

II. Introduzione ai sistemi dinamici iperbolici

(2001/2002)

1. Sistemi dinamici topologici Il teorema della varietà stabile

1.1. Sistemi dinamici topologici.

I sistemi dinamici *iperbolici* sono, molto vagamente, quei sistemi caratterizzati in modo essenziale dalla presenza di contrazione o dilatazione esponenziale delle traiettorie vicine. L'idea, sviluppata pienamente a partire dagli anni '60, ma già chiara nella sostanza a Poincaré, è che all'iperbolicità si accompagnino moti del sistema molto vari e complessi (in un senso precisabile matematicamente in modo chiaro), e dunque un comportamento d'insieme assai interessante. Lo studio di questi sistemi è anche la base per comprendere quello che oggi viene chiamato, con terminologia in realtà un po' vaga (e non sempre con la dovuta serietà) "caos deterministico". Vi sono diverse connessioni con la teoria ergodica esposta nel capitolo precedente, assieme però a una differenza di fondo: qui *non* si fa riferimento a una misura di probabilità conservata, mentre invece giocano un ruolo fondamentale le proprietà metriche o topologiche.¹⁾

Gli oggetti di cui parleremo sono dunque i cosiddetti *sistemi dinamici topologici*, cui già abbiamo accennato nell'ambito della teoria ergodica, e la cui definizione qui riprendiamo aggiungendo qualche precisazione.

Definizione. Sia M uno spazio metrico separabile completo, e Φ un homeomorfismo: $M \rightarrow \Phi(M) \subset M$ [ovvero un gruppo a un parametro $\{\Phi^t, t \in \mathbb{R}\}$ di homeomorfismi: $M \rightarrow \Phi^t(M) \subset M$]. La coppia (M, Φ) si dice *sistema dinamico topologico* discreto [ovvero continuo] di classe C^0 . Se in più M ha la struttura di varietà differenziabile di classe C^r , $r \geq 1$, e Φ [ovvero ciascuno dei Φ^t] è un diffeomorfismo: $M \rightarrow \Phi(M) \subset M$ di uguale classe C^r , allora la coppia (M, Φ) si dice sistema dinamico topologico discreto [ovvero continuo], di classe C^r .

¹⁾ Queste note seguono in larga parte O. Lanford III, *Introduction to the mathematical theory of dynamical systems*, lezioni tenute alla Scuola di Les Houches 1981: "Chaotic behaviour of deterministic systems", volume edito da G. Jooss, R.H.G Helleman, R. Stora (North-Holland, Amsterdam 1983). Dello stesso autore si veda anche *Introduction to hyperbolic sets*, lezioni tenute alla scuola di Erice 1983 *Regular and chaotic motions in dynamical systems*, volume edito da G. Velo e A.S. Wightmann (Plenum Press, New York 1985). Un ottimo manuale di riferimento è A. Katok e B. Hasselblatt, *Introduction to the modern theory of dynamical systems*, Cambridge Univ. Press 1995.

Nel caso continuo si usa anche l'espressione *flusso topologico*. Nel seguito ci occuperemo prevalentemente del caso discreto, cui anzi faremo implicitamente riferimento in assenza di indicazioni.

Molti dei sistemi dinamici considerati nell'ambito della teoria ergodica sono anche sistemi dinamici topologici. Sono di classe C^r , $r \geq 1$, tutti i sistemi classici, quando si escludano singolarità; tra essi i sistemi hamiltoniani, le traslazioni su \mathbb{T}^n , la mappa del gatto, il flusso geodetico. Gli schemi di Bernoulli sono invece un esempio in classe C^0 (con la distanza introdotta nel paragrafo 8 sullo spazio Σ delle sequenze a n simboli).

Si osservi che non si è chiesta l'invertibilità di Φ su tutto M , ma solo su $\Phi(M) \subset M$ (mappa $M \rightarrow M$ iniettiva ma non suriettiva; l'importante è escludere il caso due o molti a uno). Il caso tipico in cui $\Phi(M)$ è contenuto propriamente in M è quello degli attrattori (sistemi dissipativi, in cui il volume di M si contrae) che studieremo tra breve.

1.2. Operatori iperbolici.

Sia E uno spazio vettoriale reale, $\dim E = n$, e L un operatore lineare su E .

Definizione. L'operatore lineare $L : E \rightarrow E$ si dice *iperbolico* se il suo spettro non contiene autovalori di modulo uno.

Proposizione 1.1. *Se L è iperbolico e invertibile, allora esistono due sottospazi E^s, E^u di E tali che:*

- i. $E = E^s \oplus E^u, E^s \cap E^u = \{0\}$;
- ii. E^s e E^u sono invarianti: $LE^u = E^u, LE^s = E^s$;
- iii. le restrizioni $L|_{E^s}$ e $L|_{E^u}$ hanno autovalori rispettivamente dentro e fuori il cerchio unitario;
- iv. per $t \rightarrow \infty$ si ha

$$L^t \xi \rightarrow 0 \quad \text{per } \xi \in E^s; \quad L^{-t} \xi \rightarrow 0 \quad \text{per } \xi \in E^u.$$

E^s e E^u si dicono rispettivamente sottospazio *stabile*, o contraente, e sottospazio *instabile*, o dilatante, di L (u è il ricordo dell'inglese *unstable*). Si osservi che dilatante significa, propriamente, contraente nel passato.

Dimostrazione. È banale se L è diagonalizzabile: in tal caso E^s (ovvero E^u) è la parte reale dello spazio scandito dagli autovettori corrispondenti agli autovalori λ_j in modulo minori di uno (ovvero maggiori di uno); si osservi che se $u \pm iv$ è una coppia di autovettori complesso coniugati di L , corrispondenti ad esempio all'autovalore $|\lambda_j| < 1$, allora per scandire E^s si combinano linearmente u e v (si ricordi che E^s è reale). Se L non è diagonalizzabile, basta far riferimento alla matrice in forma di Jordan associata a L , e aggiungere gli spazi associati a ogni blocco di Jordan con $|\lambda_j| < 1$ (ovvero $|\lambda_j| > 1$). I punti i. – iii. sono ancora banali; per il quarto basta osservare che per ogni blocco di

Jordan \mathcal{L} di dimensione k associato ad esempio a λ_j reale, $0 < \lambda_j < 1$,

$$\mathcal{L} = \begin{pmatrix} \lambda_j & 1 & & & \\ & \lambda_j & 1 & & \\ & & \cdot & \cdot & \\ & & & \cdot & 1 \\ & & & & \lambda_j \end{pmatrix}, \quad (1.1)$$

e per ogni $x \in \mathbb{R}^k$, risulta (come si verifica per ricorrenza)

$$(\mathcal{L}^t x)_l = \sum_{i=0}^{k-l} \binom{t}{i} \lambda_j^{t-i} x_{l+i} \quad l = 1, \dots, k;$$

ciascuno dei coefficienti $\binom{t}{i} \lambda_j^{t-i}$ tende a zero per $t \rightarrow \infty$, e questo basta. ■

Si verifica immediatamente che la quarta proprietà caratterizza E^s e E^u , ovvero che si ha

$$E^s = \{\xi \in E : L^t \xi \rightarrow 0 \text{ per } t \rightarrow \infty\}, \quad E^u = \{\xi \in E : L^{-t} \xi \rightarrow 0 \text{ per } t \rightarrow \infty\};$$

con pari evidenza si ha anche

$$E^s = \{\xi \in E : L^t \xi \text{ è limitato } \forall t > 0\}, \quad E^u = \{\xi \in E : L^{-t} \xi \text{ è limitato } \forall t > 0\}.$$

Infine, si vede che per scambio di Φ con Φ^{-1} i sottospazi E^s e E^u si scambiano tra loro.

Le figure 1.1 – 1.3 illustrano alcuni esempi elementari di operatori iperbolici; in ciascuna figura sono messi in evidenza E^s e E^u , ed è simbolicamente indicato il comportamento di L .

Figura 1.1. $n = 2$, $L = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$, $0 < |a| < 1 < |b|$.

Una comune sella in $E = \mathbb{R}^2$, nella base degli autovettori di L .

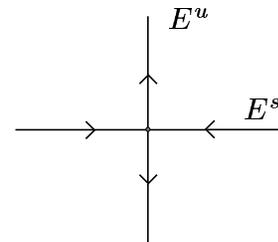


Figura 1.2. $n = 2$, $L = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$, $a^2 + b^2 < 1$.

Si ha $\lambda_{1,2} = a \pm ib$, $|\lambda_{1,2}| < 1$, e dunque E^s coincide con E .

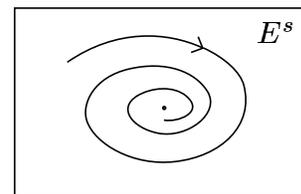
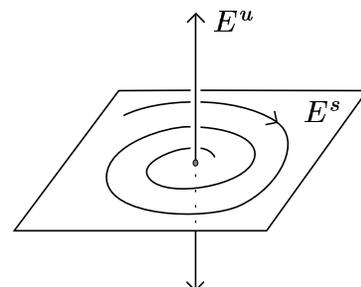


Figura 1.3. $n = 3$, $L = \begin{pmatrix} a & -b & 0 \\ b & a & 0 \\ 0 & 0 & c \end{pmatrix}$, $a^2 + b^2 < 1 < c$.

Qui E^s è un piano, E^u è una retta.



La contrazione o dilatazione nei due sottospazi è uniforme, precisamente

Proposizione 1.2. *Se L è iperbolico, allora esistono costanti C e λ , con $C > 0$ e $0 < \lambda < 1$, tali che per ogni $t > 0$*

$$\|L^t \xi\| \leq C \lambda^t \|\xi\| \quad \text{se } \xi \in E^s, \quad \|L^{-t} \xi\| \leq C \lambda^t \|\xi\| \quad \text{se } \xi \in E^u. \quad (1.2)$$

La scelta della norma in E chiaramente non è rilevante (scelte diverse della norma si compensano con scelte diverse di C).

Dimostrazione. La dimostrazione è immediata se L è diagonalizzabile: siano e_1, \dots, e_n gli autovettori di L , e $\lambda_1, \dots, \lambda_n$ i corrispondenti autovalori, con $|\lambda_j| < 1$ per $j \leq k$ e $|\lambda_j| > 1$ per $j > k$, $0 \leq k \leq n$. Se $\xi = \sum_j \xi_j e_j$, poniamo $\|\xi\| := \sum_j |\xi_j|$. Si verifica all'istante che le (1.2) sono soddisfatte con $C = 1$ e $\lambda = \max(|\lambda_1|, \dots, |\lambda_k|, |\lambda_{k+1}^{-1}|, |\lambda_n^{-1}|)$. La generalizzazione al caso di L non diagonalizzabile è lasciata come esercizio (si usa la forma di Jordan). ■

E' facile vedere che in generale, a meno di una scelta mirata della norma, ci vuole $C > 1$, e dunque non si ha contrazione o dilatazione a ogni singola iterazione ma solo per t sufficientemente grande.

Esempio 1.1. Prendiamo $n = 2$ e

$$L = \begin{pmatrix} a & 1 \\ 0 & b \end{pmatrix}, \quad 0 < a, b < 1,$$

cosicché $L : (\xi_1, \xi_2) \mapsto (a\xi_1 + \xi_2, b\xi_2)$; facciamo riferimento alla norma $\|\xi\| = |\xi_1| + |\xi_2|$. Allora per $\xi = (0, 1)$ si ha $\|L\xi\| = 1 + b > \|\xi\|$, e solo a partire dalle iterazioni successive la norma comincia a calare. Volendo però si può "adattare" la norma a L , in modo tale che la prima delle (1.2) sia soddisfatta con $C = 1$: ad esempio, poniamo $\|\xi\| = \gamma|\xi_1| + |\xi_2|$, con $0 < \gamma < 1 - b$; per ogni scelta di ξ si ottiene subito $\|L\xi\| \leq \gamma a|\xi_1| + (\gamma + b)|\xi_2| \leq \lambda\|\xi\|$, con $\lambda = \max(a, b + \gamma) < 1$.

Si dimostra facilmente che

Proposizione 1.3. *Per ogni operatore iperbolico L esistono sempre una norma adattata e una costante λ , $0 < \lambda < 1$, tali che la costante C nella (1.2) si può prendere uguale a uno, e dunque si ha*

$$\|L\xi\| \leq \lambda\|\xi\| \quad \text{per } \xi \in E^s, \quad \|L^{-1}\xi\| \leq \lambda\|\xi\| \quad \text{per } \xi \in E^u. \quad (1.3)$$

Dimostrazione. E' banale se L è diagonalizzabile, si veda la dimostrazione della Proposizione 1.2. Se L non è diagonalizzabile, si lavora preliminarmente sui singoli blocchi di Jordan, e si dimostra che se \mathcal{L} è un blocco corrispondente all'autovalore λ_j , e \mathcal{E} denota il

corrispondente sottospazio, esistono una norma in \mathcal{E} e una costante $\lambda_{\mathcal{E}}$, $0 < \lambda_{\mathcal{E}} < 1$, tali che per ogni $x \in \mathcal{E}$ risulta

$$\|\mathcal{L}x\| \leq \lambda_{\mathcal{E}}\|x\| \quad \text{se } |\lambda_j| < 1, \quad \|\mathcal{L}^{-1}x\| \leq \lambda_{\mathcal{E}}\|x\| \quad \text{se } |\lambda_j| > 1; \quad (1.4)$$

ammessa la (1.4) la conclusione poi è immediata: basta prendere come norma in tutto E la più grande delle norme nei singoli sottospazi \mathcal{E} , e come costante λ la peggiore (la più vicina a uno) delle costanti $\lambda_{\mathcal{E}}$ relative ai singoli sottospazi. Per verificare la (1.4), supponiamo per semplicità λ_j reale, $0 < \lambda_j < 1$; ponendo

$$\|x\| = \sum_{l=1}^k \gamma^{k-l} |x_l|,$$

con $k = \dim \mathcal{E}$ e $0 < \gamma < 1 - \lambda_j$, segue subito

$$\|\mathcal{L}\| \leq \sum_{l=1}^k (\gamma^{k-l} |\lambda| + \gamma^{k-l+1}) |x_l|,$$

e la (1.4) è soddisfatta con $\lambda_{\mathcal{E}} = |\lambda_j| + \gamma < 1$. ■

L'uso della norma adattata semplifica un po' le dimostrazioni, e ne faremo uso.

1.3. Punti iperbolici; varietà stabile e varietà stabile locale.

Consideriamo ora un sistema dinamico topologico discreto (M, Φ) di classe C^r , $r \geq 1$, e sia z^* un punto fisso di M , $\Phi(z^*) = z^*$. Denotiamo con $E := T_{z^*}M$ lo spazio tangente a M in z^* , e con $L := D\Phi_{z^*}$, $L : E \rightarrow E$, l'applicazione tangente a Φ in z^* .

Definizione. Il punto fisso $z^* \in M$ si dice *iperbolico*, se l'operatore L sopra definito è iperbolico.

Esempio 1.2. Sul cilindro $S^1 \times \mathbb{R}$ consideriamo la mappa

$$\Phi : (x, y) \mapsto (x + y \pmod{1}, y + c \sin 2\pi(x + y));$$

è una discretizzazione del pendolo, nota in letteratura come *standard map*.²⁾ I punti fissi di Φ sono $z_1^* = (0, 0)$ e $z_2^* = (\pi, 0)$, e si ha

$$L(z_1^*) = \begin{pmatrix} 1 & 1 \\ c & 1+c \end{pmatrix}, \quad L(z_2^*) = \begin{pmatrix} 1 & 1 \\ -c & 1-c \end{pmatrix}.$$

Si vede subito che per $c > 0$ il punto z_1^* è iperbolico, mentre non lo è z_2^* .

²⁾ La mappa preserva le aree, ed è usata come esempio tipico di mappa conservativa sul cilindro (o sul toro \mathbb{T}^2 , aggiungendo $\pmod{1}$ anche alla seconda componente).

Esempio 1.3. Data un'equazione differenziale $\dot{z} = f(z)$ in \mathbb{R}^n , denotiamo con $\{\Psi^t, t \in \mathbb{R}\}$ il flusso, e con $\Phi := \Psi^T$ la corrispondente mappa al tempo T fissato. Sia z^* punto di equilibrio, $f(z^*) = 0$, e

$$\mathcal{L} = \left(\frac{\partial f_i}{\partial z_j} \right)_{z^*}, \quad L = D\Phi_{z^*}.$$

Si verifica senza difficoltà che risulta³⁾ $L = e^{T\mathcal{L}}$, e dunque, detti μ_j e λ_j gli autovalori di \mathcal{L} e L rispettivamente, risulta $|\lambda_j| = e^{T\operatorname{Re}\mu_j}$. Pertanto z^* è iperbolico se e solo se \mathcal{L} non ha autovalori con parte reale nulla. Gli autovettori di L , che determinano E^s e E^u , coincidono evidentemente con quelli di \mathcal{L} . Un caso tipico è quello di un sistema hamiltoniano a un grado di libertà della forma $H(p, q) = \frac{1}{2}p^2 + V(q)$. Per ogni massimo q^* di V , con $V''(q^*) < 0$, il punto $z^* = (0, q^*)$ è iperbolico: infatti si trova subito

$$\mathcal{L} := \frac{\partial(\dot{q}, \dot{p})}{\partial(q, p)}(z^*) = \begin{pmatrix} 0 & 1 \\ -V''(q^*) & 0 \end{pmatrix},$$

e posto $\mu = [-V''(q^*)]^{1/2} \in \mathbb{R}$ si ha $\mu_{1,2} = \pm\mu$, $\lambda_{1,2} = e^{\pm\mu T}$. Gli autovettori comuni sono $e_{1,2} = (1, \pm\mu)$. Se invece q^* è un minimo di V , allora z^* non è iperbolico.

Sia allora $z^* \in M$ iperbolico; è spontaneo chiedersi fino a che punto le proprietà della mappa nonlineare Φ , o delle sue iterate Φ^t , assomiglino a quelle della mappa lineare associata, $L = D\Phi_{z^*}$. La risposta è tutt'altro che scontata, perché se da un lato è evidente che in qualche modo l'azione di Φ somiglia a quella di L in prossimità di z^* , dall'altro è anche chiaro che i termini non lineari possono produrre effetti molto rilevanti per t grande. Si dimostra tuttavia, ed è un fatto notevole ricco di conseguenze, che esistono in M due varietà immerse W^s e W^u , dette rispettivamente *varietà stabile* e *varietà instabile* di z^* , con proprietà analoghe a E^s e E^u : precisamente, tangenti a E^s e rispettivamente E^u (e dunque tra loro trasverse) in z^* ; invarianti sotto Φ ; e caratterizzate dal fatto che nel limite $t \rightarrow \infty$

$$\Phi^t(z) \rightarrow z^* \quad \text{per } z \in W^s, \quad \Phi^{-t}(z) \rightarrow z^* \quad \text{per } z \in W^u.$$

Sarà sufficiente nel seguito occuparsi della sola W^s , definendola con precisione e dimostrandone l'esistenza: W^u si potrà poi semplicemente definire come la varietà stabile di z^* , relativa al diffeomorfismo inverso Φ^{-1} .

Cominciamo allora con una costruzione locale della varietà stabile, in un intorno di z^* sufficientemente piccolo. Per questo mettiamoci in una carta (cosicché i punti di M e quelli di $T_{z^*}M$ si confondono), e anzi in una carta adattata a L : precisamente,⁴⁾ se $\dim E^s = k$, introduciamo in \mathcal{U} coordinate $(x, y) = (x_1, \dots, x_k, y_1, \dots, y_{n-k})$ tali che l'origine sia su z^* (ovvero $z^* = (0, 0)$), e inoltre E^s e E^u abbiano equazione rispettivamente $y = 0$ e $x = 0$. In queste coordinate l'operatore L è rappresentato da una matrice a blocchi

$$L = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix};$$

³⁾ Posto $\hat{\mathcal{L}}(z) = \left(\frac{\partial f_i}{\partial z_j}(z) \right)$, $\hat{L}(t, z) = D\Psi_z^t$, con uno scambio di derivate si trova, del tutto in generale,

$$\frac{\partial \hat{L}}{\partial t}(t, z) = \hat{\mathcal{L}}(\Psi^t(z)) \hat{L}(t, z), \quad \frac{\partial \hat{L}}{\partial t}(t, z^*) = \mathcal{L} \hat{L}(t, z^*);$$

la conclusione è immediata.

⁴⁾ L'unico caso interessante è con evidenza quello $0 < k < n$: ma formalmente tutto quello che segue, se letto con intelligenza, resta vero anche per $k = 0$ e $k = n$ (assenza rispettivamente della varietà stabile e della varietà instabile).

corrispondentemente la mappa Φ si scrive

$$\Phi(x, y) = (Ax + f(x, y), By + g(x, y)) ,$$

con f, g nulle nell'origine assieme alle derivate prime. In E^s e E^u utilizzeremo norme adattate a A e B nel senso della proposizione 1.3, cioè tali che risulti

$$\|Ax\| \leq \lambda \|x\| , \quad \|B^{-1}y\| \leq \lambda \|y\| , \quad 0 < \lambda < 1 .$$

Infine, per $\varepsilon > 0$ denotiamo

$$E_\varepsilon^s = \{x \in E^s : \|x\| \leq \varepsilon\} , \quad E_\varepsilon^u = \{y \in E^u : \|y\| \leq \varepsilon\} , \quad \mathcal{U}_\varepsilon = E_\varepsilon^s \times E_\varepsilon^u .$$

Vale allora la seguente fondamentale proposizione, detta *teorema della varietà stabile locale*, o anche *Teorema di Perron–Hadamard*.

Proposizione 1.4 (Teorema della varietà stabile locale). *Se ε è sufficientemente piccolo, allora esiste unica una funzione $w : E_\varepsilon^s \rightarrow E_\varepsilon^u$, regolare quanto Φ , tale che:*

- i. *si ha $w(0) = 0$, e $\frac{\partial w_i}{\partial x_j}(0) = 0$ per ogni i, j ; ovvero: la varietà*

$$W_\varepsilon^s := \{(x, y) \in \mathcal{U}_\varepsilon : y = w(x)\}$$

(il grafico di w) è tangente a E^s nell'origine.

- ii. *Se (x, y) appartiene a W_ε^s , allora anche $\Phi(x, y)$ appartiene a W_ε^s ; ovvero: la varietà W_ε^s è invariante, $\Phi(W_\varepsilon^s) \subset W_\varepsilon^s$.*
- iii. *Si ha $\Phi^t(x, w(x)) \rightarrow (0, 0)$ per $t \rightarrow \infty$.*
- iv. *Se $\Phi^t(x, y) \in \mathcal{U}_\varepsilon$ per ogni $t \geq 0$, allora $y = w(x)$; ovvero: per tutti i punti $(x, y) \in \mathcal{U}_\varepsilon \setminus W_\varepsilon^s$ esiste $t > 0$ tale che $\Phi^t(x, y) \notin \mathcal{U}_\varepsilon$.*

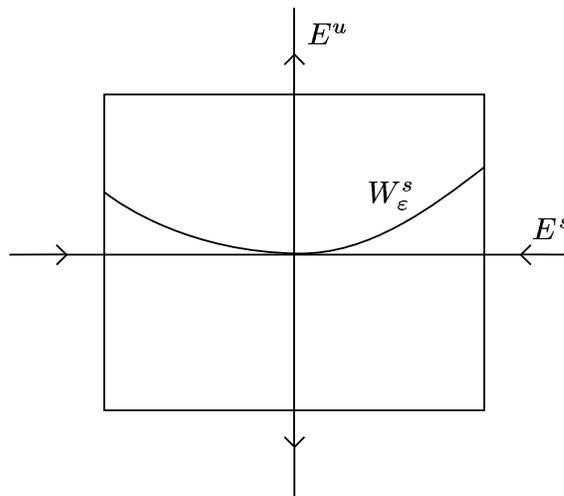


Figura 1.4. Il quadrato \mathcal{U}_ε e la varietà stabile locale W_ε^s .

Si veda la figura 1.4. La varietà W_ε^s si dice *varietà stabile locale* di z^* . Si osservi che i punti ii. e iv. implicano l'unicità di w .

La dimostrazione della proposizione è demandata al prossimo paragrafo. Vediamo subito invece come a partire da W_ε^s si arrivi immediatamente a costruire W^s .

Definizione. La varietà W^s definita da

$$W^s = \bigcup_{t \geq 0} \Phi^{-t}(W_\varepsilon^s) \tag{1.5}$$

è detta *varietà stabile* di z^* .

