' = (1 - \epsilon^2)(\xi_1 \cos(\theta) + \xi_2 \sin(\theta))(-\xi_1 \sin(\theta) + \xi_2 \cos(\theta)) \end{cases} \quad (3.5)$$

Vedremo fra poco perchè abbiamo sviluppato i conti fino a questo punto.

3.2 Geodetiche subriemanniane e contorni soggettivi

La metrica che abbiamo introdotto è solo un'approssimazione del comportamento reale della corteccia. Infatti avevamo verificato che tutte le curve liftate nella corteccia hanno l'espressione

$$\gamma' = X_1 + kX_2,$$

e quindi dovremmo passare al limite per $\epsilon \rightarrow 0$ per ottenere un modello più efficace del processo di lifting. È chiaro che passando al limite nel problema precedente non si ottiene una metrica Riemanniana, ma una metrica

totalmente degenerare. Infatti la metrica non è piú definita, mentre la matrice inversa della metrica diviene al limite per ϵ che tende a 0:

$$(g^{ij}) = \begin{pmatrix} \cos^2(\theta) & \cos(\theta)\sin(\theta) & 0 \\ \cos(\theta)\sin(\theta) & \cos^2(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Anche i simboli di Christoffel non sono definiti al limite per ϵ che tende a 0. Si può invece provare che le geodetiche non scoppiano per ϵ che tende a 0, ma hanno limite finito. Queste curve sono le candidate naturali ad essere considerate le geodetiche nella metrica limite, e si dicono geodetiche subriemanniane.

Vediamo un grafico che le illustra, ricavato con il programma *Mathematica*. Le equazioni sono state ottenute considerando quelle in (3.5): si calcola il limite per $\epsilon \rightarrow 0$ e applicando la sostituzione diretta della due costanti:

$$\xi_1 = \rho \cos(\beta), \quad \xi_2 = \rho \sin(\beta) \quad (3.6)$$

abbiamo:

$$\begin{cases} \dot{x} = \cos \theta (\rho \cos \beta \cos \theta) + \rho \sin \beta \sin \theta \\ \dot{y} = \sin \theta (\rho \cos \beta \cos \theta) + \rho \sin \beta \sin \theta \\ \theta'' = (\rho \cos \beta \cos \theta + \rho \sin \beta \sin \theta)(-\rho \cos \beta \sin \theta + \rho \sin \beta \cos \theta) \end{cases} \quad (3.7)$$

Si sono infine applicate le formule di somma e sottrazione del coseno ottenendo:

$$\begin{cases} \dot{x} = \rho \cos \theta \cos(\theta - \beta) \\ \dot{y} = \rho \sin \theta \cos(\theta - \beta) \\ \theta'' = (\rho \cos(\theta - \beta))(\rho \sin(\theta - \beta)) \end{cases} \quad (3.8)$$

Da cui imponendo $\beta = 0$ per la proprietà di invarianza per Rototraslazioni si ottiene:

$$\begin{cases} \dot{x} = \rho \cos^2 \theta \\ \dot{y} = \rho \sin \theta \cos \theta = \frac{1}{2} \rho \sin(2\theta) \\ \dot{\theta} = p_\theta \\ \theta'' = \dot{p}_\theta = \frac{1}{2} \rho^2 \sin(2\theta) \end{cases} \quad (3.9)$$