Usando la proposizione 1.4 si verifica immediatamente che W^s così definita è effettivamente una varietà immersa in M , e inoltre che W^s è connessa, non ha punti doppi, e localmente è tanto regolare quanto W_ε^s (quanto Φ). Globalmente però può essere molto complicata, anche densa in M . Una caratterizzazione alternativa di W^s , esattamente equivalente alla (1.5), è offerta dalla seguente proposizione:

Proposizione 1.5. *Risulta*

$$W^s = \{z \in M : \Phi^t(z) \rightarrow z^* \text{ per } t \rightarrow \infty\} . \tag{1.6}$$

La (1.6) è frequentemente presentata come definizione stessa di W^s .

Dimostrazione. Se $z \in W^s$, allora in base alla definizione (1.5) esiste $\tau \geq 0$ tale che $\Phi^\tau(z) \in W_\varepsilon^s$, e allora $\Phi^t(z) \rightarrow z^*$ per $t \rightarrow \infty$. Viceversa, se $\Phi^t(z) \rightarrow z^*$, allora per τ opportuno e ogni $t \geq 0$ si ha $\Phi^t(\Phi^\tau(z)) \in W_\varepsilon^s$; ma allora per il punto iv. della Proposizione 1.4 si ha $\Phi^\tau(z) \in W_\varepsilon^s$, e corrispondentemente $z \in \Phi^{-\tau}(W_\varepsilon^s) \subset W^s$. ■

Infine definiamo formalmente W^u :

Definizione. Si dice varietà instabile W^u di z^* la varietà stabile di z^* relativa al diffeomorfismo inverso Φ^{-1} .

Esempio 1.4. Consideriamo l'automorfismo algebrico di \mathbb{T}^2 definito da

$$\Phi(x, y) = (x + y, x + 2y) \pmod{1}$$

(Gatto di Arnol'd). L'origine è un punto fisso, e $L = D\Phi_{(0,0)}$ coincide con la matrice $A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$. I due autovalori sono $\lambda^s = \lambda := (3 - \sqrt{5})/2 < 1$ e $\lambda^u = \lambda^{-1} > 1$, con autovettori $e^s = (1, \lambda - 1)$ e $e^u = (1, \lambda^{-1} - 1)$; gli spazi lineari E^s e E^u sono le rette che contengono e^s e e^u , mentre le varietà W^s, W^u sono le stesse rette riportate al toro con l'operazione di modulo. Si osservi che W^s e W^u localmente sono ben fatte, ma globalmente sono molto complicate (come si è visto nel primo capitolo, ciascuna curva è densa in M). La situazione è schematicamente ritratta in figura 1.5.

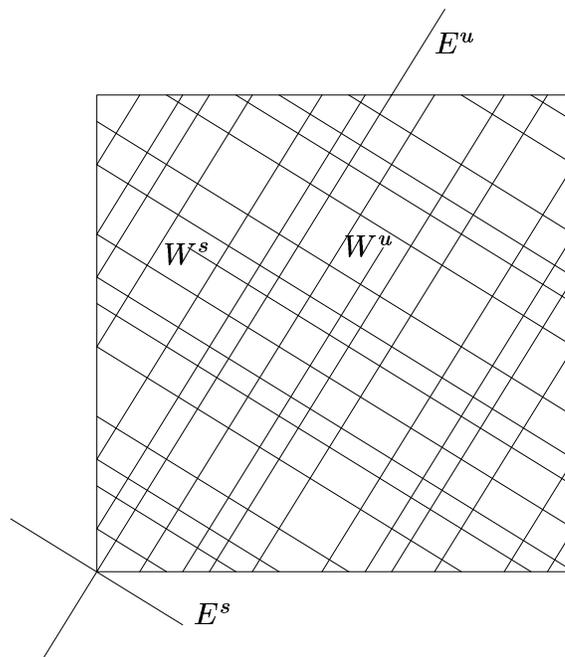


Figura 1.5. Gli spazi E^s , E^u e le corrispondenti varietà W^s , W^u avvolte su \mathbb{T}^2 , per il Gatto di Arnol'd.

Esempio 1.5. Si ponga $\Phi =$ mappa al tempo uno del pendolo. Esistono due punti fissi iperbolici $z_{\pm}^* = (\pm\pi, 0)$ (si riveda l'esempio 1.3), che per fissare le idee pensiamo qui distinti. Le loro varietà stabili e instabili altro non sono che le separatrici del pendolo, e una stessa curva è contemporaneamente varietà stabile di un punto e varietà instabile dell'altro. Si veda la figura 1.6a. Una situazione analoga si presenta per la mappa al tempo uno di un qualunque sistema hamiltoniano a un grado di libertà che abbia un punto di sella. Ad esempio, la figura 1.6b si riferisce all'hamiltoniana $H(p, q) = \frac{1}{2}p^2 + V(q)$, con $V(q) = -\frac{1}{2}q^2 + q^3$; anche qui la separatrice è contemporaneamente varietà stabile e varietà instabile di z^* (l'origine).

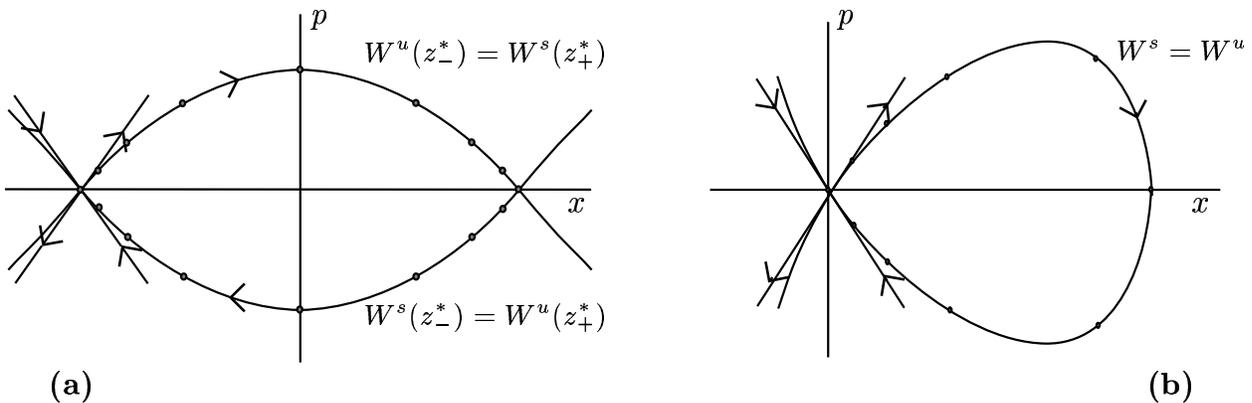


Figura 1.6. W^s e W^u per la mappa al tempo uno del pendolo (a) e del sistema hamiltoniano sopra descritto (b).

1.4. Dimostrazione del teorema della varietà stabile locale.

Ci limitiamo qui al caso $n = 2$, $\dim E^s = \dim E^u = 1$, e dunque a

$$\Phi(x, y) = (ax + f(x, y), by + g(x, y)) , \quad 0 < |a| < 1 < |b| ,$$

ma la dimostrazione si potrebbe estendere in dimensione qualsiasi senza variazioni di sostanza. Ci limitiamo inoltre a mostrare che w è di classe C^1 , supponendo che Φ sia di classe C^1 . Per semplificare la notazione supponiamo $a, b > 0$.

Per eliminare il fastidio di lavorare nel dominio \mathcal{U}_ε piccolo, conviene eseguire preliminarmente un riscaldamento che porti il dominio a dimensione uno. Per questo sfruttiamo la continuità delle derivate f'_x, f'_y, g'_x, g'_y per dire che per η arbitrario, se ε è sufficientemente piccolo e $(x, y) \in \mathcal{U}_\varepsilon$, si ha

$$|f(x, y)|, |g(x, y)| < \varepsilon\eta , \quad |f'_x(x, y)|, \dots, |g'_y(x, y)| < \eta$$

(la seconda è banale, la prima è ovvia conseguenza della seconda). Poniamo allora

$$x = \varepsilon\tilde{x} , \quad y = \varepsilon\tilde{y} ,$$

cosicch  $(\tilde{x}, \tilde{y}) \in \mathcal{U}_1$, e poi

$$\tilde{\Phi}(\tilde{x}, \tilde{y}) = \varepsilon^{-1}\Phi(\varepsilon\tilde{x}, \varepsilon\tilde{y}) , \quad \tilde{f}(\tilde{x}, \tilde{y}) = (\varepsilon\eta)^{-1}f(\varepsilon\tilde{x}, \varepsilon\tilde{y}) , \quad \tilde{g}(\tilde{x}, \tilde{y}) = (\varepsilon\eta)^{-1}g(\varepsilon\tilde{x}, \varepsilon\tilde{y}) .$$

Si verifica immediatamente che la nuova mappa $\tilde{\Phi} : \mathcal{U}_1 \rightarrow \mathcal{U}_1$ ha la forma

$$\tilde{\Phi}(\tilde{x}, \tilde{y}) = (a\tilde{x} + \eta\tilde{f}(\tilde{x}, \tilde{y}), b\tilde{y} + \eta\tilde{g}(\tilde{x}, \tilde{y})) , \tag{1.7}$$

e i termini non lineari \tilde{f} e \tilde{g} sono limitati assieme alle loro derivate, pi  precisamente si ha

$$\|\tilde{f}\|, \|\tilde{g}\|, \|\tilde{f}'_x\|, \dots, \|\tilde{g}'_y\| < 1 , \tag{1.8}$$

con $\|\cdot\| = \text{sup-norma}$. Resta ovviamente vero che \tilde{f} e \tilde{g} si annullano nell'origine assieme alle derivate prime. Nel seguito sopprimeremo la tilde.

La Proposizione 1.4 (nel contesto limitato nel quale ci siamo proposti di dimostrarla) si traduce ora nel modo seguente: *per η sufficientemente piccolo esiste unica una funzione $w : [-1, 1] \rightarrow [-1, 1]$, di classe C^1 , tale che:*

- i. $w(0) = 0, w'(0) = 0$.
- ii. $W_1^s := \{(x, y) \in \mathcal{U}_1 : y = w(x)\}$   invariante: $\Phi(x, w(x)) \in W_1^s$ per ogni $x \in [-1, 1]$.
- iii. Per ogni $x \in [-1, 1]$ risulta $\Phi^t(x, w(x)) \rightarrow (0, 0)$ per $t \rightarrow \infty$.
- iv. Se $(x, y) \in \mathcal{U}_1$ e $y \neq w(x)$, allora esiste $t > 0$ tale che $\Phi^t(x, y) \notin \mathcal{U}_1$.

La dimostrazione si articola in quattro passi.

Primo passo: L'esistenza di w si riporta alla soluzione di una opportuna equazione funzionale. Per questo osserviamo che grazie all'espressione (1.7) di Φ , la propriet  ii. di w si pu  cos  riformulare:

$$w(x) = b^{-1} [w(ax + \eta f(x, w(x))) - \eta g(x, w(x))] . \tag{1.9}$$

La funzione w si presenta allora come soluzione dell'equazione funzionale

$$v = \mathcal{F} v , \quad (1.10)$$

nella quale \mathcal{F} denota l'operatore che agisce sulle funzioni $v : [-1, 1] \rightarrow [-1, 1]$ nel modo seguente:

$$(\mathcal{F} v)(x) = b^{-1} [v(ax + \eta f(x, v(x))) - \eta g(x, v(x))] . \quad (1.11)$$

Si osservi che per η piccolo la quantità $ax + \eta f(x, v(x))$, ad argomento di v , non supera in modulo $a + \eta < 1$ (si ricordi $\|f\| < 1$), dunque \mathcal{F} è ben definito. Si osservi anche che per $\eta = 0$ la (1.10) è risolta da $w(x) = 0$: in assenza di nonlinearità la varietà stabile coincide con il sottospazio contraente (che qui è l'asse x , di equazione $y = 0$).

Secondo passo: Mostriamo che se η è piccolo l'insieme di funzioni

$$\mathcal{V} := \{v \in C^1([-1, 1]) : \|v\| \leq 1, \|v'\| \leq 1\}$$

è invariante sotto \mathcal{F} , ovvero che $v \in \mathcal{V} \Rightarrow \|\mathcal{F}v\| \leq 1$ e $\|(\mathcal{F}v)'\| \leq 1$, e inoltre che esiste $\mu < 1$ tale che

$$\|\mathcal{F}v_2 - \mathcal{F}v_1\| \leq \mu \|v_2 - v_1\| \quad (1.12)$$

per ogni scelta di $v_2, v_1 \in \mathcal{V}$ (\mathcal{F} è una contrazione in norma C^0).

Infatti:

a) se $v \in \mathcal{V}$ e $|x| \leq 1$, allora dalle (1.8) segue subito, per η piccolo,

$$b^{-1} |v(ax + \eta f(x, v(x))) - \eta g(x, v(x))| \leq b^{-1}(1 + \eta) \leq 1 ,$$

ovvero $\|\mathcal{F}v\| \leq 1$.

b) In modo analogo si procede per la derivata, la cui espressione è [notazione breve: $f = f(x, v(x))$, $g = g(x, v(x))$]

$$(\mathcal{F}v)'(x) = b^{-1} \{v'(ax + \eta f) [a + \eta(f'_x + v'(x)f'_y)] - \eta(g'_x + v'(x)g'_y)\} . \quad (1.13)$$

Procedendo con pazienza come sopra, si trova $|v'(ax + \eta f)| \leq 1$, $|f'_x + v'(x)f'_y| \leq 2$, $|g'_x + v'(x)g'_y| \leq 2$, e infine

$$\|(\mathcal{F}v)'\| \leq b^{-1}(a + 4\eta) \leq 1 .$$

c) Si ha [notazione breve: $f_i := f(x, v_i(x))$, $g_i := g(x, v_i(x))$, $i = 1, 2$]

$$\mathcal{F}v_2 - \mathcal{F}v_1 = b^{-1} [v_2(ax + \eta f_2) - v_1(ax + \eta f_1) - \eta g_2 + \eta g_1] ;$$

aggiungendo e togliendo entro la parentesi quadra $v_2(ax + \eta f_1)$, e usando uno come costante di Lipschitz di f e g in ciascun argomento, segue subito

$$\begin{aligned} \|\mathcal{F}v_2 - \mathcal{F}v_1\| &\leq b^{-1} [\eta \|f_2 - f_1\| + \|v_2 - v_1\| + \eta \|g_2 - g_1\|] \\ &\leq b^{-1}(1 + 2\eta) \|v_2 - v_1\| , \end{aligned}$$

e dunque la (1.12) è soddisfatta con $\mu = b^{-1}(1 + 2\eta) < 1$.

Terzo passo: Mostriamo⁵⁾ che la successione v_0, v_1, \dots definita da⁶⁾

$$v_0(x) = 0, \quad v_{k+1} = \mathcal{F}v_k \quad (1.14)$$

converge a una funzione $w \in \mathcal{V}$, che risolve la (1.10) e soddisfa $w(0) = w'(0) = 0$.

Per questo osserviamo innanzitutto che per il punto precedente risulta $\|v_k\| \leq 1$ per ogni k , e inoltre

$$\|v_{k+1} - v_k\| \leq \mu^k \|v_1 - v_0\|;$$

questo basta per la convergenza puntuale della successione a una funzione limite w , con $\|w\| \leq 1$. Dalla forma di \mathcal{F} si vede poi che se $v_k(0) = 0$ allora $v_{k+1}(0) = 0$, pertanto tutte le v_k si annullano nell'origine, e di conseguenza $w(0) = 0$. Proviamo ora che il limite w della successione soddisfa l'equazione funzionale (1.10); per questo basta mostrare che risulta

$$\mathcal{F}w = \lim_{k \rightarrow \infty} \mathcal{F}v_k$$

(a destra si ha infatti $\mathcal{F}v_k = v_{k+1} \rightarrow w$). A questo scopo scriviamo

$$\begin{aligned} (\mathcal{F}w)(x) &= b^{-1} [w(ax + \eta f(x, w(x))) - \eta g(x, w(x))] \\ (\mathcal{F}v_k)(x) &= b^{-1} [v_k(ax + \eta f(x, v_k(x))) - \eta g(x, v_k(x))] , \end{aligned}$$

e osserviamo che

$$\begin{aligned} &|w(ax + \eta f(x, w(x))) - v_k(ax + \eta f(x, v_k(x)))| \\ &\leq |w(ax + \eta f(x, w(x))) - v_k(ax + \eta f(x, w(x)))| \\ &\quad + |v_k(ax + \eta f(x, w(x))) - v_k(ax + \eta f(x, v_k(x)))| \\ &\leq \|w - v_k\| + \eta |w(x) - v_k(x)| \leq (1 + \eta) \|w - v_k\| \\ &\rightarrow 0 \end{aligned}$$

per $k \rightarrow \infty$; nello stesso limite si ha $g(x, v_k(x)) \rightarrow g(x, w(x))$, e la conclusione è immediata.

Veniamo ora alla differenziabilità di w . Questa segue dal teorema di Ascoli–Arzelà, secondo il quale per ogni successione $u_0, u_1, \dots, u_k, \dots$ equicontinua (ovvero continua uniformemente in k oltre che in x) e equilimitata (ciascuna u_k è limitata da una costante indipendente da k) esiste una sottosuccessione $u_{k_i}, i \in \mathbb{N}$, che converge uniformemente. Il teorema si usa così:

- a) si pone $u_k = v'_k$, e si mostra che la successione soddisfa le ipotesi del teorema;
- b) dalla conseguente convergenza uniforme della sottosuccessione $v'_{k_i}, i \in \mathbb{N}$, si deduce che la funzione w , cui già sappiamo convergere (in particolare) la sottosuccessione

⁵⁾ La proprietà (1.12) dimostrata al punto precedente non consente purtroppo una conclusione rapida basata sul lemma della contrazione. Il punto è che l'insieme \mathcal{V} in cui stiamo lavorando, e nel quale cerchiamo la soluzione della (1.10), è di classe C^1 , mentre la contrazione sopra dimostrata è solo in norma C^0 .

⁶⁾ Si osservi che la prima approssimazione della varietà stabile è l'asse x .

v_{k_i} , è derivabile, e anzi la sua derivata w' coincide con il limite della sottosuccessione delle derivate:

$$w' = \lim_{i \rightarrow \infty} v'_{k_i} .$$

Poiché per ogni k , come è evidente per ricorrenza dalla (1.13), si ha $v'_k(0) = 0$, certamente anche al limite risulta $w'(0) = 0$.

Per verificare che la successione delle v'_k è equicontinua (che sia equilimitata è ovvio) consideriamo innanzitutto f'_x, \dots, g'_y ; esse sono continue in $[-1, 1]$, dunque per ogni $\rho > 0$ esiste $\delta = \delta(\rho) > 0$ tale che, se uno o entrambi gli argomenti di f'_x, \dots, g'_y variano al più di δ , allora ciascuna funzione varia al più di ρ . Mostriamo che con la stessa δ indipendente da k , e per ogni k , risulta anche

$$|v'_k(x + \xi) - v'_k(x)| < \rho \quad \text{se} \quad |\xi| < \delta .$$

Per $k = 0$ la proprietà è banalmente verificata; supposta vera per k , dimostriamola per $k + 1$. Usiamo per questo l'espressione (1.13) di $(\mathcal{F}v)'$, per $v = v_k$. Si vede facilmente che per incremento ξ alla variabile x , $|\xi| < \delta$,

la quantità	varia meno di	
$v_k(x)$	δ	(v ha cost. Lipshitz 1)
$ax + \eta f(x, v_k(x))$	$(a + \eta)\delta < \delta$	(f ha cost. Lipshitz 1)
$v'_k(ax + \eta f(x, v_k(x)))$	ρ	(ipotesi induttiva)
$f'_x(x, v_k(x)), \dots, g'_y(x, v_k(x))$	ρ	(ipotesi su $\delta(\rho)$)

Se ne deduce senza difficoltà⁷⁾

$$|v'_{k+1}(x + \xi) - v'_{k+1}(x)| \leq b^{-1}[\rho(a + 2\eta) + 3\eta\rho + 3\eta\rho] < \rho .$$

Quarto passo: Fino ad ora abbiamo dimostrato l'esistenza (non ancora l'unicità) di una funzione w , di classe C^1 , che soddisfa i primi due punti della proposizione (o della sua riscrittura dopo il riscaldamento). Concludiamo allora la dimostrazione provando il punto iv., che a sua volta come già si è osservato implica l'unicità di w , e per ultimo il punto iii.

A questo scopo, diciamo che un segmento $[z_1, z_2]$, con estremi $z_1 = (x_1, y_1)$ e $z_2 = (x_2, y_2)$ in \mathcal{U}_1 , è "di tipo verticale", o semplicemente verticale, se $|y_2 - y_1| > |x_2 - x_1|$. Si verifica facilmente che Φ manda segmenti verticali in segmenti verticali: infatti, posto $(x'_i, y'_i) = \Phi(x_i, y_i)$, $i = 1, 2$, e introdotte le notazioni agili

$$\Delta_x = |x_2 - x_1| , \quad \Delta_y = |y_2 - y_1| , \quad \Delta'_x = |x'_2 - x'_1| , \quad \Delta'_y = |y'_2 - y'_1| ,$$

si trova subito

$$\begin{aligned} \Delta'_x &\leq a\Delta_x + \eta|f(x_2, y_2) - f(x_1, y_1)| \leq a\Delta_x + \eta(\Delta_x + \Delta_y) \\ \Delta'_y &\geq b\Delta_y - \eta|g(x_2, y_2) - g(x_1, y_1)| \geq b\Delta_y - \eta(\Delta_x + \Delta_y) , \end{aligned}$$

⁷⁾ Si usa, per i prodotti, la proprietà banale, valida per funzioni α e β qualsiasi,

$$|\alpha(x + \xi)\beta(x + \xi) - \alpha(x)\beta(x)| \leq \|\alpha\| |\beta(x + \xi) - \beta(x)| + \|\beta\| |\alpha(x + \xi) - \alpha(x)| .$$

e dunque

$$\Delta'_y - \Delta'_x \geq (b - 2\eta)\Delta_y - (a + 2\eta)\Delta_x > \Delta_y - \Delta_x .$$

Anzi: si vede anche che se un segmento è verticale, allora la sua componente y cresce, precisamente si trova

$$\Delta'_y \geq \sigma \Delta_y , \quad \sigma := (b - 2\eta) > 1 .$$

Il quarto punto è ora immediato: se infatti $y \neq w(x)$, allora il segmento di estremi $z_1 = (x, y)$ e $z_2 = (x, w(x))$, ovviamente verticale, resta con un estremo su W_ε^s (che già sappiamo essere invariante), ma iterando Φ cresce illimitatamente, dunque prima o poi l'altro estremo esce da \mathcal{U}_1 .

Infine, mostriamo che per ogni $z \in W_\varepsilon^s$ l'iterato $\Phi^t(z)$ converge all'origine. Per questo osserviamo che il segmento che congiunge l'origine a $\Phi^t(z)$ non può essere verticale: altrimenti, per quanto detto sopra, la sua lunghezza crescerebbe illimitatamente, e W_ε^s non sarebbe invariante. Posto dunque $\Phi^t(z) = (x_t, y_t)$, si ha sempre $y_t \leq x_t$. Ma allora

$$|x_{t+1}| = |ax_t + \eta f(x_t, y_t)| \leq (a + 2\eta)|x_t| ,$$

e la conclusione è immediata. ■

1.5. Dipendenza regolare da un parametro.

Sia z^* un punto fisso iperbolico di $\Phi : M \rightarrow M$, ovvero $\Phi(z^*) = z^*$. Mettendosi in coordinate in un intorno \mathcal{U} sufficientemente piccolo di z^* , ha senso introdurre la mappa Ψ definita da

$$\Psi(z) = \Phi(z) - z ,$$

che manda z^* in zero, e \mathcal{U} in un intorno \mathcal{V} dello zero. Ora, gli autovalori della matrice jacobiana $D\Psi_{z^*}$ sono tutti diversi da zero, dunque (se \mathcal{U} non è troppo grande) Ψ è un diffeomorfismo, e l'unico punto che va nello zero — ovvero l'unico punto fisso di Φ in \mathcal{U} — è z^* stesso. Se ne deduce una elementare proposizione:⁸⁾

Proposizione 1.6. *I punti fissi iperbolici sono isolati.*

Consideriamo ora un diffeomorfismo Φ_α , dipendente da un parametro reale α preso in un intorno dello zero.

Proposizione 1.7. *Se $\Phi_\alpha : M \rightarrow M$ è un diffeomorfismo di classe C^r anche nel parametro α , e Φ_0 ammette un punto fisso iperbolico z^* , allora per α sufficientemente piccolo:*

- i. esiste un punto fisso iperbolico $z^*(\alpha)$ di Φ_α , che dipende da α in classe C^r ;*
- ii. la varietà stabile $W^s(\alpha)$ di $z^*(\alpha)$ dipende essa stessa da α in classe C^r .*

Dire che $W^s(\alpha)$ dipende da α in classe C^r significa che se si taglia $W^s(\alpha)$ con un “piano” fisso Π , $\dim \Pi = \dim W_\alpha^u$, trasverso a $W^s(0)$ (una retta trasversa, nel caso semplice $n = 2$, $\dim W^s(\alpha) = \dim W^u(\alpha) = 1$), il punto di intersezione $z(\alpha)$ è funzione regolare di α .

⁸⁾ Si osservi che la proposizione è anche una conseguenza immediata del punto iv. della Proposizione 1.4: ma dedurlo in questo modo sarebbe stata cosa sproporzionata a un risultato così elementare.

Dimostrazione. Il primo punto è facile: si definisce localmente Ψ_α ponendo $\Psi_\alpha(z) = \Phi_\alpha(z) - z$, per z in un intorno \mathcal{U} di z^* . Per α sufficientemente piccolo l'immagine di \mathcal{U} è un intorno \mathcal{V} dello zero prossimo a $\Psi_0(\mathcal{U})$, e $\Psi_\alpha : \mathcal{U} \rightarrow \mathcal{V}$ è ancora un diffeomorfismo, di classe C^r anche in α (si usa il fatto ovvio che il determinante della matrice jacobiana $D\Psi_{\alpha,z}$ dipende con continuità dagli elementi di matrice, dunque resta diverso da zero per α piccolo e \mathcal{U} piccolo). Allora l'inversa Ψ_α^{-1} esiste ed è di classe C^r anche in α , e dunque $z_\alpha^* = \Psi_\alpha^{-1}(0)$ è il punto fisso cercato. L'iperbolicità di z_α^* segue dal fatto che gli autovalori della matrice jacobiana sono funzioni continue degli elementi di matrice (benché non differenziabili, in caso di autovalori multipli).

Per la seconda parte ci limitiamo, come abbiamo fatto per il teorema della varietà stabile locale, al caso $n = 2$, $\dim E^s = \dim E^u = 1$, e inoltre ci accontentiamo di lavorare in classe C^1 . Osserviamo anche che è sufficiente dimostrare la regolarità di $W^s(\alpha)$ in un intorno del punto fisso, ovvero verificare la regolarità in α della varietà stabile locale W_ε^s : infatti, qualunque porzione finita di W^s si ottiene da W_ε^s con un numero finito di applicazioni di Φ_α^{-1} , e pertanto la regolarità di W_ε^s si trasferisce (ovviamente senza uniformità) a tutta W^s . Infine, osserviamo che per studiare W_ε^s ci si può mettere in coordinate adattate per ogni α , senza perdere di regolarità in α : infatti, sia z_α^* , su cui si trasla l'origine, sia $E^s(\alpha)$ e $E^u(\alpha)$, che si prendono come nuovi assi coordinati, dipendono con regolarità⁹⁾ da α .

La dimostrazione a questo punto si riduce alla dimostrazione che la funzione w_α , il cui grafico fornisce la varietà stabile locale, è regolare in classe C^1 anche in α , e a sua volta ciò si ottiene con modesti ritocchi alla dimostrazione già vista del teorema della varietà stabile locale. Precisamente, al terzo passo, dopo aver definito per ogni α la successione analoga alla (1.14):

$$v_0^\alpha(x) = 0, \quad v_{k+1}^\alpha = \mathcal{F}v_k^\alpha,$$

si osserva che ciascuna funzione eredita dalla precedente la differenziabilità anche in α ; usando allora, con identica procedura, il teorema di Ascoli–Arzelà, si conclude senza difficoltà che la funzione limite w^α è differenziabile anche in α . ■

- **Osservazione.** Come si è osservato nel corso della dimostrazione, benché ogni tratto finito $\Phi_\alpha^{-k}(W_\varepsilon^s(\alpha))$ di $W^s(\alpha)$ sia regolare in α , in particolare continuo, in generale *non c'è da aspettarsi uniformità in k* : per α comunque piccolo ma fissato, tratti di $W^s(\alpha)$ sufficientemente lontani dall'origine (per la precisione: raggiungibili dalla varietà stabile locale con un numero elevato di iterazioni di Φ_α^{-1}) possono scostarsi di molto da $W^s = W^s(0)$. Si ritorni su questo punto dopo aver visto, nel prossimo paragrafo, come si comportano W^s e W^u in presenza di “punti omoclini”.

⁹⁾ Nel caso che stiamo trattando abbiamo due autovalori distinti, ed è immediato verificare che allora sia gli autovalori sia gli autovettori di $L(\alpha) = D\Phi_{z_\alpha^*}$ sono regolari in α . In dimensione più alta vi possono essere autovalori multipli, ciononostante, si potrebbe vedere, E^s e E^u restano regolari in α .

2. Fenomeni omoclini

2.1. Punti omoclini.

Riprendendo l'esempio 1.5, consideriamo un sistema hamiltoniano a un grado di libertà della forma elementare

$$H(p, q) = \frac{1}{2}p^2 + V(q) ,$$

con $p \in \mathbb{R}$ e $q \in \mathbb{R}$ (oppure $q \in S^1$, come per il pendolo); lo spazio delle fasi è $M = \mathbb{R}^2$ (oppure $M = \mathbb{R} \times S^1$). Supponiamo che V abbia un massimo quadratico in un punto q^* , e corrispondentemente $z^* = (0, q^*) \in M$ sia punto di equilibrio instabile (punto di sella) del sistema. Sia $E^* := H(0, q^*) = V(q^*)$; la corrispondente curva di livello di H (o "superficie" di energia costante) di equazione $H(p, q) = E^*$ contiene, oltre al punto di equilibrio z^* , le sue separatrici: precisamente le curve costituite dai punti che per $t \rightarrow \infty$, oppure $t \rightarrow -\infty$, tendono a z^* . Fissato $T > 0$ qualsiasi, sia Φ la mappa al tempo T del sistema. Allora, come già si è osservato, z^* è punto fisso iperbolico di (M, Φ) , e le sue varietà asintotiche W^s e W^u coincidono con le separatrici. In casi tipici, in particolare nei due sistemi considerati nell'esempio 1.5, la separatrice forma un laccio, ovvero una curva chiusa che va da z^* a z^* stesso ed è contemporaneamente varietà stabile e instabile di z^* . Questa situazione semplice è dovuta alla conservazione dell'energia, che non offre alternative: nella sostanza, W^s e W^u devono essere curve di livello dell'hamiltoniana, e se questa è una funzione regolare non possono essere troppo complicate.

Come per primo comprese Poincaré, perturbazioni anche molto semplici di Φ possono tuttavia condurre a una situazione diversa in modo drastico: W^s e W^u , anziché raccordarsi, possono avere un'intersezione trasversa in un punto z_0 , detto¹⁰⁾ *punto omoclineo*, o *intersezione omoclinea (trasversa)*, si veda la figura 2.1a. Rinviando al prossimo paragrafo la dimostrazione che ciò effettivamente accade per perturbazioni anche semplicissime di Φ (corrispondenti all'introduzione nelle equazioni di Hamilton di un termine forzante dipendente esplicitamente dal tempo con periodo T : in questo modo si sfugge alla conservazione dell'energia), vediamo qui quali conseguenze comporti la presenza dell'intersezione omoclinea z_0 in una generica mappa Φ .

Per $k \rightarrow \infty$ il punto $\Phi^k(z_0)$ tende a z^* lungo W^s , e allo stesso modo per $k \rightarrow -\infty$ tende a z^* lungo W^u . Gli iterati di z_0 , denotati

$$z_k := \Phi^k(z_0) , \quad k \in \mathbb{Z} ,$$

sono indicati in figura 2.1a; essi costituiscono, si dice, un'*orbita omoclinea*. Ora, comunque si prenda $l \in \mathbb{Z}$, risulta evidentemente $\Phi^{\pm k}(z_l) \rightarrow z^*$ per $k \rightarrow \infty$: ma ciò vuol dire che ciascuno degli z_l appartiene contemporaneamente a W^s e W^u , e (se $D\Phi$ non è singolare) l'intersezione è ancora trasversa: *ciascun punto dell'orbita omoclinea è a sua volta punto omoclineo*. Consideriamo, per essere definiti, il caso in cui Φ preserva l'orientazione (come certamente avviene se Φ è la mappa a un tempo T fissato di un flusso, o una sua deformazione); allora W^u , prolungata oltre z_0 , taglia W^s in z_1, z_2, \dots "da dentro a fuori", e

¹⁰⁾ Il nome sta a indicare che per applicazione di Φ^k il punto z_0 piega (gr. $\kappa\lambda\acute{\iota}\nu\omega$), per $k \rightarrow \pm\infty$, verso il medesimo punto fisso z^* . L'eventuale intersezione della varietà instabile di un punto iperbolico z_1^* con la varietà stabile di un diverso punto iperbolico z_2^* è detta invece *punto eteroclineo*.

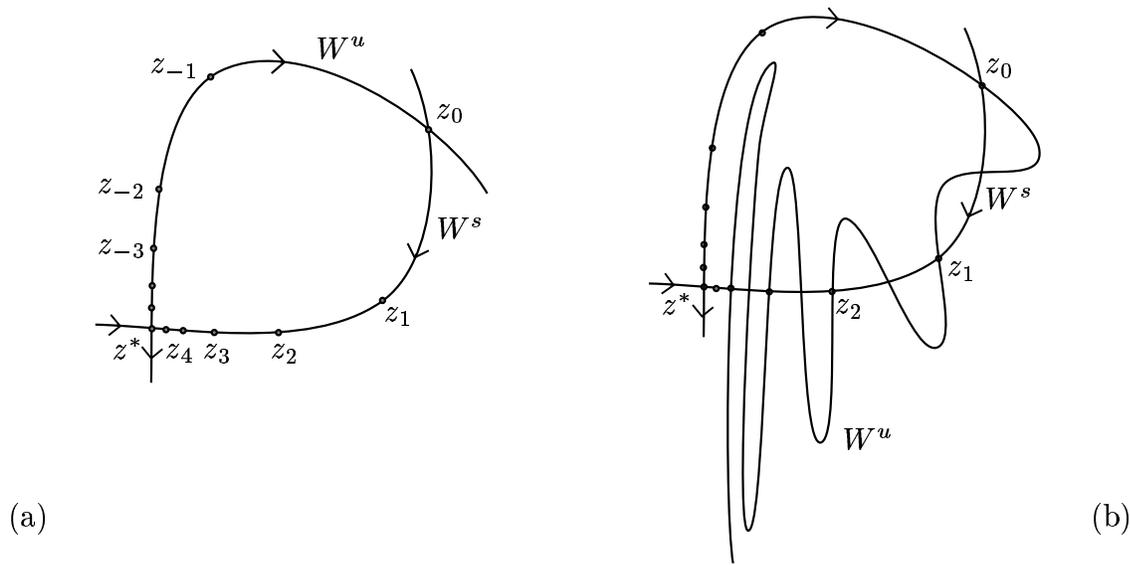


Figura 2.1. (a) Il punto omoclino z_0 e i suoi iterati $z_k = \Phi^k(z)$. (b) Le prime anse formate dalla varietà instabile W^u .

forma anse simili a quelle rappresentate in figura 2.1b (resta così intercalata una seconda orbita omoclinica, della quale però non ci occupiamo). Man mano che $z_k, k \rightarrow \infty$, si avvicina a z^* , i punti omoclini si addensano: la distanza tra l'uno e l'altro decresce esponenzialmente, grosso modo come a^k , ove $a < 1$ denota il minore dei due autovalori di $L := D\Phi_{z^*}$. Corrispondentemente, i tratti di varietà instabile W^u si dispongono “parallelamente” a W^u stessa (altrimenti W^u intersecherebbe se stessa!), accumulandosi su di essa, e le anse sempre più schiacciate necessariamente si allungano: grosso modo come b^k , se $b > 1$ denota l'altro autovalore di L .

Ora, come è illustrato in figura 2.2a, per k abbastanza grande accade necessariamente (su questo punto è bene riflettere un po'; si usa la continuità di Φ in un intorno di z^*) che W^u , correndo lungo W^u stessa, intersechi nuovamente W^s in due punti z'_0 e z''_0 : tali intersezioni sono anch'esse punti omoclini, benché non appartenenti alla famiglia, o orbita omoclinica $\{z_k\}$ sopra considerata (né a quella intercalata); anzi, d'ora in poi si formeranno punti omoclini nuovi a ogni iterazione, come i punti z'''_0 e z^{iv}_0 rappresentati in figura 2.2b. Tutti questi punti, per iterazione di Φ , daranno luogo a nuove famiglie di punti omoclini, e man mano che i punti si avvicineranno a z^* vi saranno corrispondentemente nuove anse di W^u , e ancora nuove intersezioni con W^s e nuove famiglie... con un livello di complicazione che è impossibile illustrare in un disegno, e forse anche immaginare. Analoga situazione si ha evidentemente per k negativo: questa volta è la varietà stabile W^s a formare le anse, e la situazione complessiva è quella vagamente delineata in figura 2.3. La figura 2.4 mostra invece le anse di W^s e W^u per il pendolo, quando i punti instabili corrispondenti a $-\pi, \pi, \dots$ si pensino distinti.

Tutto questo — è bene ricordarlo — deriva dalla “semplice” ipotesi che esista un punto omoclino trasverso z_0 , ed era già chiaro a Poincaré. Studi successivi di Smale ed altri, a partire dagli anni '60, hanno permesso di capire più a fondo i “fenomeni omoclini”, e anche di dare un senso preciso all'idea — che spontaneamente si affaccia alla mente di chiunque abbia visto figure come quelle che abbiamo sopra tracciato — che a tali fenomeni si accompagnino moti in qualche modo “caotici”.

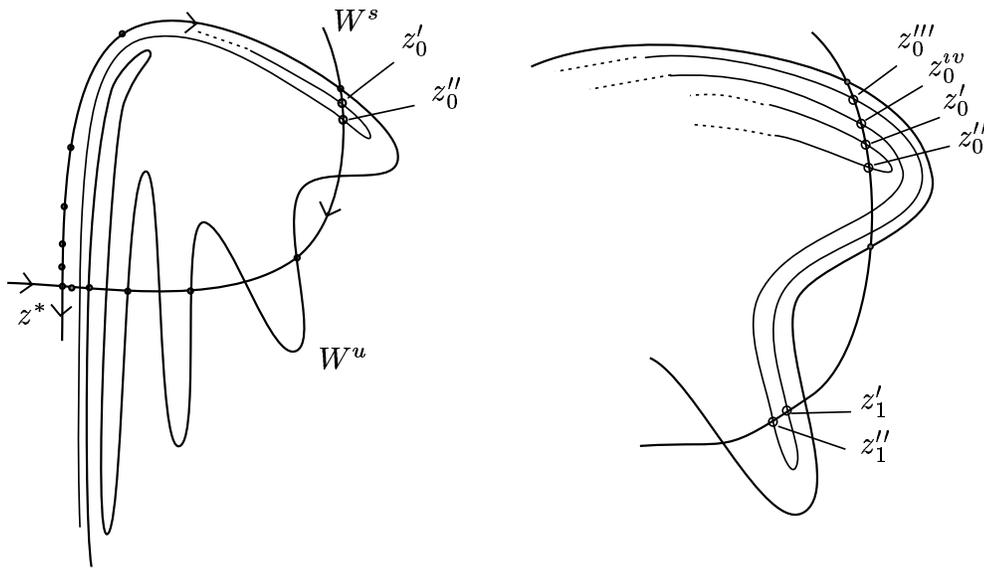


Figura 2.2. Anse successive di W^u , e nuovi punti omoclini (a). Dettaglio (b).

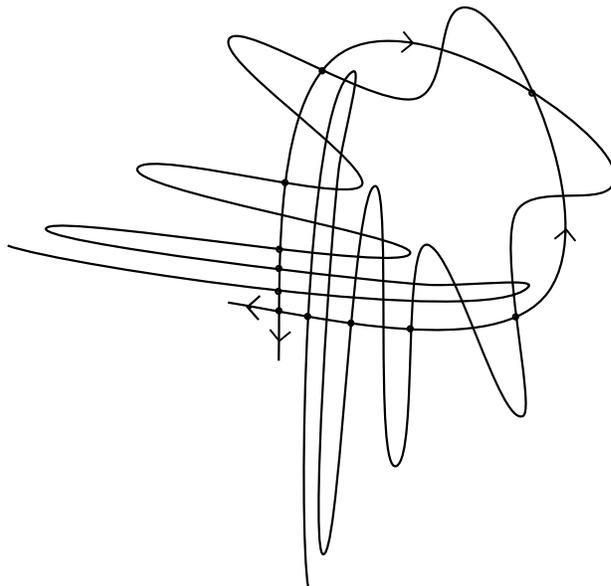


Figura 2.3. Una visione d'insieme di W^s e W^u .

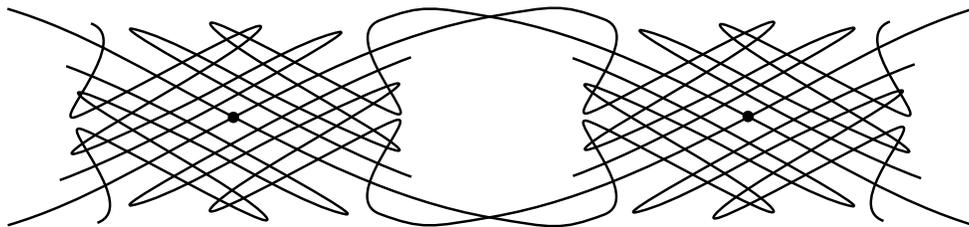


Figura 2.4. Il caso del pendolo.

2.2. Il metodo di Poincaré–Melnikov per il pendolo forzato

Consideriamo un pendolo forzato, di equazione

$$\ddot{q} + \omega^2 \sin q = \alpha F(q, t) ,$$

o equivalentemente

$$\dot{q} = p , \quad \dot{p} = -\omega^2 \sin q + \alpha F(q, t) ; \quad (2.1)$$

F si suppone regolare in entrambe le variabili, periodica di periodo 2π in q , periodica di periodo T in t . Il caso più semplice, per il quale svilupperemo i conti in dettaglio, è quello della semplice forzante sinusoidale,

$$F = \sin \Omega t , \quad T = 2\pi/\Omega . \quad (2.2)$$

Il sistema (2.1) è hamiltoniano, precisamente si ha

$$H_\alpha(p, q, t) = h(p, q) + \alpha f(q, t) , \quad (2.3)$$

con $h(p, q) = \frac{1}{2}p^2 - \omega^2 \cos q$, e f tale che $\frac{\partial f}{\partial q} = -F$; per la forzante (2.2) si ha in particolare¹¹⁾

$$f(q, t) = -q \sin \Omega t . \quad (2.4)$$

Denotiamo con $\{\Phi_\alpha^t, t \in \mathbb{R}\}$ il flusso dell'hamiltoniana H_α . Si osservi che pur essendo il sistema non autonomo, grazie alla periodicità in t delle equazioni è ben definita la mappa¹²⁾ $\Phi^T : M \rightarrow M$, con $T = \text{periodo}$, e si ha $(\Phi^T)^k = \Phi^{kT}$. Per $\alpha = 0$ sappiamo che vi sono i punti iperbolici $z_0^\pm = (0, \pm\pi)$, congiunti da una separatrice W_0 che è contemporaneamente varietà instabile dell'uno e varietà stabile dell'altro; per quanto visto nel paragrafo 1.5, per α piccolo vi saranno ancora due punti iperbolici¹³⁾ z_α^\pm , $z_\alpha^+ = z_\alpha^- + 2\pi$. Siano allora W_α^+ e W_α^- rispettivamente varietà stabile di z_α^+ e instabile di z_α^- . Il proposito è di mostrare che *se f soddisfa una opportuna condizione di genericità* (che sarà soddisfatta in particolare dalla (2.4)), e *se $\alpha \neq 0$ è sufficientemente piccolo, allora W_α^+ e W_α^- hanno un'intersezione trasversa* (con angolo $\mathcal{O}(\alpha)$); si veda la figura 2.5.

Per $\alpha = 0$ l'equazione della separatrice è evidentemente

$$h(p, q) = h(z_0^\pm) = \omega^2 ,$$

ovvero (limitandoci al ramo superiore della curva)

$$p = \eta_0(q) , \quad \eta_0(q) = 2\omega \cos \frac{q}{2} ;$$

¹¹⁾ Si osservi che H non è periodica in q (resta definita solo sul ricoprimento di S^1), ma ciò non ha rilevanza.

¹²⁾ Un buon modo di vedere le cose è quello di rendere autonomo il sistema, sostituendo il tempo nelle equazioni con una variabile τ , e completando le equazioni del moto con $\dot{\tau} = 1$. Lo spazio delle fasi appare tridimensionale (come è realmente: solo in questo spazio un dato iniziale determina univocamente il moto), e τ corre sul cerchio $\mathbb{R} \text{ mod } T$. La mappa Φ^T appare allora come mappa di Poincaré della sezione $\tau = 0$, che è M , in sé.

¹³⁾ Si osservi che z_α^\pm sono punti fissi per la mappa Φ_α^T , ma in generale non per il flusso (la corrispondente orbita è periodica di periodo T).

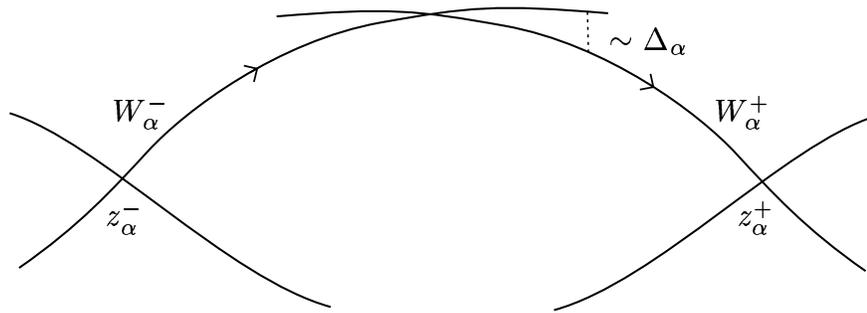


Figura 2.5. W_α^+ e W_α^- per il pendolo forzato.

per ogni q si ha allora

$$\lim_{k \rightarrow \infty} \Phi^{\pm kT}(\eta_0(q), q) = z_0^\pm .$$

Come per tutti i sistemi conservativi a un grado di libertà, il moto del pendolo (cioè la soluzione del sistema (2.1) con $\alpha = 0$) si determina con una quadratura. Per dati iniziali generici ci si imbatte nei cosiddetti integrali ellittici, non riducibili a funzioni più elementari. Fanno eccezione i dati iniziali sulla separatrice, per i quali invece la soluzione si riesce a esprimere esplicitamente in termini di funzioni esponenziali e trigonometriche: in particolare, per il dato iniziale $\bar{z} = (2\omega, 0) \in W_0$, denotando $\Phi_0^t = (P^t, Q^t)$, si trova¹⁴⁾

$$P^t(\bar{z}) = \frac{4\omega}{e^{\omega t} + e^{-\omega t}} , \quad Q^t(\bar{z}) = \pi - 4 \arctan e^{-\omega t} . \tag{2.5}$$

L'integrale di Poincaré–Melnikov.

Dalla proposizione 1.7 sappiamo che W_α^\pm sono curve regolari in α , che per $\alpha \neq 0$ piccolo si discostano poco dalla separatrice W_0 ; un minimo di riflessione (si riveda l'osservazione al termine della dimostrazione) mostra che al pari di W_0 tali curve si possono scrivere come grafico di una funzione, ovvero

$$p = \eta_\alpha^+(q) \quad \text{e} \quad p = \eta_\alpha^-(q)$$

(con η_α^\pm regolari), pur di escludere per la prima un intorno di z_α^- , e per la seconda un intorno di z_α^+ (un intorno piccolo per α piccolo). Per definizione di varietà stabile e instabile avremo

$$\lim_{k \rightarrow \infty} \Phi_\alpha^{\pm kT}(\eta_\alpha^\pm(q), q) = z_\alpha^\pm .$$

¹⁴⁾ Sulla separatrice si ha $\dot{q} = p = 2\omega \cos \frac{q}{2}$, e dunque prendendo $q = 0$ a $t = 0$ si trova facilmente

$$\begin{aligned} t &= \int_0^{Q^t} \frac{dq}{2\omega \cos \frac{q}{2}} = \frac{1}{2\omega} \int_0^{Q^t} \frac{dq}{\sin(\frac{\pi}{2} - \frac{q}{2})} = \frac{1}{4\omega} \int_0^{Q^t} \frac{dq}{\sin \frac{\pi-q}{4} \cos \frac{\pi-q}{4}} \\ &= \frac{1}{4\omega} \int_0^{Q^t} \frac{dq}{\tan \frac{\pi-q}{4} \cos^2 \frac{\pi-q}{4}} = -\frac{1}{\omega} \log \tan \frac{\pi - Q^t}{4} . \end{aligned}$$

Segue immediatamente l'espressione (2.5) di Q^t , e poi per derivazione ($p = \dot{q}$) quella di P^t .