Nella pagina seguente il codice che le ha calcolate a partire dall'ultimo sistema, imponendo $x(0) = y(0) = 0$ e $\theta_0 = 0$.

```

In[1]= $HistoryLength = 1;
In[2]= Clear[geo]
In[3]= geo[ρ_, c_, x0_, y0_, θ0_] := {x'[t] == ρ Cos[θ[t]]2,
  y'[t] == .5 ρ Sin[2 θ[t]], θ'[t] == pθ[t], pθ'[t] == .5 ρ2 Sin[2 θ[t]],
  x[0] == x0, y[0] == y0, θ[0] == θ0, pθ[0] == Sqrt[c - ρ Cos[θ0] ]};
In[4]= GeoCurve[ρ_, c_, x0_, y0_, θ0_] :=
  Module[{s}, s = NDSolve[geo[ρ, c, x0, y0, θ0], {x, y, θ, pθ}, {t, 0, 10}];
  {x[x], y[x], θ[x]} /. s[[1]];

In[5]= ρ = 1;
  c = 2;
  θ0 = -1. π / 2;
  x0 = 0;
  y0 = 0;
  g = GeoCurve[ρ, c, x0, y0, θ0];
  g1 = GeoCurve[ρ, 1.5, x0, y0, θ0];
  g2 = GeoCurve[ρ, 1.1, x0, y0, θ0];
  g3 = GeoCurve[ρ, 1.002, x0, y0, θ0];
  g4 = GeoCurve[ρ, 1, x0, y0, θ0];

In[15]=

In[16]= ParametricPlot[
  {g[[1 ;; 2]], g1[[1 ;; 2]], g2[[1 ;; 2]], g3[[1 ;; 2]], g4[[1 ;; 2]]}, {x, 0, 10}]
Out[16]=

In[17]= g = GeoCurve[ρ, 0.1, x0, y0, θ0];
  g1 = GeoCurve[ρ, 0.5, x0, y0, θ0];
  g2 = GeoCurve[ρ, 0.9, x0, y0, θ0];
  g3 = GeoCurve[ρ, 0.99, x0, y0, θ0];
  g4 = GeoCurve[ρ, 1, x0, y0, θ0];

ParametricPlot[
  {g[[1 ;; 2]], g1[[1 ;; 2]], g2[[1 ;; 2]], g3[[1 ;; 2]], g4[[1 ;; 2]]}, {x, 0, 10}]
Out[22]=


```

Figura 3.1: Geodetiche subriemanniane

Non tutte le geodetiche subriemanniane che abbiamo determinato possono essere interpretate come contorni soggettivi. Infatti si vede sperimentalmente che in un contorno illusorio la differenza fra l'orientazione del punto iniziale e quello finale non può eccedere $\pi/2$. Dobbiamo quindi troncare opportunamente queste curve, per rappresentare il fascio di contorni possibili uscenti da un punto con orientazione fissata.

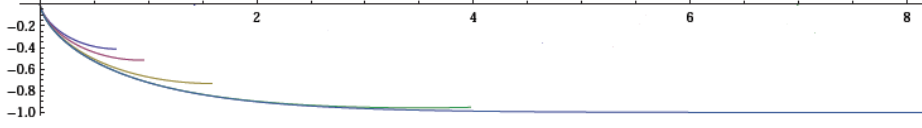


Figura 3.2: Fascio di contorni possibili uscenti da un punto con orientazione fissata

Per analizzare ulteriormente tali curve utilizziamo l'espressione delle geodetiche come minimi del funzionale delle lunghezze. Per il teorema 2.3.4 un secondo metodo per calcolare le geodetiche considerate è quello di ottenerle come punti critici del funzionale delle lunghezze

$$\begin{aligned} \mathcal{L} &= \mathcal{L}(x, y, \theta, \dot{x}, \dot{y}, \dot{\theta}) = \\ &= \int \sqrt{(\dot{x} \cos \theta + \dot{y} \sin \theta)^2 + \frac{(-\dot{x} \sin \theta + \dot{y} \cos \theta)^2}{\varepsilon^2} + (\dot{\theta})^2} \end{aligned}$$

Fatta la variazione prima di questo funzionale è immediato ritrovare le equazioni ottenute nella (3.2). Si osservi inoltre che con questo metodo non si incappa nel problema dovuto ai simboli di Christoffel, che scoppiano per $\varepsilon \rightarrow 0$.