Poniamo allora, nel dominio comune,

$$\begin{aligned}\Delta_\alpha(q) &= \frac{1}{2} [(\eta_\alpha^+(q))^2 - (\eta_\alpha^-(q))^2] \\ &= h(\eta_\alpha^+(q), q) - h(\eta_\alpha^-(q), q) ;\end{aligned}$$

si ha $\Delta_\alpha = \frac{1}{2}(\eta_\alpha^+ + \eta_\alpha^-)(\eta_\alpha^+ - \eta_\alpha^-)$, pertanto è chiaro che *ci sarà una intersezione trasversa di W_α^+ e W_α^- , se esiste un valore \tilde{q}_α tale che risulti $\Delta_\alpha(\tilde{q}_\alpha) = 0$, con derivata $\Delta'_\alpha(\tilde{q}_\alpha) \neq 0$* . A sua volta, lo studio di Δ_α è oggetto del seguente lemma:

Lemma 2.1. *Vale per Δ_α l'espressione*

$$\Delta_\alpha(q) = \alpha J(q) + \mathcal{O}(\alpha^2) , \quad J(q) = \int_{-\infty}^{+\infty} \{f, h\}(\Phi_0^t(\eta_0(q), q), t) dt , \quad (2.6)$$

ove $\{ , \}$ indica la parentesi di Poisson.¹⁵⁾

Il lemma vale in realtà in condizioni più generali, ogniqualvolta si ha un sistema del tipo (2.3), e per $\alpha = 0$ vi sono due punti iperbolici z_0^\pm connessi da una separatrice, purché, come nel caso del pendolo, anche per $\alpha \neq 0$ risulti $h(z_\alpha^+) = h(z_\alpha^-)$. Il caso tipico è quello in cui, come per il pendolo, i due punti z_α^\pm in realtà coincidono.

Dimostrazione. Per calcolare $\Delta_\alpha(q)$ servono $h(\eta_\alpha^+(q), q)$ e $h(\eta_\alpha^-(q), q)$. Si procede così: ricordando che $\dot{h} = \{h, H\} = \alpha\{h, f\}$, per ogni $k > 0$ si ha l'identità banale

$$h(\Phi_\alpha^{kT}(\eta_\alpha^+(q), q)) - h(\eta_\alpha^+(q), q) = \alpha \int_0^{kT} \{h, f\}(\Phi_\alpha^t(\eta_\alpha^+(q), q), t) dt ,$$

e prendendo il limite $k \rightarrow \infty$

$$h(z_\alpha^+) - h(\eta_\alpha^+(q), q) = \alpha \int_0^\infty \{h, f\}(\Phi_\alpha^t(\eta_\alpha^+(q), q), t) dt .$$

Grazie alla proposizione 1.7, a sinistra dell'uguale, e dunque a destra, c'è regolarità in α (cosa non ovvia se ci si limitasse a guardare il solo integrale illimitato). Uno sviluppo di Taylor troncato al primo ordine dà allora

$$h(z_\alpha^+) - h(\eta_\alpha^+(q), q) = \alpha \int_0^\infty \{h, f\}(\Phi_0^t(\eta_0(q), q), t) dt + \mathcal{O}(\alpha^2) .$$

Allo stesso modo si trova

$$h(z_\alpha^-) - h(\eta_\alpha^-(q), q) = \alpha \int_0^{-\infty} \{h, f\}(\Phi_0^t(\eta_0(q), q), t) dt + \mathcal{O}(\alpha^2) .$$

¹⁵⁾ La convenzione dei segni qui adottata è $\{f, h\} = \frac{\partial f}{\partial q} \frac{\partial h}{\partial p} - \frac{\partial f}{\partial p} \frac{\partial h}{\partial q}$.

Sottraendo, e ricordando che z_α^+ e z_α^- differiscono per una traslazione di 2π in q , cosicché $h(z_\alpha^+) = h(z_\alpha^-)$, segue immediatamente la (2.6); un cambio di segno è dovuto allo scambio di h e f nella parentesi di Poisson. ■

Grazie a questo lemma il problema dell'intersezione trasversa di W_α^+ e W_α^- si riporta allo studio dell'integrale J : se J ha uno zero trasverso (cioè con derivata prima non nulla) in \tilde{q}_0 , e α è sufficientemente piccolo, allora¹⁶⁾ $J + \mathcal{O}(\alpha)$ ha uno zero trasverso in \tilde{q}_α prossimo a \tilde{q}_0 , e corrispondentemente anche Δ_α ha uno zero trasverso in \tilde{q}_α . Si osservi che la derivata $\Delta'_\alpha(\tilde{q}_\alpha)$ sarà piccola con α , e dunque l'angolo con cui le varietà si intersecano sarà anch'esso piccolo con α , benché diverso da zero per $\alpha \neq 0$ (sufficientemente piccolo).¹⁷⁾ L'integrale J è comunemente chiamato *integrale di Poincaré–Melnikov*. La condizione di genericità su f , cui abbiamo accennato nelle considerazioni introduttive, è precisamente che l'integrale di Poincaré–Melnikov J abbia uno zero trasverso.

- **Osservazione.** Si osservi — ed è questa l'idea essenziale della dimostrazione, anzi di ogni metodo perturbativo — che *per il calcolo di J* (il primo ordine in α) è *sufficiente conoscere il moto del sistema imperturbato* (l'ordine zero in α), e anzi il solo moto sulla separatrice.

Il caso del pendolo con forzante sinusoidale.

Studiamo qui l'integrale di Poincaré–Melnikov nel caso del pendolo con forzante sinusoidale, precisamente per f della forma (2.4), cosicché si ha semplicemente

$$\{f, h\}(\Phi_0^t(\eta_0(q), q), t) = -P^t(\eta_0(q), q) \sin \Omega t .$$

Conviene qui eseguire un cambio di variabile da q a $\tau =$ tempo di percorrenza da $(\eta(0), 0)$ a $(\eta(q), q)$ lungo la separatrice, ovvero scrivere

$$q = Q^\tau(\bar{z}) , \quad \bar{z} = (2\omega, 0) \in W_0 ,$$

con Q^τ definito nella (2.5). Posto allora $\bar{J}(\tau) = J(Q^\tau(\bar{z}))$, usando l'espressione (2.5) per P^t si trova subito

$$\bar{J}(\tau) = - \int_{-\infty}^{+\infty} P^{t+\tau}(\bar{z}) \sin \Omega t \, dt = -4\omega \int_{-\infty}^{+\infty} \frac{\sin \Omega(t - \tau)}{e^{\omega t} + e^{-\omega t}} \, dt$$

(si tenga conto che $P^t(P^\tau(\bar{z}), Q^\tau(\bar{z})) = P^{t+\tau}(\bar{z})$). Si vede bene che $\bar{J}(\tau)$ si annulla per $\tau = lT$ e $\tau = \frac{1}{2}T + lT$, $l \in \mathbb{Z}$ (infatti, per questi valori di τ il numeratore è dispari in t , mentre il denominatore è pari). Quanto alla derivata, in $\tau = lT$ e $\tau = \frac{1}{2}T + lT$ si trova rispettivamente

$$\bar{J}' = \mp 4\omega \Omega \int_{-\infty}^{+\infty} \frac{\cos \Omega t}{e^{\omega t} + e^{-\omega t}} \, dt ,$$

¹⁶⁾ Del tutto in generale: se una funzione regolare $J : \mathbb{R} \rightarrow \mathbb{R}$ si annulla in \tilde{q} , con derivata $J'(\tilde{q}) \neq 0$, allora \tilde{q} è uno zero isolato di J , e inoltre ogni funzione J_α , α -vicina a J almeno in classe C^1 , per α piccolo ha uno zero \tilde{q}_α , con $J'(\tilde{z}_\alpha) \neq 0$.

¹⁷⁾ L'ipotesi che α sia piccolo è evidentemente indispensabile per poter affermare che il termine di ordine α^2 è trascurabile rispetto a quello di ordine α (in linea di principio, potrebbero esserci cancellazioni), e costituisce la maggior debolezza del metodo di Poincaré–Melnikov.

e l'integrale non si annulla; in effetti, l'integrale si calcola esplicitamente (col metodo dei residui) e si trova

$$\bar{J}' = \pm \frac{2\pi\Omega}{e^{\frac{\pi\Omega}{2\omega}} + e^{-\frac{\pi\Omega}{2\omega}}} \neq 0; \quad (2.7)$$

corrispondentemente si trova, per la derivata J' di J rispetto a q ,

$$J' = \frac{\bar{J}'}{P^\tau(\bar{z})} \neq 0.$$

Si osservi che, come ci si attende, la derivata J' cresce man mano che, avvicinandosi ai punti fissi, il denominatore diminuisce.

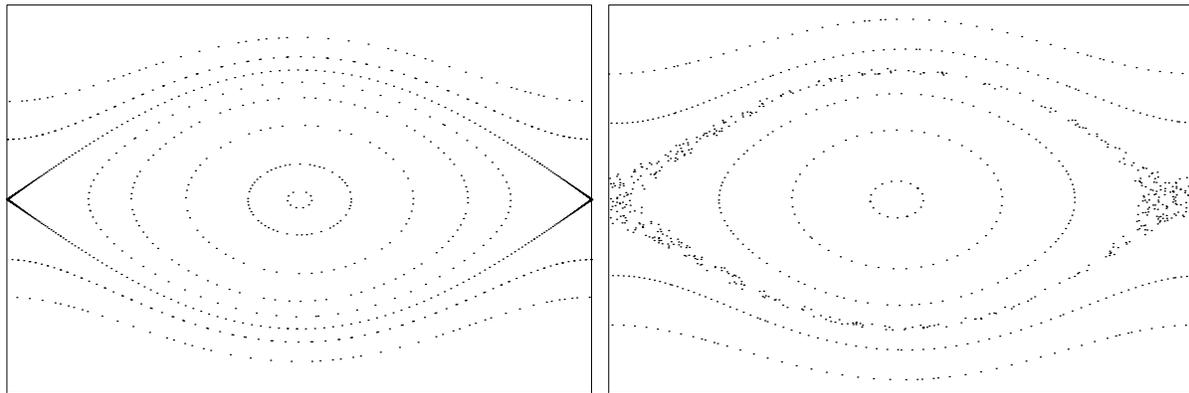
- **Osservazione.** Il caso del pendolo è particolarmente felice, perché tutto, a cominciare dal moto imperturbato sulla separatrice, si scrive esplicitamente in termini di funzioni elementari. Ma il metodo è del tutto generale, e riduce la dimostrazione di esistenza del punto omoclinico, con tutto quello che ne consegue, allo studio di funzioni che sono “semplici” integrali di funzioni elementari.
- **Osservazione.** La (2.7) mostra che \bar{J}' è diverso da zero per ogni scelta di Ω , ma per Ω elevato decresce esponenzialmente con Ω . Corrispondentemente l'intersezione di W^s e W^u è trasversa per ogni Ω , ma l'angolo di intersezione, pur di ordine α per ogni Ω , è tuttavia esponenzialmente piccolo in Ω .¹⁸⁾

Elementari simulazioni numeriche mostrano che effettivamente ai fenomeni omoclini si accompagnano, nel pendolo forzato e in altri sistemi, moti in qualche modo “complicati”. Ad esempio, la figura 2.6 mostra la mappa Φ del pendolo al tempo $T = 2\pi/\Omega$, per diversi valori di α e diversi dati iniziali. Si vede bene che per α piccolo, al di fuori di un sottile intorno delle separatrici,¹⁹⁾ i diversi punti di una stessa orbita sono ben allineati in curve invarianti, come nel caso imperturbato, ma attorno alle separatrici il comportamento è sostanzialmente diverso: una singola orbita sembra invadere una regione bidimensionale, la cui area cresce al crescere di α (in ciascuna figura, tutti i punti non allineati in curve appartengono a una singola orbita). La regione delle curve invarianti è detta *regione ordinata* del sistema, l'altra *regione caotica* (ma sono nomi dal valore euristico). La fenomenologia, con assoluta evidenza, è la stessa osservata nel capitolo precedente a proposito del modello di Hénon e Heiles (capitolo I, paragrafo 7.2); in effetti, si potrebbe vedere che il metodo di Poincaré–Melnikov qui descritto si può applicare, seppure con qualche complicazione in più, anche al modello di Hénon e Heiles.

Torneremo nei paragrafi successivi sul pendolo forzato, e daremo un senso più preciso all'idea di “moto caotico” basandoci, come vedremo, sulla dinamica simbolica; invece nel

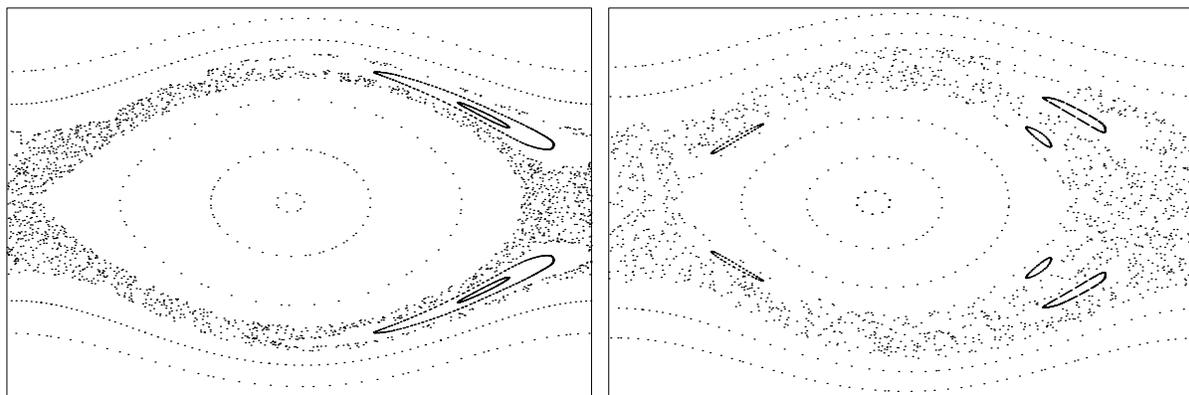
¹⁸⁾ In queste condizioni tuttavia l'applicazione del metodo è delicata, e conduce a un risultato in un certo senso povero: se al primo ordine in α l'angolo è $\sim \alpha e^{-c\Omega}$, $c = \frac{\pi}{2\omega}$, allora per poter trascurare gli ordini superiori (escludendo così cancellazioni spurie) bisogna supporre $\alpha \ll e^{-c\Omega}$. Ciò rende assai problematica l'applicazione del risultato ai casi (tipici in problemi fisici) in cui α e Ω , qui trattati come parametri indipendenti, sono in realtà funzioni l'uno dell'altro, e ad esempio si ha $\Omega = \alpha^{-\gamma}$, $\gamma > 0$. Questo problema, cui si accennerà nuovamente nel terzo capitolo a proposito della cosiddetta “diffusione di Arnol'd”, è di *grande* importanza, e negli ultimi anni si è visto uno sforzo non indifferente per andar oltre il risultato modesto qui riportato (è indispensabile, come si intuisce, tenere sotto controllo ogni ordine in α , e non solamente il primo).

¹⁹⁾ Delle separatrici, si intende, del sistema imperturbato.



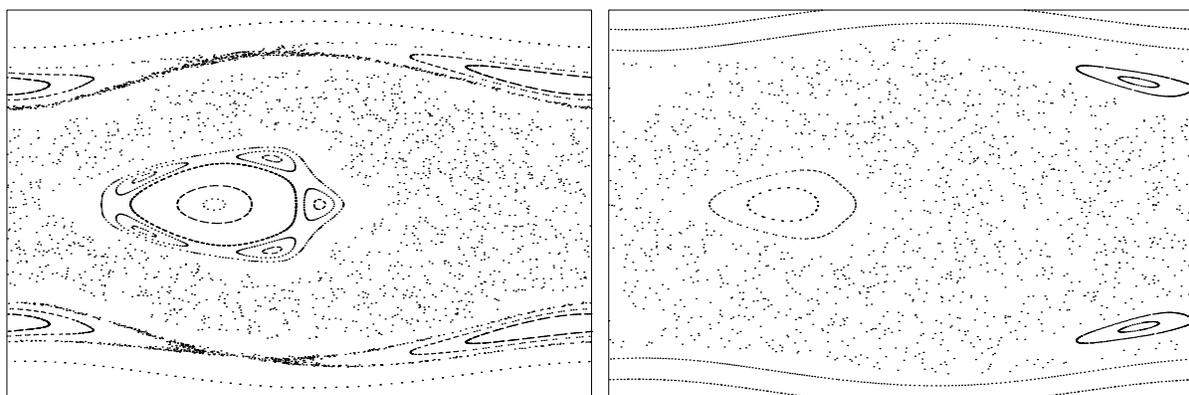
$\alpha = 0$

$\alpha = 0.05$



$\alpha = 0.1$

$\alpha = 0.3$



$\alpha = 1$

$\alpha = 1.5$

Figura 2.6. La mappa al tempo $2\pi/\Omega$ per il pendolo con forzante sinusoidale, con $\omega = 1$ e $\Omega = \sqrt{2}$, per diversi valori di α .

terzo capitolo, dedicato ai sistemi hamiltoniani, descriveremo in modo più preciso i “moti ordinati”. La differenza tra i due tipi di moto, in queste figure, è piuttosto evidente: ma sarebbe ancora più evidente in una simulazione “dal vivo”, nella quale si vedrebbero i punti della regione ordinata avanzare in modo perfettamente regolare su ciascuna curva invariante, e invece i punti della regione caotica saltellare da una parte all'altra, senza alcun ordine apparente. La differenza è infine confermata dal calcolo numerico del massimo esponente di Lyapunov, apparentemente nullo nella regione ordinata e invece positivo, e crescente con α , nella regione caotica.

- **Osservazione.** Come già si è detto, il metodo di Melnikov, qui impiegato per dimostrare la presenza delle intersezioni omocline nel pendolo forzato, è di tipo perturbativo, e pertanto la dimostrazione vale solo per $|\alpha|$ piccolo. Ma non c'è da stupirsi che di fatto per $|\alpha|$ grande i fenomeni omoclini non spariscano, e anzi, che al crescere di α la regione da essi interessata progressivamente si allarghi.

3. Insiemi iperbolici

3.1. Nozione di insieme iperbolico.

Consideriamo l'automorfismo algebrico di \mathbb{T}^2

$$\Phi \begin{pmatrix} x \\ y \end{pmatrix} = A \begin{pmatrix} x \\ y \end{pmatrix} \pmod{1}, \quad A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$$

(“gatto di Arnol'd”). Per ogni $z = (x, y) \in \mathbb{T}^2$ si vede immediatamente che, denotato con $E(z)$ lo spazio tangente a \mathbb{T}^2 in z , e con $E^s(z), E^u(z) \subset E(z)$ gli autospazi contraente e dilatante di $D\Phi_z = A$, restano soddisfatte le seguenti quattro proprietà:

- $E^s(z)$ e $E^u(z)$ sono trasversi, ovvero si ha

$$E(z) = E^s(z) \oplus E^u(z) ;$$

- $E^s(z)$ e $E^u(z)$ sono invarianti, precisamente

$$E^s(\Phi(z)) = D\Phi_z(E^s(z)), \quad E^u(\Phi(z)) = D\Phi_z(E^u(z))$$

(detto in altro modo: $\xi \in E^s(z) \Leftrightarrow D\Phi_z \xi \in E^s(\Phi(z))$, e analogamente per E^u);

- a $E^s(z)$ e $E^u(z)$ corrispondono contrazione e dilatazione esponenziale, precisamente esistono costanti positive²⁰⁾ C e λ , $\lambda < 1$, tali che

$$\xi \in E^s(z) \Rightarrow \|D\Phi_z^t \xi\| < C\lambda^t \|\xi\|, \quad \xi \in E^u(z) \Rightarrow \|D\Phi_z^{-t} \xi\| < C\lambda^t \|\xi\| \quad (3.1)$$

²⁰⁾ In questo esempio elementare si può evidentemente prendere $C = 1$.

per ogni $t > 0$;

iv. $E^s(z)$ e $E^u(z)$ dipendono con continuità²¹⁾ da z .

Queste quattro proprietà sono alla base della nozione di insieme iperbolico. Precisamente

Definizione. Sia (M, Φ) un sistema dinamico topologico di classe C^1 . L'insieme $\Lambda \subset M$, compatto e invariante, si dice *insieme iperbolico* di (M, Φ) , se per ogni punto $z \in \Lambda$ esistono $E^s(z), E^u(z) \subset E(z) := T_z M$, in modo che le proprietà i. – iv. sopra enunciate siano soddisfatte.²²⁾

Esempio 3.1. Per il Gatto di Arnol'd (per gli automorfismi algebrici di \mathbb{T}^n) l'insieme iperbolico Λ coincide con tutto M .

Definizione. (M, Φ) si dice *sistema di Anosov*, se il suo insieme iperbolico Λ coincide con tutto M .

Sia $z^* \in M$ un punto fisso iperbolico. Allora $\Lambda = \{z^*\}$ è un insieme iperbolico, ma assai poco interessante. Se però esiste un'intersezione omoclina z_0 , allora

Proposizione 3.1. *L'insieme*

$$\{z^*\} \cup \left(\bigcup_{t \in \mathbb{Z}} \Phi^t(z_0) \right)$$

(chiusura di un'orbita omoclina mediante l'aggiunta di z^*) è *iperbolico*.

Dimostrazione. Posto $z_t = \Phi^t(z_0)$, $t \in \mathbb{Z}$, prendiamo $E^s(z_t) =$ retta tangente a W^s in z_t , $E^s(z^*) =$ retta tangente a W^s in z^* , e similmente per E^u ; allora la trasversalità è evidente, e così anche l'invarianza (basta prendere archi di W^s e W^u come rappresentativi dei vettori tangenti). Il problema della continuità si pone solo per z^* , che è l'unico punto di accumulazione di Λ : ma è evidente che ad esempio $E^u(z_t)$, quando per $t \rightarrow \infty$ il punto z_t tende a z^* , assume in z_t tangente parallela a $E^u(z^*)$, altrimenti W^u intersecherebbe se stessa. L'unico punto non banale dunque è la verifica della proprietà iii. Per questo, preso $\xi \in E^s(z_t)$ qualsiasi, indichiamo con

$$\gamma_t := \frac{\|D\Phi_{z_t} \xi\|}{\|\xi\|}$$

il “fattore di contrazione” da z_t a z_{t+1} . Segue immediatamente, per ogni $z_s \in \Lambda$ e $\xi \in E^s(z_s)$,

$$\frac{\|D\Phi_{z_s}^t \xi\|}{\|\xi\|} = \prod_{k=s}^{s+t-1} \gamma_k ,$$

²¹⁾ Al di là dell'esempio, e anche in dimensione più elevata, la continuità di E^s, E^u si può definire così: ci si mette in una carta qualsiasi, e attraverso le coordinate si identificano tra loro tutti gli spazi tangenti $E(z)$ in punti z diversi; si dice che $E^s(z)$ e $E^u(z)$ dipendono con continuità da z , se preso un qualsiasi campo vettoriale costante ξ (in coordinate ha senso), e dette $\xi^s(z), \xi^u(z)$ le sue componenti su $E^s(z), E^u(z)$, queste sono funzioni continue di z . La scelta della carta non è rilevante. Nell'esempio la continuità è banale: E^s e E^u (nella carta naturale) sono addirittura costanti.

²²⁾ Per la continuità, resta evidentemente inteso che essa è richiesta, nel senso precisato nella nota precedente, solo per i punti di accumulazione di Λ .

e ci basta mostrare che esistono C, λ tali che per ogni s e ogni $t > 0$ (l'uniformità è essenziale!) si ha

$$\prod_{k=s}^{s+t-1} \gamma_k < C\lambda^t .$$

Tracciamo allora una palla \mathcal{U}_ε di raggio ε attorno a z^* ; solo un numero finito di punti di Λ , diciamo z_k con $k' \leq k \leq k''$, cade fuori di \mathcal{U}_ε , mentre tutti gli altri stanno in \mathcal{U}_ε . Per questi ultimi, in forza della continuità, γ_k (se ε è abbastanza piccolo) è prossimo a $\gamma^* =$ fattore di contrazione di $z^* =$ modulo del minore dei due autovalori di $D\Phi_{z^*}$, $\gamma^* < 1$ strettamente. Esiste pertanto $\lambda, 0 < \lambda < 1$, tale che $\gamma_k < \lambda < 1$ per $k < k'$ e $k > k''$. Posto allora

$$C = \prod_{k=k'}^{k''} \max\left(1, \frac{\gamma_k}{\lambda}\right) < \infty ,$$

si ha

$$\prod_{k=s}^{s+t-1} \gamma_k = \lambda^t \prod_{k=s}^{s+t-1} \left(\frac{\gamma_k}{\lambda}\right) < C\lambda^t ,$$

come richiesto. In modo analogo si procede per la direzione dilatante. ■

Un esempio particolarmente interessante di insieme iperbolico, per sé e in quanto tipico, sarà oggetto di studio del prossimo paragrafo.