Ricordiamo che nel calcolo delle geodetiche si può sempre assumere $\|\gamma\| \leq 1$. Pertanto in questo caso possiamo assumere che

$$\sqrt{(\dot{x} \cos \theta + \dot{y} \sin \theta)^2 + \frac{(-\dot{x} \sin \theta + \dot{y} \cos \theta)^2}{\varepsilon^2} + (\dot{\theta})^2} = 1.$$

Se imponiamo la condizione $\|\gamma\| \leq 1$, allora deduciamo che le geodetiche della metrica degenerare verificano

$$\frac{(-\dot{x} \sin(\theta) + \dot{y} \cos(\theta))}{\varepsilon^2} \leq 1$$

In altre parole al limite per $\varepsilon \rightarrow 0$ si ha

$$-\dot{x} \sin(\theta) + \dot{y} \cos(\theta) = 0$$

Questo si interpreta dicendo che le curve minimali della metrica limite sono del tipo:

$$\gamma' = \gamma_{X_1} X_1 + \gamma_{X_2} X_2,$$

che come abbiamo visto nel capitolo 1.2.2 è esattamente l'espressione delle curve liftate dalla corteccia.

Per testare la bontà del modello riconosciamo che fornisce risultati confrontabili con il metodo di minimizzazione del più classico funzionale delle elastiche.

3.3 Il funzionale delle elastiche modificato

Come abbiamo ricordato nell'introduzione un funzionale classico che era stato introdotto da Mumford per la modellazione dei contorni soggettivi è il funzionale delle elastiche definito direttamente sul piano immagine, ed è indipendente dalla funzionalità della corteccia. Il funzionale si scrive:

$$\int_{\tilde{\gamma}} (1 + k^2) ds,$$

dove $\tilde{\gamma}$ è una curva sul piano bidimensionale delle variabili (x, y) , e k è la curvatura di $\tilde{\gamma}$. È stato osservato che i minimi di questo funzionale non possono presentare spigoli. Infatti un modo naturale per approssimare uno spigolo, per esempio

$$\{0\} \times [-1, 0] \cup [0, 1] \times \{0\}$$

è approssimarlo con una curva

$$\{0\} \times [-1, -r] \cup \tilde{\gamma}_r \cup [r, 1] \times \{0\}$$

dove $\tilde{\gamma}_r$ è un arco di cerchio di raggio r (si veda figura seguente).

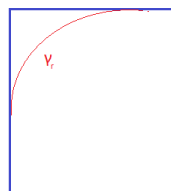


Figura 3.3: Spigolo

Quando il raggio del cerchio tende a 0, si ottiene l'approssimazione dello spigolo. Lungo l'arco di cerchio la curvatura è costantemente uguale a $\frac{1}{r}$ e il funzionale delle elastiche assume il valore

$$\int_0^{\pi/2} r \left(1 + \frac{1}{r^2}\right),$$

Pertanto si vede subito che il funzionale non è limitato al tendere di r a 0. Questo contrasta con l'evidenza sperimentale, perchè i contorni soggettivi possono essere curvi:

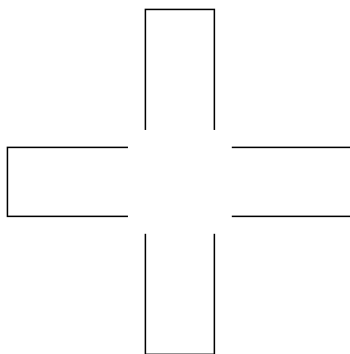


Figura 3.4: Contorni soggettivi

Mumford, Nitzberg e Shiota in [6] hanno quindi proposto di modificare il funzionale delle elastiche con una funzione che crescesse linearmente in k , per esempio:

$$\int_{\tilde{\gamma}} \sqrt{1 + k^2} ds.$$

Sui cerchi del tipo precedente, questo funzionale risulta infatti limitato (si confrontino i limiti per $r \rightarrow \infty$):

$$\int_0^{\pi/2} r \sqrt{1 + \frac{1}{r^2}} = \int_0^{\pi/2} \sqrt{1 + r^2}$$

È proprio con i minimi di questo funzionale che noi confronteremo le geodetiche determinate con la metrica Riemanniana.