3.2. Il “Ferro di Cavallo” di Smale.

Il “ferro di cavallo”, che ora andiamo a definire, è un sistema dinamico la cui definizione ricorda la trasformazione del panettiere vista nel capitolo sulla teoria ergodica; in esso però gioca un ruolo importante la dissipazione, ovvero la non conservazione del volume, inoltre a differenza della trasformazione del panettiere la dinamica non ha discontinuità.

Prendiamo un rettangolo chiuso R di lati α e β , e definiamo Φ come composizione di due trasformazioni Φ' e Φ'' , nel modo seguente:

- 1) per effetto di Φ' , R viene schiacciato e allungato fino a diventare un rettangolo R' di base $a\alpha$, $a > 2$, e altezza $b\beta$, $b \leq a^{-1} < \frac{1}{2}$;
- 2) per effetto di Φ'' , il nuovo rettangolo R' viene ripiegato in modo da intersecare R come indicato in figura 3.1, formando appunto un ferro di cavallo.

In figura 3.1 sono messi in evidenza due rettangoli Δ_0 e Δ_1 contenuti in R ; la loro unione è precisamente il sottoinsieme di R che dopo l'applicazione di Φ è inviato ancora in R . E' comodo prendere Φ in modo che su Δ_0 e Δ_1 agisca in modo lineare affine: diciamo

$$\begin{aligned} \Phi(x, y) &= (ax + c_0, by + d_0) && \text{per } (x, y) \in \Delta_0 \\ &= (-ax + c_1, -by + d_1) && \text{per } (x, y) \in \Delta_1 ; \end{aligned}$$

non c'è poi difficoltà a completare la definizione di Φ in modo che sia regolare a piacere (anche C^∞).

Formalmente (R, Φ) non è un sistema dinamico topologico, perché $\Phi(R)$ non è contenuto in R . Ma la cosa si aggiusta in un momento, come indicato in figura 3.2: si

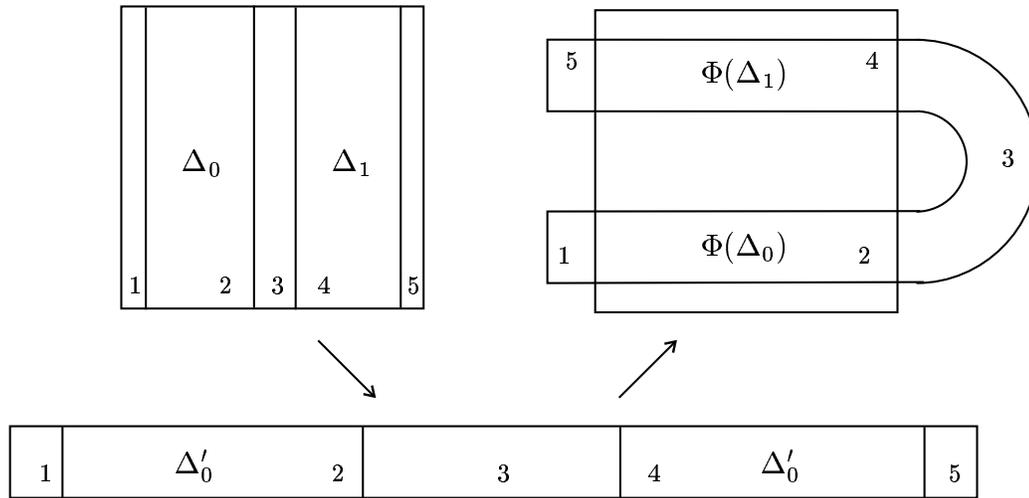


Figura 3.1. La mappa Φ del ferro di cavallo. I rettangoli Δ_0, Δ_1 sono inviati da Φ' in Δ'_0, Δ'_1 , e poi da Φ'' in $\Phi(\Delta_0), \Phi(\Delta_1)$.

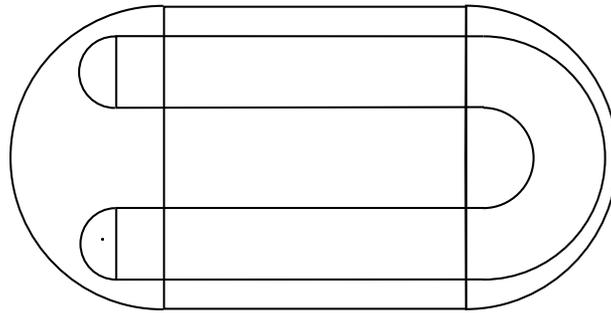


Figura 3.2. Il sistema dinamico (M, Φ) completato con le lunette.

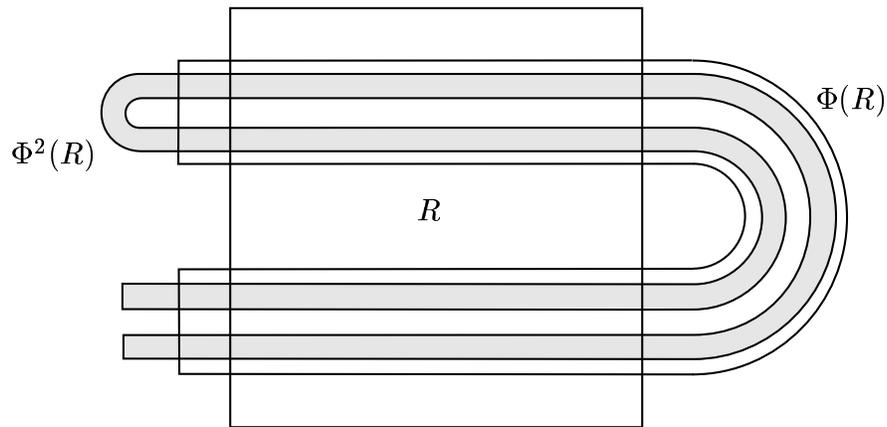


Figura 3.3. L'azione di Φ^2 su R .

aggiungono a R due lunette, costruendo così M a forma di campo sportivo, e come mostra la figura, si ottiene allora $\Phi(M) \subset M$. Ma la dinamica dei punti delle lunette è senza interesse: la lunetta di sinistra semplicemente si contrae in una lunetta più piccola (quella più bassa), e asintoticamente sarà risucchiata da un punto fisso (lemma della contrazione); quella di destra finisce a sinistra, e poi subisce la stessa sorte. Lo stesso si può dire di tutti i punti che, dopo una o più iterazioni, escono da R . Il sistema completato con le lunette è un sistema dinamico topologico, con però il requisito più debole di invertibilità di Φ solo sull'immagine $\Phi(M)$ contenuta strettamente in M . Nel seguito ometteremo le lunette. La figura 3.3 illustra l'azione di Φ^2 .

Consideriamo ora l'insieme Λ dei punti di R che non escono mai da R . Precisamente, poniamo

$$\Lambda^+ := \bigcap_{t \geq 0} \Phi^{-t}(R) = \{z \in R : \Phi^t(z) \in R \text{ per ogni } t \geq 0\},$$

poi

$$\Lambda^- := \bigcap_{t \geq 0} \Phi^t(R) = \{z \in R : \text{per ogni } t \geq 0 \text{ esiste } z' \in R \text{ tale che } \Phi^t(z') = z\},$$

e infine

$$\Lambda = \Lambda^+ \cap \Lambda^-.$$

L'insieme Λ per definizione è invariante, e su di esso Φ è invertibile; Λ è chiuso in quanto intersezione numerabile di chiusi.

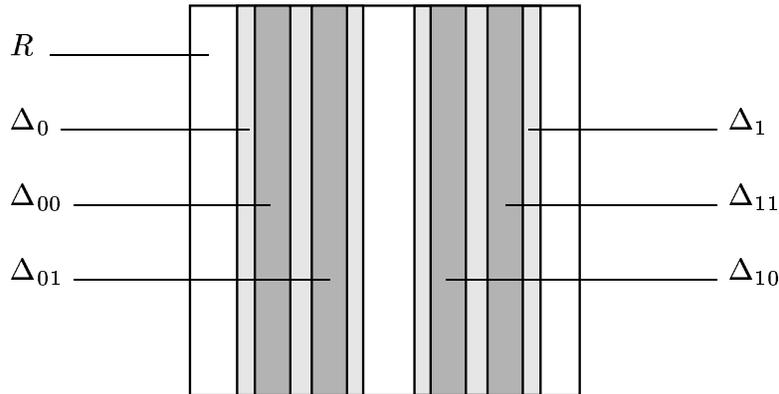


Figura 3.4. La costruzione di Λ^+ .

Per capire come è fatto Λ^+ , consideriamo

$$\Lambda_t^+ := \{z \in R : \Phi^s(z) \in R \text{ per } 0 < s \leq t\};$$

dalla definizione segue $\Lambda_{t+1}^+ \subset \Lambda_t^+$. E' chiaro che $\Lambda_1^+ = \Delta_0 \cup \Delta_1$, mentre Λ_2^+ è unione di quattro rettangolini $\Delta_{00}, \Delta_{01}, \Delta_{10}, \Delta_{11}$, di altezza β e larghezza $a^{-2}\alpha$, disposti come in figura 3.4. Un minimo di riflessione (la dimostrazione formale si fa per induzione) mostra che Λ_t^+ è unione di 2^t bande verticali di altezza β e larghezza $a^{-t}\alpha$, cosicché l'intersezione limite $\Lambda^+ = \bigcap_{t \geq 0} \Lambda_t^+$ è un insieme di infiniti segmenti verticali, che chiameremo fili, la cui sezione orizzontale ha la struttura di un insieme di Cantor ternario.

In modo analogo, posto

$$\Lambda_t^- := R \cap \Phi^t(R) ,$$

si vede (figura 3.1 e 3.3) che $\Lambda_1^- = \Phi(\Delta_0) \cap \Phi(\Delta_1)$, mentre Λ_2^- è costituito da quattro rettangolini di base α e altezza $b^2\beta$, con $\Lambda_2^- \subset \Lambda_1^-$ e dunque $\Lambda_{t+1}^- \subset \Lambda_t^-$ per ogni $t \geq 0$. Ci si convince facilmente che Λ_t^- è unione di 2^t bande orizzontali di larghezza α e altezza $b^t\beta$, e dunque che $\Lambda^- = \bigcap_{t \geq 0} \Lambda_t^-$ è un insieme di infiniti fili (segmenti) orizzontali, e tagliato verticalmente dà a sua volta un insieme di Cantor.

In conclusione, $\Lambda = \Lambda^+ \cap \Lambda^-$ è un doppio insieme di Cantor: un insieme non numerabile,²³⁾ con interno vuoto, totalmente disconnesso, che per costruzione è invariante sia sotto Φ che sotto Φ^{-1} . Per ogni $z \notin \Lambda$, la dinamica asintotica è banale: o Φ^t , per qualche $t > 0$, porta z nelle lunette, e la cosa finisce lì, oppure per qualche $t > 0$ il punto $\Phi^{-t}(z)$ non esiste. Infine, è immediata la verifica che Λ è *un insieme iperbolico*: per ogni $z \in \Lambda$ i sottospazi $E^s(z)$ e $E^u(z)$ sono (ovviamente) la retta verticale e la retta orizzontale per z , evidentemente trasversi e invarianti; la continuità è banale; e l'espansività è altrettanto ovvia, quando si osservi che si ha

$$D\Phi_z = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \quad \text{per } z \in \Delta_0 , \quad D\Phi_z = \begin{pmatrix} -a & 0 \\ 0 & -b \end{pmatrix} \quad \text{per } z \in \Delta_1 .$$

3.3. La dinamica simbolica per il ferro di cavallo.

Consideriamo lo spazio Σ delle stringhe doppiamente infinite di simboli 0 e 1,

$$\Sigma = \{0, 1\}^{\mathbb{Z}} = \{ \sigma = (\dots, \sigma_{-1}, \sigma_0, \sigma_1, \dots), \sigma_t = 0, 1 \} ;$$

munito della distanza

$$\delta(\sigma, \sigma') = \sum_{t \in \mathbb{Z}} 2^{-|t|} \delta(\sigma_t, \sigma'_t) , \quad \delta(\sigma_t, \sigma'_t) = \begin{cases} 0 & \text{per } \sigma_t = \sigma'_t \\ 1 & \text{per } \sigma_t \neq \sigma'_t \end{cases} .$$

Come dinamica su Σ prendiamo la traslazione a sinistra

$$(\Psi(\sigma))_t = \sigma_{t+1} , \quad t \in \mathbb{Z} .$$

Si verifica immediatamente che Ψ è continua con inversa continua, e dunque (Σ, Ψ) è un sistema dinamico topologico (di classe C^0).

Per proseguire introduciamo ora una nozione del tutto naturale:²⁴⁾

Definizione. Due sistemi dinamici topologici (M, Φ) e $(\widetilde{M}, \widetilde{\Phi})$ si dicono topologicamente *isomorfi*, o *coniugati*, se esiste un homeomorfismo $h : M \rightarrow \widetilde{M}$ tale che

$$\widetilde{\Phi} \circ h = h \circ \Phi . \tag{3.2}$$

²³⁾ Questo segue dalla costruzione stessa, che è topologicamente identica, in orizzontale e in verticale, alla costruzione dell'insieme ternario di Cantor. Con evidenza, ogni filo verticale di Λ^+ è in corrispondenza biunivoca a una stringa di simboli 0 e 1, e lo stesso per Λ^- .

²⁴⁾ Questa nozione si potrebbe più propriamente chiamare coniugazione in classe C^0 , e in modo analogo si potrebbe definire la coniugazione in classe C^r qualsiasi: di fatto però, in tutti i casi rilevanti, ci si deve accontentare della classe C^0 .

Consideriamo ora il ferro di cavallo, ristretto all'insieme iperbolico Λ ; su Λ prendiamo la metrica ereditata da \mathbb{R}^2 , più precisamente, per fissare le idee, facciamo riferimento alla distanza

$$d(z, z') = \max(|x - x'|, |y - y'|)$$

tra $z = (x, y)$ e $z' = (x', y')$, cosicché formalmente (Λ, Φ) è un sistema dinamico topologico (di classe C^0).

Definiamo $h : \Lambda \rightarrow \Sigma$ tramite la *dinamica simbolica*: precisamente, facendo riferimento ai rettangoli Δ_0, Δ_1 sopra definiti, diciamo che al punto $z \in \Lambda$ corrisponde la stringa $h(z) = (\dots, \sigma_{-1}(z), \sigma_0(z), \sigma_1(z), \dots) \in \Sigma$, tale che per ogni $t \in \mathbb{Z}$

$$\sigma_t(z) = l \quad \text{se} \quad \Phi^t(z) \in \Delta_l, \quad l = 0, 1$$

(la "storia" di z). Si verifica senza difficoltà che la corrispondenza tra Λ e Σ così definita è biunivoca. Precisamente:

- h è iniettiva, ovvero punti con la stessa storia coincidono. Infatti, per avere la stessa storia in avanti z e z' devono necessariamente trovarsi sullo stesso filo verticale: altrimenti la dinamica espande la separazione orizzontale fino a superare la larghezza di Δ_0 e Δ_1 , e la storia in avanti non può essere identica. Allo stesso modo, per avere la stessa storia all'indietro z e z' devono trovarsi sullo stesso filo orizzontale, dunque devono coincidere.
- h è suriettiva, ovvero tutte le stringhe (tutte le storie) sono realizzate. Infatti, presa una qualsiasi stringa $\hat{\sigma} \in \Sigma$, poniamo

$$A(\hat{\sigma}, k) = \{z \in \Lambda : \sigma_t(z) = \hat{\sigma}_t \quad \text{per} \quad |t| \leq k\};$$

in particolare, $A(\hat{\sigma}, 0) = \Delta_l$ se $\hat{\sigma}_0 = l$. Dalla definizione segue che, per ogni scelta di $\hat{\sigma}$, $A(\hat{\sigma}, 0), A(\hat{\sigma}, 1), \dots$ è sempre una successione di chiusi annidati, precisamente rettangolini di base $a^{-t}\alpha$ e altezza $b^t\beta$ collocati in posizione dipendente da $\hat{\sigma}$. Pertanto l'intersezione infinita è non vuota, ovvero c'è almeno un punto $z \in \Lambda$ (che per quanto visto sopra sarà unico) che ha per storia $\hat{\sigma}$.

Ancora senza difficoltà si verifica che h è continua, precisamente che, posto $\sigma = h(z)$, $\sigma' = h(z')$, si ottiene $d(z, z')$ piccola a piacere pur di prendere $\delta(\sigma, \sigma')$ sufficientemente piccola, e viceversa: infatti,

- se $\delta(\sigma, \sigma') < 2^{-t}$, allora le stringhe certamente coincidono tra $-t$ e t , ovvero i corrispondenti punti z e z' hanno la stessa storia tra $-t$ e t e dunque stanno nell'intersezione \mathcal{U} di una opportuna strisciolina verticale di larghezza $a^{-t}\alpha$ con una strisciolina orizzontale di altezza $b^t\beta$; per t grande allora il diametro di \mathcal{U} è arbitrariamente piccolo.
- Viceversa, pur di prendere $d(z, z')$ abbastanza piccola (precisamente: $d(z, z') < \rho_t$, ove ρ_t è la minima separazione tra le striscie verticali di Λ_t^+ e tra quelle orizzontali di Λ_t^-), si ottiene che z e z' stanno necessariamente in uno stesso rettangolino \mathcal{U} del tipo descritto sopra, e le loro storie coincidono tra $-t$ e t , perciò risulta $\delta(\sigma, \sigma') < \sum_{|k|>t} 2^{-|k|} = 2^{-t+1}$.

Infine, la (3.2) è soddisfatta per la definizione stessa di dinamica simbolica. In conclusione, abbiamo dimostrato che

Proposizione 3.2. *I sistemi dinamici (Λ, Φ) e (Σ, Ψ) sono coniugati dalla dinamica simbolica.*

- **Osservazione.** La scelta di Δ_0 e Δ_1 per fondare la dinamica simbolica non è obbligatoria: in particolare al loro posto si possono prendere, come è evidente, $\tilde{\Delta}_0 := \Phi(\Delta_0)$ e $\tilde{\Delta}_1 := \Phi(\Delta_1)$; ne approfitteremo tra non molto.

Vale la pena di mettere in evidenza alcune significative conseguenze della coniugazione.

- *Le orbite periodiche sono dense in Λ .* Infatti, sono evidentemente dense in Σ le stringhe periodiche; automaticamente allora sono dense le loro immagini in Λ . Si osservi però che per approssimare con precisione grande una sequenza non periodica con una sequenza periodica è necessario ricorrere a periodi grandi.
- *Λ contiene un'orbita densa, o come si dice, è topologicamente transitivo.* Infatti, l'insieme delle stringhe finite è numerabile, pertanto si possono allineare in un'unica stringa $\hat{\sigma}$ che a destra o a sinistra (o anche: a destra e a sinistra) le contiene tutte. Allora presa σ qualsiasi, e $k \geq 0$ arbitrario, esiste $t \in \mathbb{Z}$ tale che $\Psi^t(\hat{\sigma})$ coincide con σ tra $-k$ e k , dunque tale che $\delta(\Psi^t(\hat{\sigma}), \sigma) < 2^{-k+1}$; corrispondentemente, posto $\hat{z} = h^{-1}(\hat{\sigma})$, $z = h^{-1}(\sigma)$, si ha $d(\Phi^t(\hat{z}), z)$ piccola a piacere.
- *Comunque si prefissi una stringa di simboli 0 e 1, si trova $z \in \Lambda$ tale che $\Phi^t(z)$, $t \in \mathbb{Z}$, "obbedisce alla stringa":* precisamente, $\Phi^t(z)$ si trova in Δ_0 oppure in Δ_1 (o anche: in $\tilde{\Delta}_0$ oppure in $\tilde{\Delta}_1$) a seconda del simbolo 0 o 1 prefissato sulla stringa in posizione t .

3.4. Il meccanismo del ferro di cavallo nei fenomeni omoclini

Il principale motivo di interesse per il ferro di cavallo è che il suo meccanismo dello schiacciare – ripiegare, ripetuto indefinitamente, è un meccanismo tipico. Lo si ritrova, in particolare, nei fenomeni omoclini, come andiamo ora a illustrare.

Sia dunque (M, Φ) un sistema dinamico topologico di classe C^r , $r \geq 1$, con un punto fisso $z^* \in M$; ci restringiamo al caso $\dim M = 2$ e $\dim E^s = \dim E^u = 1$, e supponiamo che W^s e W^u si intersechino trasversalmente in un punto omoclineo z_0 . Introducendo coordinate opportune si può sempre supporre che la situazione sia quella illustrata in figura 3.5, con la varietà stabile W^s rettilinea e “verticale” tra z^* e z_0 , e la varietà instabile W^u che forma un arco come in figura, con due tratti rettilinei “orizzontali” in prossimità di z^* e z_0 (tutto questo in realtà non è affatto necessario, serve solo a far delle figure chiare). Prendiamo poi un rettangolo R attorno al segmento verticale z^*z_0 ; dopo un numero sufficiente $k \geq 1$ di iterazioni il rettangolo si è evidentemente contratto lungo W^s , e allungato a piacere lungo W^u . Un minimo di riflessione mostra che *con k opportuno, e con eventuali piccoli aggiustamenti²⁵⁾ ai lati di R , si può sempre ottenere che $\Phi^k(R)$ intersechi R come indicato in figura 3.5:* ovvero si può ottenere, per il sistema dinamico (M, Φ^k) , una dinamica a ferro di cavallo. Anzi: pur di prendere k abbastanza grande, si può prendere R stretto a piacere,

²⁵⁾ k deve essere abbastanza grande, perché $\Phi^k(R)$, allungandosi lungo W^u , arrivi a intersecare R vicino a z_0 ; gli aggiustamenti servono a evitare che $\Phi^k(R)$, seguendo le anse di W^u oltre z_0 , intersechi R in modo più complicato.

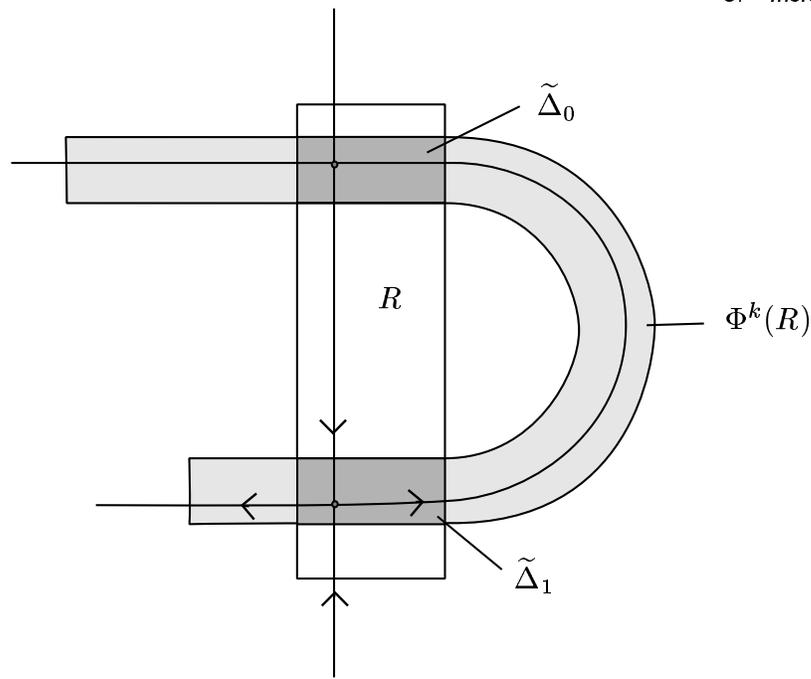


Figura 3.5. Il ferro di cavallo nei fenomeni omoclini.

in modo che i rettangoli $\tilde{\Delta}_0$ e $\tilde{\Delta}_1$ indicati in figura, intersezioni di $\Phi(R)$ con R stesso, siano piccoli a piacere. Applicando allora quanto si è detto per la dinamica simbolica del ferro di cavallo (e facendo qui riferimento agli insiemi $\tilde{\Delta}_0$ e $\tilde{\Delta}_1$), se ne deduce immediatamente un risultato rilevante che si può formalizzare nella seguente proposizione:

Proposizione 3.3. *Sia (M, Φ) , $\dim M = 2$, un sistema dinamico topologico di classe C^1 . Sia $z^* \in M$ un punto fisso, e supponiamo che le sue varietà asintotiche W^s e W^u , entrambe di dimensione uno, abbiano un'intersezione trasversa in un punto omoclinico z_0 . Allora comunque si prendano intorno U_0 e U_1 rispettivamente di z_0 e z^* , esiste $k > 0$ tale che per ogni stringa $\sigma = (\dots, \sigma_{-1}, \sigma_0, \sigma_1, \dots)$ di simboli $\sigma_l = 0, 1$ si trova $z \in M$ tale che*

$$\Phi^{kt}(z) \in U_l \quad \text{se } \sigma_t = l, \quad t \in \mathbb{Z}.$$

Vale la pena di sottolineare che questa insospettata complessità e varietà di comportamento delle orbite non richiede equazioni del moto eccezionalmente complicate, e si ritrova in “semplici” sistemi come la mappa del pendolo con forzante sinusoidale. Nel paragrafo 5.1, corollario 5.5, ritroveremo questo stesso risultato per una via diversa.

4. Attrattori e attrattori strani

4.1. Nozione di attrattore.

Sia (M, Φ) un sistema dinamico topologico.

Definizione. Un insieme chiuso $X \subset M$ si dice *attrattore* di (M, Φ) , se esiste U aperto contenente X tale che

- i. U è invariante in avanti, precisamente $\Phi(U) \subset U$;
- ii. risulta $\bigcap_{t \geq 0} \Phi^t(U) = X$.

L'insieme

$$\mathcal{B} := M \cap \left(\bigcup_{t \geq 0} \Phi^{-t}(U) \right)$$

si dice *bacino di attrazione* (o bacino di attrazione massimale) di X .

Si osservi che la prima proprietà implica che sia $\Phi^{t+1}(U) \subset \Phi^t(U)$ per ogni $t > 0$; dalla seconda proprietà segue poi immediatamente $\Phi(X) = \bigcap_{t \geq 1} \Phi^t(U) = X$, e dunque l'attrattore X è un insieme invariante. Il più semplice esempio di attrattore è un punto fisso iperbolico z^* , tale che lo spettro di $D\Phi_{z^*}$ sia tutto interno al disco unitario (un nodo stabile, oppure un fuoco stabile, in due dimensioni). Un altro caso semplice è quello del *ciclo limite stabile*, ovvero un'orbita periodica sulla quale vanno a collassare per $t \rightarrow \infty$ tutte le orbite ad essa inizialmente vicine; un esempio tra i più noti è quello dell'equazione di Van der Pol, introdotta negli anni '30 come modello matematico di un circuito elettrico auto-oscillante (si veda l'appendice A). Formalmente, anche i sistemi conservativi si può dire che abbiano un attrattore: X in tal caso coincide con tutta M , e anche \mathcal{B} coincide evidentemente con M . Il caso interessante però è quello in cui Φ *non* conserva il volume di M (o una misura ad esso equivalente).

Si è compreso relativamente di recente (anni '70; lavori importanti di Ruelle e Tackens, di Lorenz, di Hénon) che fenomeni fisicamente importanti come la turbolenza nei fluidi si possono comprendere (almeno un po') come dovuti alla presenza di un *attrattore strano*,²⁶⁾ diciamo un attrattore X che *non* è una sottovarietà di M (non è liscio, tipicamente ha una struttura di Cantor), e che è un insieme iperbolico, cosicché su di esso la dinamica può essere molto complessa e varia.²⁷⁾

²⁶⁾ Ci accontentiamo qui di una nozione vaga di attrattore strano, utile solo a comprendere l'esempio che segue; per una nozione più precisa si rinvia alla letteratura (peraltro non completamente concorde nella definizione).

²⁷⁾ Qui si aprirebbe un lungo discorso. Una delle idee (Landau) con cui si era sperato di spiegare la turbolenza era che essa fosse caratteristica dei sistemi a infiniti gradi di libertà, e consistesse sostanzialmente nella presenza di infiniti moti periodici sovrapposti. Ruelle e Tackens proposero un'idea diversa: la dissipazione fa collassare la dinamica su una sottovarietà M_0 di dimensione bassa (cosicché poche coordinate ben scelte bastano a descrivere il moto del sistema infinito dimensionale!), e all'interno di M_0 vi è un ulteriore collasso su un attrattore strano $X \subset M_0$, che con la sua dinamica ricca e complessa rende ragione dei moti complicati dei sistemi turbolenti. L'idea, molto innovativa, ebbe qualche difficoltà ad essere accolta (la pubblicazione dei lavori fu inizialmente rifiutata), ma alla fine prevalse. Qualche parola in più sull'argomento si può trovare in un gustoso

4.2. Il “Solenioide”

Il sistema dinamico che andiamo a descrivere, noto col nome di *Solenioide*, è un sistema dinamico topologico di classe C^∞ in \mathbb{R}^3 . Precisamente, poniamo

$$M := \{(z, w) \in \mathbb{C}^2 : |z| = 1, |w| \leq 1\} ;$$

M è il “toro pieno” in \mathbb{R}^3 , si veda la figura 4.1, prodotto di un cerchio, o anima del toro (su cui corre z), e di un disco, o sezione del toro (su cui corre w); le normalizzazioni a uno di $|z|$ e $|w|$ sono introdotte per praticità.

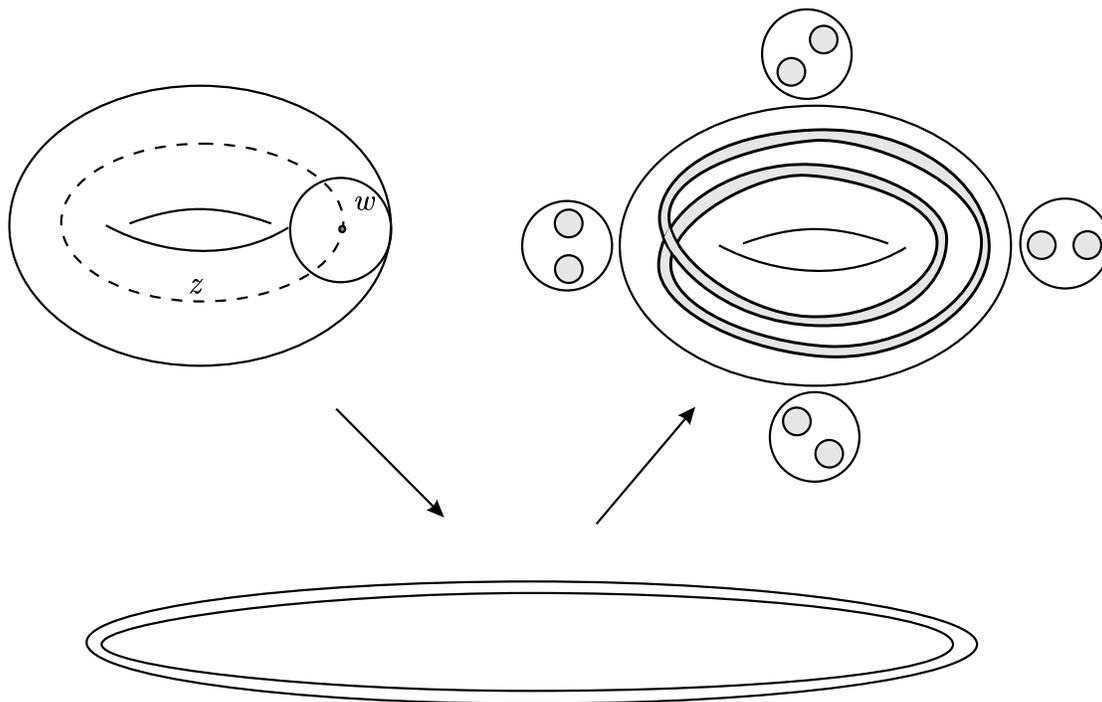


Figura 4.1. La mappa Φ del Solenoide.

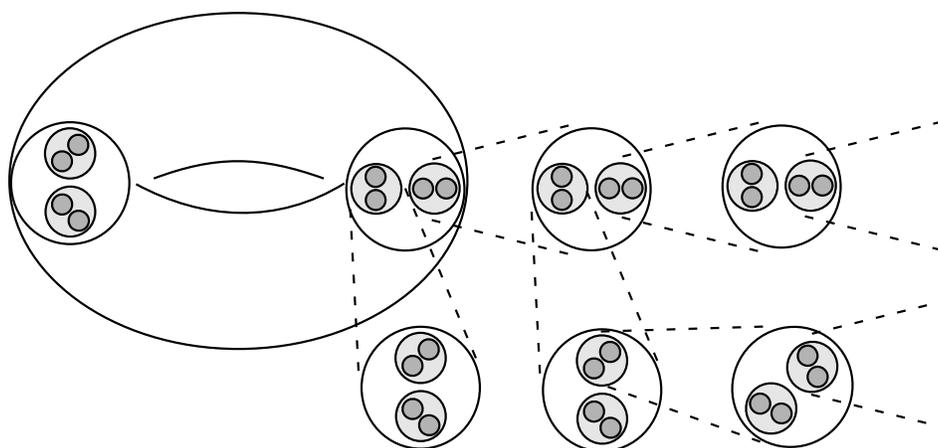


Figura 4.2. La sezione S_0 del Solenoide, con i dischi annidati.

libretto di Ruelle, che in linguaggio leggero e aneddotico comunica però idee profonde (D. Ruelle, *Caso e caos*, Bollati Boringhieri 1993).

Definiamo poi la mappa $\Phi : M \rightarrow M$ ponendo

$$\Phi(z, w) = (z^2, \frac{1}{4}w + \frac{1}{2}z) ;$$

equivalentemente, posto $z = e^{i\theta}$, $w = x + iy$, si ha

$$(\theta, x, y) \mapsto (2\theta \bmod 1, \frac{1}{4}x + \frac{1}{2} \cos \theta, \frac{1}{4}y + \frac{1}{2} \sin \theta) .$$

L'idea è quella di un elastico che si allunga di un fattore due e si strizza di un fattore (lineare) quattro,²⁸⁾ poi viene ripiegato su se stesso e immerso nel toro di partenza come illustrato in figura 4.1; si faccia attenzione in particolare alle quattro sezioni del toro corrispondenti a $\theta = 0, \frac{1}{2}\pi, \pi, \frac{3}{2}\pi$. (Al pari del ferro di cavallo, Φ è iniettiva ma non suriettiva.) Iterando Φ , si ottiene un elastico sempre più sottile e avvinghiato. E' conveniente, per comprendere il comportamento di Φ^t , considerare una sezione di M , ad esempio la sezione S_0 corrispondente a $\theta = 0$; la figura 4.2 mostra allora le tracce su S_0 delle prime iterate $\Phi^t(M)$. Un minimo di riflessione mostra che $\Phi^t(M)$ interseca S_0 in 2^t dischetti di raggio 4^{-t} , e che ad ogni iterazione si formano, all'interno di un dischetto, due altri dischetti più piccoli. Si ha così un annidamento un po' complesso: se con un indice a valori 0 o 1 si distinguono i due nuovi dischetti che nascono entro un dischetto assegnato,²⁹⁾ allora per ogni stringa di simboli 0 e 1 si ottiene una diversa successione di dischetti (chiusi) annidati, e ciascuna successione si contrae in un punto. Posto allora

$$X = \bigcap_{t \geq 0} \Phi^t(M) ,$$

si vede che i punti di $X_0 := X \cap S_0$ costituiscono un insieme non numerabile (sono in corrispondenza biunivoca con le stringhe a due simboli). Se invece della sola sezione guardiamo tutto il sistema, allora per ogni successione di simboli 0 e 1 si ha una successione di tubicini annidati, che vanno da un dischetto di S_0 al dischetto successivo ancora su S_0 , e si contraggono in un arco di curva, contenuto in X , che unisce un punto di X_0 a un successore. I punti di X sono dunque naturalmente organizzati in "fili" che si avvolgono entro il toro, e da questo viene il nome di Solenoide: ma (al contrario di quanto forse suggerisce l'intuizione) non si tratta di un unico filo, bensì di un insieme di fili non numerabile. La struttura di X è quella di un insieme di Cantor bidimensionale, moltiplicata per un cerchio. Il bacino di attrazione di X è con evidenza tutto M .

Si verifica senza difficoltà che X è un insieme iperbolico. Per questo basta osservare che, nelle coordinate θ, x, y , la mappa tangente a Φ è rappresentata da

$$L(\theta, x, y) = \begin{pmatrix} 2 & 0 & 0 \\ -\frac{1}{2} \sin \theta & \frac{1}{4} & 0 \\ \frac{1}{2} \cos \theta & 0 & \frac{1}{4} \end{pmatrix} ;$$

si vede bene allora che E^s coincide con il piano $\theta = \text{cost}$, e l'autovalore corrispondente è $\frac{1}{4}$, mentre E^u è una retta trasversa a E^s , ed ha autovalore 2; la direzione di E^u , non facile

²⁸⁾ Cosa fisicamente orribile per un elastico vero.

²⁹⁾ Ad esempio: 0 al primo e 1 al secondo, procedendo dall'asse x in senso antiorario.

da determinare, è tuttavia necessariamente tangente al filo in z , come segue dal fatto che archi di filo sono inviati da Φ in archi di filo.

Vale la pena di osservare, a conclusione dell'esempio, che la dinamica della coordinata θ è la “trasformazione dell'elastico” $\theta \mapsto 2\theta \bmod 2\pi$, un endomorfismo dotato di forti proprietà ergodiche, che ben si presta alla descrizione con la dinamica simbolica. Ad esempio, si può dividere il toro in due tramite i due piani meridiani $\theta = 0$ e $\theta = \pi$; presa allora ad arbitrio una stringa di simboli 0 e 1, si trova sempre un dato iniziale che nel suo collapsare sull'attrattore obbedisce alla stringa, precisamente il suo k -esimo iterato si trova nell'una o nell'altra metà del toro a seconda del simbolo prescelto in posizione k .

- **Osservazione.** L'attrattore del solenoide (per fissare le idee, è meglio avere in mente la sezione) è un esempio classico di insieme “frattale”: non è liscio in nessun punto, è auto-similare (una porzione ingrandita riproduce l'intero insieme), e gli si può assegnare, in un senso matematicamente preciso, una dimensione d_H , detta dimensione di Hausdorff, che tuttavia risulta essere non intera: precisamente $d_H = 1/2$ per la sezione, $d_H = 3/2$ per l'intero attrattore. Qualche considerazione generale sugli insiemi frattali, comprendente la nozione di dimensione di Hausdorff, è riportata nell'appendice B.

5. Il “lemma dell'orbita ombra” (*Shadowing lemma*)

5.1. Orbite e pseudo-orbite; enunciato del lemma.

Sia (M, Φ) un sistema dinamico topologico discreto.

Definizione. La successione $\{z_t, t \in \mathbf{Z}\}$ si dice essere una δ -pseudo-orbita per (M, Φ) , se risulta

$$\text{dist}(z_t, \Phi(z_{t-1})) \leq \delta \quad \text{per ogni } t \in \mathbf{Z} .$$

Una pseudo-orbita dunque è una collezione di punti che localmente (ma solo localmente) assomiglia a un'orbita vera. Globalmente può essere anche molto diversa: in particolare ad esempio una pseudo-orbita può contenere due volte lo stesso punto, senza per questo essere periodica. Si realizza in particolare una pseudo-orbita

- quando, ad esempio per fare un modello semplice di una situazione fisica complicata, si alterano di poco le equazioni del moto di un sistema: se l'alterazione è piccola, le orbite del sistema modificato sono pseudo-orbite per il sistema vero, e viceversa (si pensi, per fissare le idee, alla mappa a un tempo T fissato).
- Quando si introduce in un sistema dinamico un disturbo non conosciuto, eventualmente aleatorio, che si sa solo essere piccolo. Ciò è tipico dei sistemi fisici solo ap-

prossimativamente chiusi, in realtà soggetti a piccole perturbazioni da parte di oggetti esterni di cui non si vuol tenere conto.

- Quando si simula numericamente un sistema dinamico: anche nel caso felice in cui la mappa Φ sia assegnata a priori (e non debba a sua volta calcolata risolvendo numericamente sistemi di equazioni differenziali), gli errori di arrotondamento, dovuti alla precisione finita con cui i calcolatori rappresentano i numeri reali, fanno sì che le quantità calcolate siano non orbite vere, ma solo pseudo-orbite.

Le pseudo-orbite, dunque, sono oggetti assai comuni (e si potrebbe forse sostenere che sono i soli oggetti reali, mentre le orbite vere sono astrazioni).

Data una pseudo-orbita $\{z_t, t \in \mathbb{Z}\}$, è spontaneo chiedersi in che relazione essa stia con le orbite vere del sistema, $\{\Phi^t(z), t \in \mathbb{Z}\}$, con dato iniziale z uguale o vicino a z_0 . Esempi banali mostrano che orbite e pseudo-orbite possono avere un comportamento asintotico sostanzialmente diverso: si pensi a un sistema conservativo (un oscillatore armonico, un pendolo), al quale si aggiunge un piccolo termine di smorzamento, violando così una legge di conservazione; oppure si confrontino le mappe del cerchio $z \mapsto (z + \alpha) \bmod 1$, e $z \mapsto (z + \alpha + \delta) \bmod 1$, dopo un numero di iterazioni $\sim \delta^{-1}$.

Nel caso dei sistemi instabili, in particolare dei sistemi iperbolici, si aggiunge apparentemente una difficoltà ulteriore: la dinamica stessa amplifica a ritmo esponenziale gli errori (ne basta uno solo all'inizio!) e può sembrare che vi sia ancor meno speranza di trovare orbite vere che assomiglino a una data pseudo-orbita, perfino per tempi relativamente brevi. *Non è invece così*, come è garantito da una fondamentale proposizione che andiamo a enunciare, dopo aver premesso una semplice definizione.

Definizione. Si dice che il punto $z \in M$, ovvero la sua orbita $\{\Phi^t(z), t \in \mathbb{Z}\}$, ε -ombreggia³⁰⁾ la pseudo-orbita $\{z_t, t \in \mathbb{Z}\}$, se risulta

$$\text{dist}(\Phi^t(z), z_t) \leq \varepsilon \quad \text{per ogni } t \in \mathbb{Z} .$$

Si osservi che si richiede uniformità in t . Vale allora la proposizione seguente, detta *lemma dell'orbita ombra* (*Shadowing Lemma*) o *teorema di Anosov e Bowen*:

Proposizione 5.1. *Sia (M, Φ) un sistema dinamico topologico di classe C^1 , e sia $\Lambda \subset M$ un suo insieme iperbolico. Allora per ogni $\varepsilon > 0$ esiste $\delta = \delta(\varepsilon) > 0$ tale che ogni δ -pseudo-orbita $\{z_t, t \in \mathbb{Z}\}$, che sia anche δ -vicina a Λ :*

$$\text{dist}(z_t, \Lambda) \leq \delta \quad \forall t \in \mathbb{Z} ,$$

ammette un punto $z \in M$ che la ε -ombreggia. Se ε è abbastanza piccolo, $\varepsilon \leq \varepsilon_0$ con ε_0 opportuno indipendente dalla pseudo-orbita, allora z è unico.

Non daremo una dimostrazione del teorema, ma ci limiteremo a verificarlo, nel paragrafo 5.3, nel caso semplice del “Gatto di Arnol’d”. Nel prossimo paragrafo illustriamo invece la potenza della proposizione traendone qualche immediata conseguenza.

³⁰⁾ Segue come un'ombra: ingl. *to shadow*, ombreggiare ma anche pedinare, star incollato a qualcuno come un'ombra.

5.2. Alcune conseguenze del lemma dell'orbita ombra.

Corollario 5.2. *Sia $\varepsilon < \varepsilon_0$, e $\delta = \delta(\varepsilon)$ come nell'enunciato del lemma. Allora per ogni δ -pseudo-orbita periodica di periodo T esiste un'orbita vera periodica di ugual periodo T , ε -vicina alla pseudo-orbita.*

Dimostrazione. Sia $\{z_t, t \in \mathbb{Z}\}$ una δ -pseudo-orbita; per il lemma, esiste $z \in M$ che la ε -ombreggia, ovvero tale che $\text{dist}(\Phi^t(z), z_t) < \varepsilon$, e dobbiamo dimostrare che se la pseudo-orbita è periodica di periodo T allora anche l'orbita vera è periodica di periodo T , $\Phi^T(z) = z$. Per questo poniamo $z' = \Phi^T(z)$. Allora è evidente che per ogni t

$$\text{dist}(\Phi^t(z'), z_t) = \text{dist}(\Phi^{t+T}(z), z_t) = \text{dist}(\Phi^{t+T}(z), z_{t+T}) < \varepsilon ,$$

dunque anche z' ombreggia la stessa pseudo-orbita, e per l'unicità garantita da $\varepsilon < \varepsilon_0$ risulta $z' = z$. ■

Corollario 5.3. *Nei sistemi iperbolici conservativi le orbite periodiche sono dense.*

Dimostrazione. Prendiamo $\bar{z} \in M$ qualsiasi, e denotiamo con B_r la palla di raggio r centrata in \bar{z} . Fissato $\varepsilon < \varepsilon_0$, sia $\delta = \delta(\varepsilon)$ la quantità che interviene nell'enunciato del lemma dell'orbita ombra, che non è restrittivo supporre minore di ε ; vale il teorema della ricorrenza, dunque in particolare esistono $z_0 \in B_{\delta/2}$ e $T > 0$, tali che $\Phi^T(z_0) \in B_{\delta/2}$. Ma allora la successione di punti $\{z_t\}$ definita da

$$z_t = \Phi^t(z_0) \quad \text{per } 0 \leq t < T, \quad z_T = z_0 ,$$

poi ripetuta per periodicità ($z_{t+T} = z_t$ per ogni t) è una δ -pseudo-orbita periodica di periodo T (con un solo errore per periodo). Per quanto visto nel corollario precedente, esiste un'orbita vera periodica che la ombreggia, e il punto iniziale z , essendo ε -vicino a z_0 , sta in $B_{\varepsilon+\delta/2}$, dunque è arbitrariamente vicino a \bar{z} . ■

Già sapevamo che le orbite periodiche erano dense per gli automorfismi algebrici del toro (cap. I, esempio 2.5, esercizio), ma avevamo usato un trucco strettamente adattato al sistema. Come si vede invece il risultato è molto più robusto.

Corollario 5.4. *Esiste un intorno U dell'insieme iperbolico Λ nel quale la dinamica è separante: precisamente esistono U e $\varepsilon > 0$ tali che se per ogni $t \in \mathbb{Z}$ si ha $\Phi^t(z), \Phi^t(z') \in U$ e $\text{dist}(\Phi^t(z), \Phi^t(z')) < \varepsilon$, allora $z = z'$.*

Dimostrazione. Prendiamo ε uguale alla soglia ε_0 che garantisce l'unicità, e $U =$ intorno di Λ di raggio $\delta = \delta(\varepsilon_0)$. Le due orbite vere, evidentemente, sono anche δ -pseudo-orbite δ -vicine a Λ , e sia z che z' per ipotesi le ε -ombreggiano entrambe (ciascuna orbita ombreggia sia se stessa che l'altra). Per l'unicità, z e z' devono coincidere. ■

Corollario 5.5. *Vale la proposizione 3.3.*

Dimostrazione. Per evitare conflitti di notazione, denoteremo qui con \hat{z} , anziché z_0 , l'intersezione omoclina di riferimento, e con $\Phi^k(\hat{z})$, anziché z_k , i suoi iterati (l'orbita omoclina). Fissato $\varepsilon > 0$, sia $\delta = \delta(\varepsilon)$ come nell'enunciato del lemma; mostriamo che per

ogni prefissata stringa $\sigma = \{\sigma_t, t \in \mathbb{Z}\}$ di simboli 0 e 1, esistono $k = k(\delta) > 0$ e una δ -pseudo-orbita $\{z_s, s \in \mathbb{Z}\}$ tale che

$$z_{kt} = \hat{z} \quad \text{se } \sigma_t = 0, \quad z_{kt} = z^* \quad \text{se } \sigma_t = 1. \quad (5.1)$$

Ciò basta perché allora l'orbita vera che ε -ombreggia tale pseudo-orbita si trova, all'iterazione kt , a distanza ε da z_0 o z^* , in obbedienza all'arbitraria stringa σ . Per ottenere la (5.1) procediamo così: poiché $\varphi^t(\hat{z}) \rightarrow z^*$ per $t \rightarrow \pm\infty$, certamente esiste $m > 0$ tale che

$$\text{dist}(\Phi^t(\hat{z}), z^*) < \frac{\delta}{2} \quad \text{per } |t| \geq m;$$

poniamo allora $k = 2m$, e per soddisfare la (5.1) mostriamo che se si è trovata una stringa finita $\{z_s, |s| \leq kT\}$ che soddisfa la (5.1) per $t \leq T$, allora la stringa si può estendere a destra e a sinistra fino a $|s| = k(T+1)$, in modo da soddisfare ancora la (5.1). Cominciamo con la continuazione a destra, e osserviamo che dobbiamo considerare quattro casi, a seconda che sia $(\sigma_t, \sigma_{t+1}) = (0, 0), (1, 0), (0, 1), (1, 1)$, ovvero che si voglia la transizione da \hat{z}_0 a \hat{z} , ..., da z^* a z^* esattamente in k iterazioni.