3.4 Geodetiche subriemanniane ed elastiche modificate

Possiamo ora confrontare il funzionale delle lunghezze nel caso subriemanniano e il funzionale delle elastiche modificato. I due funzionali operano su spazi diversi: il funzionale delle elastiche su curve 2-dimensionali, mentre il funzionale delle lunghezze sul loro lifting corticale. Abbiamo già ricordato che l'azione delle cellule semplici della corteccia V1 può essere interpretata come un lifting di curve dal piano bidimensionale retinico allo spazio corticale 3D. Come visto se la curva $\tilde{\gamma} = (x(t), y(t))$ sul piano bidimensionale aveva rappresentazione in coordinate d'arco

$$\dot{x} = \|\dot{\tilde{\gamma}}\| \cos \theta \quad \dot{y} = \|\dot{\tilde{\gamma}}\| \sin \theta,$$

la selezione dell'orientazione opera il lifting di questa curva alla curva

$$\gamma = (x, y, \theta)$$

nello spazio $\mathbb{R} \times \mathcal{S}^1$ Allora la funzione θ verifica se

$$\theta = \arctan\left(\frac{\dot{y}}{\dot{x}}\right) \quad \dot{\theta} = \frac{\ddot{y}\dot{x} - \dot{y}\ddot{x}}{\dot{x}^2 + \dot{y}^2} = (\dot{x}^2 + \dot{y}^2)^{1/2} k$$

dove k è appunto la curvatura della curva (x, y) nel piano bidimensionale.

La lunghezza della curva rilevata nella metrica subriemanniana sarà quindi, se parametrizzata su un intervallo $[a, b]$

$$\int_a^b \sqrt{\gamma_{X_1}^2 + \gamma_{X_2}^2} =$$

poichè $\gamma_{X_3} = 0$

$$\int_a^b \sqrt{\gamma_{X_1}^2 + \gamma_{X_3}^2 + \gamma_{X_2}^2} =$$

poichè $\gamma_{X_1}^2 + \gamma_{X_3}^2 = \dot{x}^2 + \dot{y}^2$, mentre $\gamma_{X_2}^2 = k$

$$\int_a^b \sqrt{\dot{x}^2 + \dot{y}^2 + k^2(\dot{x}^2 + \dot{y}^2)}.$$

coincide pertanto con una modifica del funzionale delle elastiche di Mumford:

$$\int_a^b \sqrt{\dot{x}^2 + \dot{y}^2} \sqrt{1 + k^2} = \int_{\tilde{\gamma}} \sqrt{1 + k^2}$$

dove sono state applicate le uguaglianze precedenti. Tale proprietà permette a Citti e Sarti di semplificare il problema della ricerca delle elastiche di Mumford che viene quindi ricondotto a quello delle geodetiche nello spazio $\mathbb{R}^2 \times \mathcal{S}^1$, con una metrica opportuna. I due problemi danno risultati equivalenti, ma mentre il modello di Mumford è un modello puramente fenomenologico, il modello di Citti e Sarti è un modello corticale, che giustifica quindi su base neurale i dati macroscopici forniti dalla psicologia della visione.

Bibliografia

- [1] G.Citti e A.Sarti. A cortical based model of perceptual completion in the roto-translation space. In *Proceeding of the Workshop on Second Order Subelliptic Equations and Applications. Cortona, June, 15-21 2003*.
- [2] G.Citti e A.Sarti. Geodetiche, elastiche, superfici minime nel gruppo di rototraslazioni.
- [3] G. Kanisza. *Grammatica del vedere*. Il mulino, Bologna, 1980.
- [4] J. M. Lee *Riemannian Manifolds: An introduction to Curvature*. Springer-Verlag New York, Inc, Seattle, WA, USA, 1997.
- [5] D. Mumford. Elastica and computer vision. In *Algebraic geometry and its applications (West Lafayette, IN, 1990), pagine 491,506*. Springer, New York, 1994.
- [6] M. Nitzberg D.Mumford e T.Shiota. *Filtering, Segmentation and Depth*. Springer-Verlag New York, Inc, Secaucus, NJ,USA, 1993.
- [7] J. Petitot. *Neurogeometrie de la Vision. Modèles, Mathématiques et physiques des architectures fonctionnelles*. Ecole Polytechnique, 2008.
- [8] G. Sanguinetti. Invariant models of vision between phenomenology, image statistics and neurosciences. Universidad de la Republica, Montevideo, 2011.