- Transizione $\hat{z} \rightarrow \hat{z}$: partendo da $z_{kt} = \hat{z}$ si eseguono senza errori $m - 1$ iterazioni di Φ seguendo l'orbita omoclina, ovvero si pone $z_{kt+j} = \Phi^j(\hat{z}), 1 \leq j < m$; all'iterazione successiva, anziché su $\Phi^m(\hat{z})$, si salta su $\Phi^{-m}(\hat{z})$, commettendo un errore inferiore a δ ; poi con altre m iterazioni prive di errore si torna a \hat{z} .
- Transizione $\hat{z} \rightarrow z^*$: si procede come sopra, ma alla m -esima iterazione si salta su z^* , e lì si resta per altre m iterazioni.
- Transizione $z^* \rightarrow \hat{z}$: si sta fermi in z^* per $m - 1$ iterazioni; alla m -esima si salta su $\Phi^{-m}(\hat{z})$, e si giunge in altri m passi a \hat{z} .
- Transizione $z^* \rightarrow z^*$: banalmente si resta fermi in z^* .

Per la parte sinistra della stringa si procede evidentemente in modo analogo, e questo basta. ■

Si osservi che l'idea della dimostrazione è semplicissima: con errori ben mirati arbitrariamente piccoli vicino a z^* si controlla a piacere la pseudo-orbita.³¹⁾

Infine, proponiamo una variante a questo corollario, ovvero alla proposizione 3.3, adattata al pendolo forzato sinusoidalmente. Seguiamo qui le notazioni del paragrafo 2.2, ma per comodità pensiamo al pendolo vero, con q su S^1 e non su \mathbb{R} , cosicché i due punti iperbolici z_α^\pm della mappa Φ_α^T si identificano in un unico punto z_α^* .

Corollario 5.6. *Per α sufficientemente piccolo, comunque si prenda un intorno U di z_α^* esiste $\tau > 0$ tale che, in corrispondenza a una arbitraria stringa σ a tre simboli $l = 0, \pm 1$, si trova un dato iniziale z tale che $\Phi^{k\tau} \in U$ per ogni $k \in \mathbb{Z}$, e nel tempo tra $k\tau$ e $(k+1)\tau$ il pendolo resta in U se $\sigma_k = 0$, compie un giro in verso antiorario se $\sigma_k = 1$, in verso orario se $\sigma_k = -1$.*

³¹⁾ C. Simò sta cercando di applicare la stessa tecnica al controllo di un veicolo spaziale (un osservatorio solare) da collocarsi nel punto instabile di equilibrio tra la gravità della Terra e quella del Sole. Piccoli economici impulsi, che sfruttino bene il gioco delle separatrici, possono tenere il veicolo in prossimità della posizione prescelta, benchè instabile.

Omettiamo la dimostrazione, basata come la precedente sulla facile costruzione di una pseudo-orbita che obbedisce alla stringa: come prima, si seguono orbite vere omocline al di fuori di un piccolo intorno (di diametro δ) di z_α^* , e con un piccolo errore in prossimità di z_α^* si controlla in pratica a piacere la pseudo-orbita. Il requisito che α sia piccolo serve solo a garantire l'esistenza dell'intersezione omoclina, e con esso dell'insieme iperbolico e dell'applicabilità del lemma dell'orbita ombra.

5.3. Verifica del lemma dell'orbita ombra sul "Gatto di Arnol'd".

Il sistema, ricordiamo, è una mappa Φ del toro $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$ in sé, definita da

$$\Phi \begin{pmatrix} x \\ y \end{pmatrix} = A \begin{pmatrix} x \\ y \end{pmatrix} \pmod{1}, \quad A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix};$$

la matrice A ha autovalori λ e λ^{-1} , con $\lambda = \frac{1}{2}(3 - \sqrt{5}) < 1$.

Attorno a un punto \bar{z} qualsiasi del toro è conveniente introdurre coordinate locali ξ, η , con origine in \bar{z} e assi locali E^s e E^u . Se ξ, η sono le coordinate di z nel riferimento locale di \bar{z} , poniamo³²⁾

$$\text{dist}(z, \bar{z}) = \max(|\xi|, |\eta|);$$

in questo modo resta definita una distanza sul toro, che semplifica un po' la dimostrazione. Si osservi che con questa distanza la palla

$$U_\varepsilon(\bar{z}) := \{z \in M : \text{dist}(z, \bar{z}) \leq \varepsilon\}$$

è (nella metrica ordinaria) un quadrato chiuso di semilato ε , con i lati paralleli a E^s e E^u , centrato in \bar{z} .

Sia allora $\{z_t, t \in \mathbb{Z}\}$ una δ -pseudo-orbita, ovvero sia

$$\text{dist}(z_t, \Phi(z_{t-1})) < \delta \quad \text{per ogni } t \in \mathbb{Z}; \tag{5.2}$$

un punto z la ε -ombreggia se $\Phi^t(z) \in U_\varepsilon(z_t)$ per ogni t , ovvero se

$$z \in \Delta := \bigcap_{t \in \mathbb{Z}} \Phi^{-t}(U_\varepsilon(z_t)),$$

e si tratta di vedere che, se δ è abbastanza piccolo, allora l'intersezione è non vuota. L'unicità, per ε abbastanza piccolo, seguirà dalla costruzione stessa.³³⁾

Studiamo innanzitutto l'insieme

$$\Delta^+ := \bigcap_{t \geq 0} \Phi^{-t}(U_\varepsilon(z_t)),$$

³²⁾ Più precisamente: questa è la definizione della distanza tra punti abbastanza vicini. Se invece z è lontano da \bar{z} , si deve prendere la più piccola delle distanze tra \bar{z} e i punti di \mathbb{R}^2 equivalenti a z .

³³⁾ L'unicità per ε piccolo si può anche dedurre indipendentemente, sfruttando il fatto che il sistema già si sa essere espansivo: è evidente infatti che due diverse orbite vere, non potendo restare sempre troppo vicine tra loro, non possono neanche ombreggiare troppo bene la medesima pseudo-orbita.

e mostriamo che se $\delta \leq (1-\lambda)\varepsilon$, e se ε è sufficientemente piccolo, allora Δ^+ è un segmento di lunghezza 2ε parallelo a E^s , contenuto in $U_\varepsilon(z_0)$. A questo scopo, osserviamo che per ogni $t > 0$ l'insieme $\Phi^{-1}(U_\varepsilon(z_t))$ è un rettangolo di semilati $\lambda^{-1}\varepsilon$ e $\lambda\varepsilon$, paralleli rispettivamente a E^s e E^u , centrato in $\Phi^{-1}(z_t)$. D'altra parte, quest'ultimo punto è certamente vicino a z_{t-1} : precisamente, dette ξ_t, η_t le sue coordinate nel riferimento locale di z_{t-1} , la condizione (5.2) si traduce in

$$|\xi_t| < \lambda^{-1}\delta, \quad |\eta_t| < \lambda\delta.$$

Il rettangolo $\Phi^{-1}(U_\varepsilon(z_t))$ a sua volta sarà compreso tra le coordinate locali ξ_t^\pm e η_t^\pm , date da

$$\xi_t^\pm = \xi_t \pm \lambda^{-1}\varepsilon, \quad \eta_t^\pm = \eta_t \pm \lambda\varepsilon.$$

Supponiamo allora che per ogni $t > 0$ sia soddisfatta la condizione di inclusione illustrata in figura 5.1, precisamente che risulti

$$|\xi_t^\pm| > \varepsilon, \quad |\eta_t^\pm| < \varepsilon. \quad (5.3)$$

Se ne deduce facilmente, per ricorrenza, che l'insieme

$$\Delta_t^+ := \bigcap_{0 \leq s \leq t} \Phi^{-s}(U_\varepsilon(z_s))$$

è un rettangolo di semilati ε e $\lambda^t\varepsilon$, paralleli rispettivamente a E^s e E^u , contenuto in $U_\varepsilon(z_0)$. Infatti, per $t = 1$ la proprietà è vera, purchè ε sia abbastanza piccolo in modo da evitare che $\Phi^{-1}(U_\varepsilon(z_1))$ "girando attorno al toro" abbia con $U_\varepsilon(z_0)$ una intersezione più grande. Supposta allora vera l'affermazione per $t - 1$, dimostriamola per t , e per questo applichiamola alle $t - 1$ iterazioni tra 1 e t : ma allora

$$\bigcap_{0 \leq s \leq t-1} \Phi^{-s}(U_\varepsilon(z_{s+1}))$$

è un rettangolo di semilati ε e $\lambda^t\varepsilon$ contenuto in $U_\varepsilon(z_1)$; scrivendo

$$\Delta_t^+ = U_\varepsilon(z_0) \bigcap \Phi^{-1} \left(\bigcap_{0 \leq s \leq t-1} \Phi^{-s}(U_\varepsilon(z_{s+1})) \right),$$

e usando l'inclusione (5.3), la proprietà cercata segue immediatamente. Dalle proprietà di Δ_t^+ segue poi subito che l'intersezione infinita Δ^+ è il segmento cercato.

Ci siamo allora ridotti a verificare la (5.3), e si vede facilmente che a sua volta essa si riduce alle disuguaglianze

$$|\pm \lambda^{-1}\delta \pm \lambda^{-1}\varepsilon| > \varepsilon, \quad |\pm \lambda\delta \pm \lambda\varepsilon| < \varepsilon$$

(con scelta anche incrociata dei segni \pm), che sono verificate se

$$\lambda^{-1}(\varepsilon - \delta) > \varepsilon, \quad \lambda(\varepsilon + \delta) < \varepsilon.$$

Per entrambe la condizione $\delta < (1 - \lambda)\varepsilon$ è sufficiente.

Si procede in modo analogo per dimostrare che

$$\Delta^- := \bigcap_{t \leq 0} \Phi^{-t}(U_\varepsilon(z_t))$$

è un segmento di lunghezza 2ε parallelo a E^u , contenuto in $U_\varepsilon(z_0)$; se ne conclude che l'intersezione $\Delta = \Delta^+ \cap \Delta^-$ contiene uno e un solo punto, come richiesto.

Una condizione su ε sufficiente a garantire l'unicità è $\varepsilon < \varepsilon_0 = \frac{1}{2}\lambda$: il lato lungo di $\Phi^{-1}(U_\varepsilon(z_t))$ non supera 1, e dunque il rettangolo non ce la fa a girare attorno al toro.

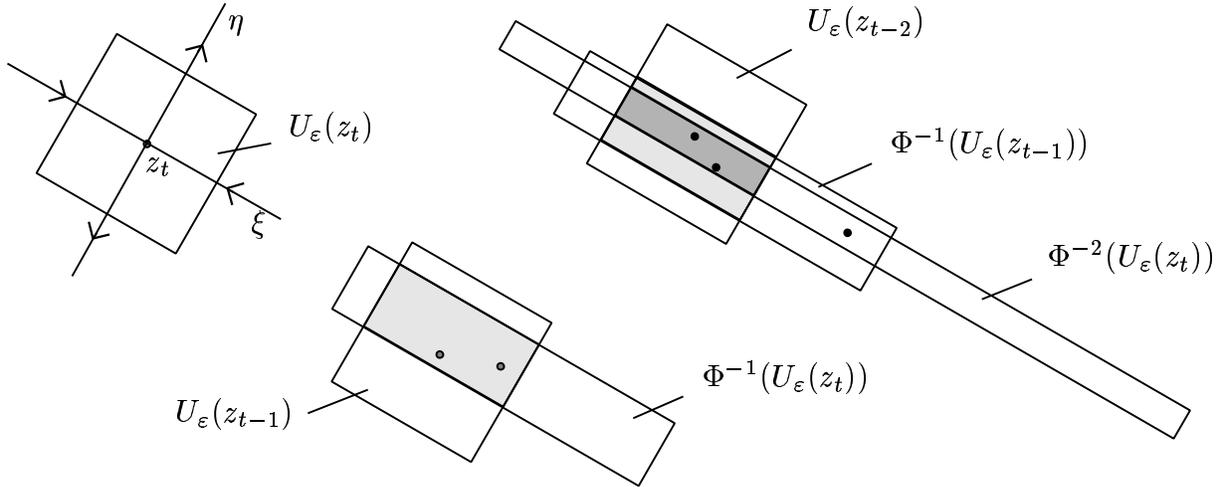


Figura 5.1. La costruzione di Δ_t^+ .

6. Proprietà ergodiche dei sistemi dinamici iperbolici (cenzo)

Consideriamo un sistema dinamico topologico (M, Φ) di classe C^r , $r \geq 1$; sia Λ un suo insieme iperbolico, e siano $E^s(z), E^u(z) \in T_z M$ i sottospazi dilatante e contraente di $z \in \Lambda$. Il teorema della varietà stabile, che abbiamo introdotto in connessione ai punti iperbolici, ha una naturale estensione agli insiemi iperbolici, precisamente si dimostra che a $E^s(z)$ e $E^u(z)$ corrispondono, su M , due varietà $W^s(z)$ e $W^u(z)$ definite da

$$W^s(z) = \{z' \in M : \text{dist}(\Phi^t(z'), \Phi^t(z)) \rightarrow 0 \text{ per } t \rightarrow \infty\}$$

$$W^u(z) = \{z' \in M : \text{dist}(\Phi^{-t}(z'), \Phi^{-t}(z)) \rightarrow 0 \text{ per } t \rightarrow \infty\},$$

con proprietà che generalizzano in modo naturale le proprietà delle varietà asintotiche dei punti iperbolici: precisamente,

Proposizione 6.1. *W^s, W^u sopra definite sono varietà di classe C^r immerse in M , di dimensione pari rispettivamente agli spazi E^s e E^u , e ad essi tangenti; sono invarianti, nel senso che per ogni t si ha*

$$\Phi^t(W^s(z)) = W^s(\Phi^t(z)), \quad \Phi^t(W^u(z)) = W^u(\Phi^t(z));$$

contraggono e dilatano esponenzialmente, precisamente esistono costanti positive C e $\lambda, \lambda < 1$, tali che

$$\text{dist}(\Phi^t(z'), \Phi^t(z)) < C \lambda^t \text{dist}(z', z) \quad \text{per } z' \in W^s(z)$$

$$\text{dist}(\Phi^{-t}(z'), \Phi^{-t}(z)) < C \lambda^t \text{dist}(z', z) \quad \text{per } z' \in W^u(z).$$

Consideriamo allora un sistema di Anosov (M, Φ) , cosicché il suo insieme iperbolico Λ coincide con tutta la varietà M , e supponiamo che, come per il Gatto di Arnol'd, vi sia una misura μ regolare conservata; in tal caso (M, μ, Φ) è un sistema dinamico classico. Si dimostra che

Proposizione 6.2. *Il sistema dinamico (M, μ, Φ) è ergodico.*

Interviene nella dimostrazione un lemma interessante:

Lemma 6.3. *Per un sistema dinamico classico, se la media temporale coincide con la media in fase quasi ovunque per tutte le funzioni continue, allora il sistema è ergodico.*

Dimostrazione. Supponiamo per assurdo che il sistema non sia ergodico. Allora esiste $A \subset M$ invariante non banale, $0 < \mu(A) < 1$. Ma A si può approssimare esternamente a piacere con un aperto B , la cui misura sarà ancora minore di uno. Con evidenza poi si può costruire una funzione continua $f : M \rightarrow \mathbb{R}$ (anche C^∞) che vale 1 in B , e approssima a piacere la funzione caratteristica di B : in particolare tale che la media $\langle f \rangle$ sia prossima a $\mu(B)$ e dunque minore di uno. Per tale funzione si ha però $\bar{f}(z) = 1$ in A di misura positiva, in contraddizione con l'ipotesi del lemma. ■

Veniamo ora alla dimostrazione della proposizione.

Dimostrazione. Ci limitiamo a una traccia della dimostrazione, che presenta un punto sottile. Grazie al lemma possiamo limitarci a verificare che la media temporale coincide quasi ovunque con la media in fase per ogni funzione continua f . L'idea, nota come "argomento di Hopf", è semplice: se z' è abbastanza vicino a z , allora $W^u(z')$ taglia trasversalmente $W^s(z)$ in un punto \tilde{z} (figura 6.1), e risulta

$$\begin{aligned} \text{dist}(\Phi^t(\tilde{z}), \Phi^t(z)) &\rightarrow 0 && \text{per } t \rightarrow +\infty \\ \text{dist}(\Phi^t(\tilde{z}), \Phi^t(z')) &\rightarrow 0 && \text{per } t \rightarrow -\infty . \end{aligned}$$

Se ne deduce facilmente che

$$\bar{f}^+(\tilde{z}) = \bar{f}^+(z) , \quad \bar{f}^-(\tilde{z}) = \bar{f}^-(z') ,$$

ove \bar{f}^\pm denotano le medie temporali in avanti e all'indietro di f , che per il teorema di Birkhoff–Kinchin esistono quasi ovunque, e quasi ovunque coincidono. Da quest'ultima proprietà segue³⁴⁾ che per quasi ogni $z \in M$, e quasi ogni z' vicino a z , $\bar{f}(z') = \bar{f}(z)$, e dunque la media temporale di f è quasi ovunque costante su M . Segue immediatamente, ancora dal teorema di Birkhoff–Kinchin, $\bar{f} = \langle \bar{f} \rangle = \langle f \rangle$. ■

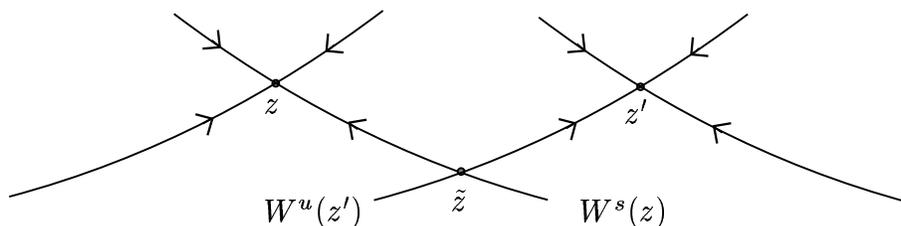


Figura 6.1. $W^u(z')$ interseca trasversalmente $W^s(z)$ in \tilde{z} .

³⁴⁾ Qui in realtà ci sarebbe un po' di lavoro, sul quale soprassediamo, per passare dal quasi ovunque in un intorno di $z \in M$ al quasi ovunque lungo le varietà (si usa il teorema di Fubini).

Veniamo ora al caso dissipativo, più precisamente al caso di sistemi dinamici con un attrattore iperbolico X di misura di Lebesgue nulla, eventualmente un attrattore strano come nel caso del “Solenioide” (paragrafo 4.2). Possiamo sempre pensare che il bacino di attrazione \mathcal{B} di X coincida con M (basta eventualmente restringere M); è allora evidente che qualunque misura μ conservata da (M, Φ) ha necessariamente supporto sull’attrattore.

- **Esercizio.** Si dimostri esplicitamente questa affermazione. [*Suggerimento:* se \mathcal{B} è bacino di attrazione di X , allora $\mathcal{B} \setminus \Phi(\mathcal{B})$ ha necessariamente misura nulla.]

Questa situazione è abbastanza spiacevole dal punto di vista fisico (ma non solo) se ci limitiamo ad applicare lo schema della teoria ergodica così come lo abbiamo descritto in precedenza: si finisce infatti con il non affermare nulla per un insieme assai rilevante di moti, tutti quelli con dato iniziale nell’insieme $M \setminus X$, che ha misura di Lebesgue piena pur essendo $\mu(M \setminus X) = 0$. A questa difficoltà si aggiunge un secondo fatto spiacevole, e cioè che, come abbiamo avuto modo di discutere a suo tempo, (M, Φ) ammetterà in generale diverse misure conservate (tutte a supporto sull’attrattore), e non si vede un criterio per preferirne una: sistemi con un attrattore (tipicamente: sistemi con turbolenza) sembrerebbero così mancare di una trattazione probabilistica “naturale”.

Entrambe le difficoltà si superano (Sinai, Ruelle e Bowen, anni ’70) introducendo nella teoria ergodica una variante significativa, nella quale la misura di Lebesgue nel bacino dell’attrattore (benché non conservata) assume un ruolo privilegiato. L’idea non è difficile: innanzitutto introduciamo una naturale

Definizione. Diremo che (M, Φ) è statisticamente regolare in $z \in M$, se per ogni funzione continua $f : M \rightarrow \mathbb{R}$ esiste la media temporale

$$\bar{f}(z) = \lim_{t \rightarrow \infty} \frac{1}{t} \sum_{s=0}^{t-1} f(\Phi^s(z)) .$$

Considerazioni di carattere generale (cioè del tutto indipendenti dall’esistenza o meno di un attrattore), nelle quali non entriamo, consentono di dedurre la seguente proposizione:

Proposizione 6.4. *Se (M, Φ) è statisticamente regolare in $z \in M$, allora esiste una misura μ_z su M , preservata dalla dinamica, tale che*

$$\bar{f}(z) = \langle f \rangle_{\mu_z} \tag{6.1}$$

per ogni $f : M \rightarrow \mathbb{R}$ continua.

L’idea è che la traiettoria stessa, col suo tempo di permanenza in un aperto, attribuisce una misura all’aperto. Si vede facilmente dalla (6.1) che punti z e z' sulla stessa orbita danno luogo alla stessa misura, e che se, da un certo t in poi, l’orbita non entra mai in un aperto A , allora A ha necessariamente misura nulla (infatti, ogni funzione continua f con supporto in A viene ad avere $\bar{f}(z) = \langle f \rangle_{\mu_z}$ nulla, dunque A deve avere misura nulla); segue in particolare che se z appartiene al bacino \mathcal{B} di un attrattore X , allora il supporto di μ_z è necessariamente contenuto nell’attrattore (come deve essere, perché μ_z possa essere invariante).

Definizione. Diremo che (M, Φ) è statisticamente regolare nel bacino \mathcal{B} di un attrattore $X \subset M$, se (M, Φ) è statisticamente regolare Lebesgue–quasi ovunque in \mathcal{B} , e inoltre per ogni f continua la media temporale \bar{f} è Lebesgue–quasi ovunque costante.

La locuzione “Lebesgue–quasi ovunque” significa, evidentemente, quasi ovunque rispetto alla misura di Lebesgue (in coordinate qualsiasi su M).

- **Osservazione.** La nozione di sistema statisticamente regolare in \mathcal{B} ricorda quella di sistema ergodico: ma il “quasi ovunque”, che nell’ordinaria nozione di ergodicità si riferisce alla misura μ conservata dalla dinamica, qui si riferisce invece alla misura di Lebesgue, e inoltre l’uguaglianza della media temporale alla media in fase è sostituita dall’uguaglianza delle medie temporali nei diversi punti tra di loro. Sulla base della proposizione 6.4, si può anche dire che per un sistema statisticamente regolare esiste una misura μ , con supporto sull’attrattore, tale che per ogni f continua si ha “quasi ovunque”, come per i sistemi ergodici,

$$\bar{f}(z) = \langle f \rangle_\mu ,$$

ma il “quasi ovunque” è riferito qui alla misura di Lebesgue in \mathcal{B} , e non, come nella teoria ergodica, alla misura μ preservata dalla dinamica.

- **Osservazione.** Se l’attrattore X ammette la decomposizione invariante $X = X_0 \cup X_1$, e corrispondentemente il bacino di attrazione di X ammette la decomposizione $\mathcal{B} = \mathcal{B}_0 \cup \mathcal{B}_1$ (con \mathcal{B}_i bacino di X_i , $i = 0, 1$), allora (M, Φ) può essere regolare in uno o anche in ciascuno dei \mathcal{B}_i , ma non in tutto \mathcal{B} (esempi elementari: due punti fissi attrattivi o cicli limite attrattivi disgiunti). La situazione ricorda i sistemi metricamente decomponibili della teoria ergodica. Tale situazione è esclusa se X contiene un’orbita densa, ovvero se, come si dice, è *topologicamente transitivo*.

Vediamo allora alcuni esempi elementari.

Esempio 6.1. Sia (M, Φ) un sistema dinamico topologico con un punto fisso attrattivo z^* . Allora per ogni z nel bacino di attrazione \mathcal{B} di z^* , e per ogni f continua, si ha

$$\bar{f}(z) = f(z^*) ;$$

perciò il sistema è statisticamente regolare in \mathcal{B} , e μ è la misura atomica che assegna misura uno a tutti gli insiemi contenenti z^* , misura zero a tutti gli altri.

Esempio 6.2. Sia $M \subset \mathbb{R}^2$ l’anello $\{(x, y) \in \mathbb{R}^2 : 1 - \rho < \sqrt{x^2 + y^2} < 1 + \rho\}$, $0 < \rho < 1$, ovvero in coordinate polari (r, θ)

$$M = \{(r, \theta) : 1 - \rho < r < 1 + \rho, \theta \in S^1\} .$$

Poniamo

$$\Phi(r, \theta) = (1 + \frac{1}{2}(r - 1), \theta + \alpha) , \quad \frac{\alpha}{2\pi} \text{ irrazionale} .$$

Il sistema ha evidentemente come attrattore X il circolo $r = 1$, e il suo bacino \mathcal{B} coincide con M . Si dimostra senza difficoltà che (M, Φ) è regolare in M , e la misura μ è la misura

di Lebesgue $\frac{1}{2\pi}d\theta$ sul circolo: ovvero, per ogni $f : M \rightarrow \mathbb{R}$ continua e ogni $(r, \theta) \in M$, risulta

$$\bar{f}(r, \theta) = \frac{1}{2\pi} \int_0^{2\pi} f(1, \theta') d\theta' . \quad (6.2)$$

Infatti, posto

$$\mathcal{B}_t = \{(r, \theta) \in M : |r - 1| < 2^{-t}\rho\} \subset \mathcal{B} ,$$

per ogni $z = (r, \theta) \in \mathcal{B}$ si ha $\Phi^t(z) \in \mathcal{B}_t$; sfruttando la continuità di f segue allora che per ogni $\varepsilon > 0$ esiste $K > 0$ tale che, se $t > K$,

$$|f(\Phi^t(r, \theta)) - f(1, \psi^t(\theta))| < \varepsilon ,$$

ove ψ è la traslazione su S^1 definita da $\psi(\theta) = \theta + \alpha \pmod{2\pi}$. Segue immediatamente

$$\frac{1}{T} \sum_{s=0}^{T-1} |f(\Phi^{s+K}(r, \theta)) - f(1, \psi^{s+K}(\theta))| < \varepsilon ,$$

e passando al limite $T \rightarrow \infty$, ricordando l'ergodicità della traslazione su S^1 , si ottiene subito la (6.2).

- **Osservazione.** La dimostrazione funziona perchè, come è immediato verificare, ogni punto dell'attrattore X attrae la sua varietà contraente, in questo caso un segmento radiale: tutte le orbite con dato iniziale (r, θ) , al variare di r , "collassano" sull'orbita dell'attrattore che ha dato iniziale in $(1, \theta)$. Questa stessa proprietà è alla base del prossimo meno banale esempio.

Esempio 6.3. Consideriamo il Solenoide (M, Φ) , descritto nel paragrafo 4.2; mostriamo che il sistema è statisticamente regolare in tutto M . Per questo sfruttiamo il fatto che:

- la dinamica per la coordinata $z = e^{i\theta}$ si scrive, per la variabile $\theta \in S^1$, nella forma $\theta \mapsto \psi(\theta) = 2\theta \pmod{2\pi}$, e il sistema dinamico (S^1, ν, ψ) , con ν misura di Lebesgue $\frac{1}{2\pi}d\theta$ su S^1 , è ergodico;
- per ogni fissato z , il disco $D := \{w \in \mathbb{C} : |w| \leq 1\}$ si contrae, dopo t iterazioni di Φ , in un dischetto di diametro 4^{-t} , cosicchè l'oscillazione di qualunque fissata funzione continua sulle immagini di D tende necessariamente a zero per t grande.

Precisamente, presa $F : M \rightarrow \mathbb{R}$ continua, poniamo

$$f_k^+(\theta) = \max_{w \in D} F(\Phi^k(e^{i\theta}, w)) , \quad f_k^-(\theta) = \min_{w \in D} F(\Phi^k(e^{i\theta}, w)) .$$

Per l'ergodicità di ψ si ha che, per dato iniziale θ al di fuori di un opportuno Θ_0 di misura nulla,

$$\lim_{t \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} f_k^\pm(\psi^t(\theta)) = \langle f_k^\pm \rangle_\nu$$

Ma per quanto detto sopra (punto ii) per ogni fissato $\varepsilon > 0$, se k è sufficientemente grande, si ha

$$f_k^+(\theta) - f_k^-(\theta) < \varepsilon . \quad (6.3)$$

Seguono senza difficoltà, per ogni z, w , le disuguaglianze

$$\langle f_k^+ \rangle_\nu \geq \limsup \frac{1}{T} \sum_{t=0}^{T-1} F(\Phi^t(z, w)) \geq \liminf \frac{1}{T} \sum_{t=0}^{T-1} F(\Phi^t(z, w)) \geq \langle f_k^- \rangle_\nu .$$

Infatti, f_k^+ e f_k^- sono sommabili, e ad esempio per la prima di esse

$$\begin{aligned} \langle f_k^+ \rangle_\nu &= \lim \frac{1}{T} \sum_{t=0}^{T-1} f_k^+(\psi^t(\theta)) && \text{[ergodicità di } (M, \nu, \psi)] \\ &= \lim \frac{1}{T} \sum_{t=0}^{T-1} \max_w F(\Phi^k(e^{i\psi^t(\theta)}, w)) && \text{[definizione di } f_k^+] \\ &\geq \limsup \frac{1}{T} \sum_{t=0}^{T-1} F(\Phi^{t+k}(z, w)) && \text{[si elimina max; lim sup per l'esistenza]} \\ &= \limsup \frac{1}{T} \sum_{t=0}^{T-1} f(\Phi^t(z, w)) && \text{[un tratto finito non contrib. alla media];} \end{aligned}$$

le ultime due relazioni si intendono verificate per ogni w . In modo analogo si deduce l'ultima disuguaglianza. Usando la (6.3), si conclude immediatamente che \limsup e \liminf sono arbitrariamente vicini, ovvero coincidono. Di conseguenza la media temporale

$$\bar{F}(z, w) = \lim \frac{1}{T} \sum_{t=0}^{T-1} F(\Phi^t(z, w))$$

esiste ed è indipendente da (z, w) in un insieme di misura piena, precisamente per $(z, w) \in M \setminus M_0$ con $M_0 = \Theta_0 \times D$

Il sistema (M, Φ) è così statisticamente regolare in M . Si potrebbe dimostrare (usando in modo essenziale la dinamica simbolica) che la misura μ implicitamente definita sull'attrattore X è così costruita:

- si scrive $X = [0, 2\pi) \times \tilde{X}$, ove \tilde{X} è un sottoinsieme della sezione $\theta = 0$ con struttura di Cantor, in corrispondenza biunivoca (nel modo visto precedentemente) con le stringhe semi-infinita $\sigma = \{\sigma_0, \sigma_1, \dots\}$;
- si pone $\mu = \nu \times \tilde{\mu}$, ove ν è la misura di Lebesgue sopra utilizzata su S^1 , mentre $\tilde{\mu}$ è una misura su \tilde{X} corrispondente, nello spazio di stringhe, alla misura dello schema di Bernoulli $B_{\frac{1}{2}, \frac{1}{2}}$.

L'esempio del Solenoide è generalizzato dalla seguente rilevante

Proposizione 6.5 (Teorema di Sinai-Ruelle-Bowen). *Sia (M, Φ) un sistema dinamico topologico, e sia $X \in M$ un attrattore iperbolico. Se X contiene un'orbita densa, allora (M, Φ) è statisticamente regolare nel bacino di attrazione B di X .*

Questa proposizione, si può dire, costituisce la base per una trattazione probabilistica dei sistemi dinamici dissipativi.

APPENDICI

A. L'equazione di Van der Pol

L'equazione di Van der Pol, proposta negli anni '30 per far da modello alle oscillazioni che si instaurano in certe condizioni in un circuito elettrico amplificato, si scrive

$$\ddot{x} + \varepsilon(x^2 - 1)\dot{x} + x = 0, \quad \varepsilon > 0, \quad x \in \mathbb{R}, \quad (\text{A.1})$$

o equivalentemente

$$\dot{x} = v, \quad \dot{v} = -x - \varepsilon(x^2 - 1)v; \quad (\text{A.2})$$

nella sostanza, è un oscillatore armonico smorzato di pulsazione uno, con coefficiente di smorzamento $\mu(x, \varepsilon) = \varepsilon(x^2 - 1)$ positivo per $|x| > 1$ (attrito ordinario), e negativo per $|x| < 1$. La forma semplice di μ consente di eseguire i calcoli in dettaglio, ma è sostituibile con una classe assai più vasta di funzioni.

Daremo due dimostrazioni dell'esistenza di un ciclo limite stabile per l'equazione di Van der Pol: una di tipo perturbativo, valida solo per ε piccolo (ma interessante, perché costituisce una delle più semplici applicazioni di un metodo tipico della teoria delle perturbazioni) e una seconda più generale, valida per ε positivo qualsiasi.

Cominciamo allora con il caso di ε piccolo, e mostriamo che

Proposizione A.1. *Se $\varepsilon > 0$ è sufficientemente piccolo, allora esiste un ciclo limite stabile prossimo al circolo di raggio 2, di periodo prossimo a 2π .*

Dimostrazione. Per $\varepsilon = 0$ si ha un oscillatore armonico di pulsazione uno, e per ogni dato iniziale (x, v) la traiettoria è un circolo percorso con velocità uniforme, nel periodo $T_0 = 2\pi$. Tale moto appare semplice in coordinate polari. Al posto del raggio tuttavia è più conveniente usare l'energia dell'oscillatore armonico $E = \frac{1}{2}(x^2 + v^2)$, cosicché il cambiamento di coordinate si scrive

$$x = \sqrt{2E} \cos \theta, \quad v = \sqrt{2E} \sin \theta. \quad (\text{A.3})$$

Le equazioni del moto nelle variabili (E, θ) si ottengono facilmente: si trova³⁵⁾

$$\dot{E} = \varepsilon f(E, \theta), \quad \dot{\theta} = -1 + \varepsilon g(E, \theta), \quad (\text{A.4})$$

con³⁶⁾

$$f(E, \theta) = -2E \sin^2 \theta (2E \cos^2 \theta - 1), \quad g = \sin \theta \cos \theta (2E \cos^2 \theta - 1) \quad (\text{A.5})$$

³⁵⁾ Si vede bene che (per ε piccolo) E è una *variabile lenta*, mentre θ è una *variabile veloce*. La distinzione delle variabili in lente e veloci (come avremo modo di vedere ampiamente nel terzo capitolo) è la vera anima di ogni procedimento perturbativo.

³⁶⁾ Per f la deduzione è immediata: \dot{E} è la potenza delle forze dissipative, $\dot{E} = -\varepsilon(x^2 - 1)v^2$; usando le (A.3) segue subito l'espressione (A.5) di f .

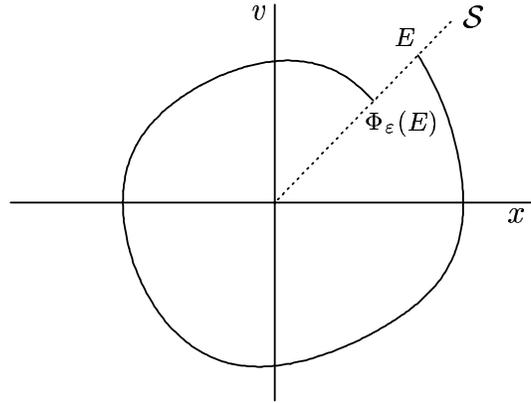


Figura A.1. La semiretta \mathcal{S} come sezione di Poincaré.

(l'espressione di g , aggiunta qui per completezza, di fatto non giocherà alcun ruolo).

Fissiamo ora l'attenzione su una qualunque semiretta \mathcal{S} uscente dall'origine nel piano xv (figura A.1), con $\theta = \bar{\theta}$ assegnato. Per $\varepsilon = 0$ tutte le traiettorie tagliano \mathcal{S} trasversalmente, e impiegano il tempo T_0 a fare un giro completo (\mathcal{S} è una buona sezione di Poincaré). Per ε piccolo e dato iniziale $(E, \bar{\theta})$ su \mathcal{S} , se denotiamo con $\mathcal{E}_\varepsilon(t, E)$, $\Theta_\varepsilon(t, E)$ la corrispondente soluzione, si ha evidentemente³⁷⁾

$$\mathcal{E}_\varepsilon(t, E) = E + \mathcal{O}(\varepsilon) , \quad \Theta_\varepsilon(t, E) = \bar{\theta} - t + \mathcal{O}(\varepsilon) ,$$

e dunque la traiettoria interseca nuovamente \mathcal{S} in un punto $E' = \Phi_\varepsilon(E)$, con $\Phi_0(E) = E$ e dunque

$$\Phi_\varepsilon(E) = E + \varepsilon F_\varepsilon(E) ,$$

dopo un tempo $T_\varepsilon = T_0 + \mathcal{O}(\varepsilon)$. La ricerca di un'orbita periodica si riduce allora alla ricerca di un punto fisso di Φ_ε , ovvero di uno zero di F_ε .

Si verifica ora senza difficoltà che si ha

$$F_\varepsilon(E) = -aE(E - b) + \mathcal{O}(\varepsilon) \tag{A.6}$$

con $a, b > 0$, precisamente $a = \pi$, $b = 2$. Infatti, scrivendo in forma integrale l'equazione per E , e approssimando la soluzione per $\varepsilon \neq 0$ con quella per $\varepsilon = 0$, si trova subito

$$\begin{aligned} F_\varepsilon(E) &= \int_0^{T_\varepsilon} f(\mathcal{E}_\varepsilon(t, E), \Theta_\varepsilon(t, E)) dt = \int_0^{T_0} f(\mathcal{E}_0(t, E), \Theta_0(t, E)) dt + \mathcal{O}(\varepsilon) \\ &= -2E \int_0^{2\pi} \sin^2 t (2E \cos^2 t - 1) dt + \mathcal{O}(\varepsilon) , \end{aligned} \tag{A.7}$$

e la (A.6) è immediata (il calcolo dell'integrale serve solo a determinare esattamente a e b ; che si tratti di costanti positive è ovvio).

Per $\varepsilon = 0$, la curva F_ε si riduce alla parabola $F_0(E) = -a E(E - b)$, che ha uno zero trasverso in $E_0^* = b$. Per ε piccolo allora esiste sicuramente uno zero trasverso in un punto

³⁷⁾ Si sta qui usando la sola regolarità della soluzione al variare dei parametri dell'equazione (per t limitato).

vicino $E_\varepsilon^* = E_0^* + \mathcal{O}(\varepsilon)$, e tale zero è unico in un intervallo³⁸⁾ I contenente E_ε^* . Si ha così $\Phi_\varepsilon(E_\varepsilon^*) = E_\varepsilon^*$, e corrispondentemente il sistema (A.4) ha un'orbita periodica con dato iniziale $(E_\varepsilon^*, 0)$.

Si dimostra poi facilmente che tale orbita è attrattiva: precisamente, dalla forma di F_ε si vede subito che per ogni fissato $E^0 \in I$, $E^0 \neq E_\varepsilon^*$, posto $E^k = \Phi_\varepsilon(E^{k-1})$, la successione E^0, E^1, E^2, \dots è monotona crescente se $E^0 < E_\varepsilon^*$, e decrescente se $E^0 > E_\varepsilon^*$; ma il limite della successione è necessariamente un punto fisso (per questo basta la continuità di Φ al variare del dato iniziale), e dunque coincide con E_ε^* .

L'andamento qualitativo di F_ε e Φ_ε , assieme alla costruzione grafica degli E_k , è rappresentato in figura A.2. ■

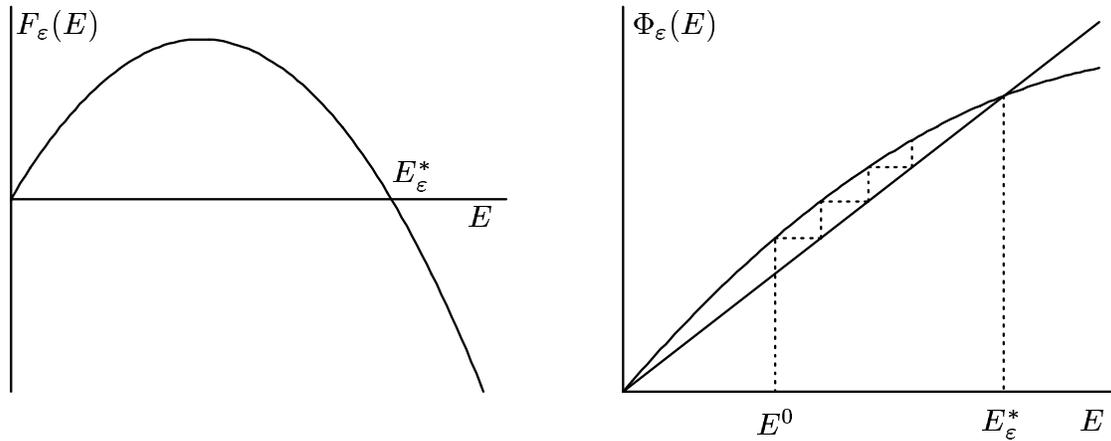


Figura A.2. L'andamento di $F_\varepsilon(E)$ e di $\Phi_\varepsilon(E)$.

Abbandoniamo ora lo studio perturbativo, e dimostriamo del tutto in generale che

Proposizione A.2. *Per ogni $\varepsilon > 0$ esiste unico nel piano (x, v) un ciclo limite stabile, il cui bacino di attrazione coincide con l'intero piano privato dell'origine.*

Dimostrazione. Convieni introdurre al posto di v una diversa coordinata³⁹⁾

$$y = \varepsilon^{-1}v + f(x), \quad f(x) = \frac{x^3}{3} - x; \tag{A.8}$$

il sistema che sostituisce (A.2) è allora

$$\dot{x} = \varepsilon(y - f(x)), \quad \dot{y} = -\varepsilon^{-1}x. \tag{A.9}$$

³⁸⁾ Anzi: in qualunque intervallo staccato dall'origine, se ε è abbastanza piccolo.

³⁹⁾ La coppia di coordinate x, y al posto di x, v è spontanea, quando si scriva l'equazione di Van der Pol (A.1) nella forma

$$\frac{d}{dt} [\dot{x} + \varepsilon(x^3/3 - x)] + x = 0.$$

Si vede bene che l'asse y e la cubica γ di equazione $y = f(x)$ dividono il piano in quattro regioni I – IV simili a quadranti, ove \dot{x} e \dot{y} hanno segno definito, nell'ordine $(+, -)$, $(+, +)$, $(-, +)$, $(-, -)$; si veda la figura A.3a. Ciò è già sufficiente a garantire che ciascuna traiettoria con origine nel semiasse y positivo, che denotiamo y^+ , taglia successivamente γ^+ , y^- , γ^- (con ovvio significato dei simboli), poi ancora y^+ , e così via. Ha senso allora considerare l'applicazione $\sigma : y^+ \rightarrow y^+$, corrispondente al primo ritorno della traiettoria su y^+ (sezione di Poincaré).

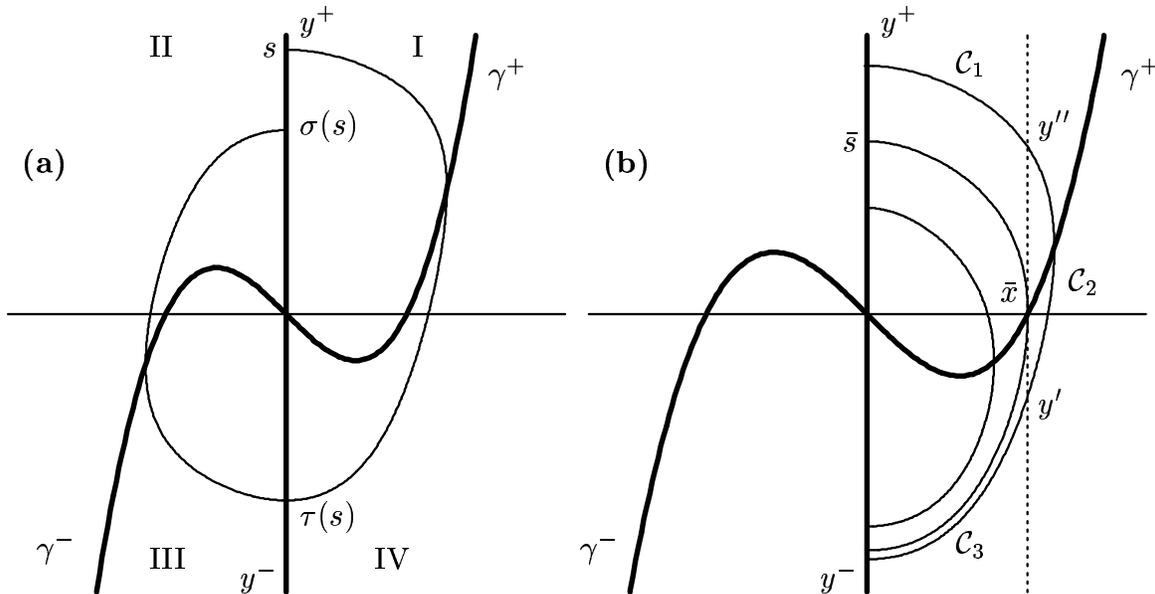


Figura A.3. (a) I quadranti e le mappe σ e τ . (b) Illustrazione del Lemma.

Se usiamo una variabile s per l'asse y^+ , possiamo dire che σ è funzione continua di s , e monotona crescente (le orbite non si possono intersecare); perciò vi sono, per ogni $s > 0$, tre sole possibilità:

- i) $\sigma(s) = s$, e l'orbita è periodica;
- ii) $\sigma(s) > s$, e allora $s < \sigma(s) < \sigma^2(s) < \dots$ (l'orbita procede a spirale verso l'esterno);
- iii) $\sigma(s) < s$, e allora $s > \sigma(s) > \sigma^2(s) > \dots$ (l'orbita procede a spirale verso l'interno).

Accanto alla mappa σ , corrispondente a un giro completo dell'orbita, è utile considerare una mappa τ corrispondente a mezzo giro: precisamente, se l'orbita di origine $(0, s) \in y^+$ taglia per la prima volta y^- in $(0, -s')$, $s' > 0$, poniamo $\tau(s) = s'$. La mappa τ è, come σ , continua e monotona crescente, e inoltre la simmetria delle equazioni mostra che σ si ottiene componendo due volte τ , $\sigma(s) = \tau(\tau(s))$. Le tre possibilità sopra considerate corrispondono allora rispettivamente a $\tau(s) = s$, $\tau(s) > s$, e $\tau(s) < s$.

Poniamo ora $\delta(s) = \frac{1}{2}(\tau(s)^2 - s^2)$, e dimostriamo il seguente

Lemma A.3. *Esiste un valore \bar{s} , tale che: (a) $\delta(s) > 0$ per $0 < s \leq \bar{s}$; (b) $\delta(s)$ è strettamente decrescente per $s \geq \bar{s}$; (c) $\delta(s)$ tende a $-\infty$ per $s \rightarrow +\infty$.*

Dimostrazione. Per ogni $s > 0$, sia t_s il tempo al quale la traiettoria di origine $(0, s)$ taglia per la prima volta l'asse y^- ; introduciamo la funzione $W(x, y) = \frac{1}{2}(\varepsilon^{-2}x^2 + y^2)$, e

osserviamo che risulta (banalmente)

$$\delta(s) = W(0, Y_s(t_s)) - W(0, s) = \int_0^{t_s} \dot{W}(X_s(t), Y_s(t)) dt, \quad (\text{A.10})$$

ove $(X_s(t), Y_s(t))$ indica la traiettoria di origine $(0, s)$. Dalle equazioni del moto (A.9) si ottiene subito $\dot{W} = \varepsilon^{-2}x\dot{x} + y\dot{y} = -\varepsilon^{-1}xf(x)$, perciò

$$\delta(s) = -\varepsilon^{-1} \int_0^{t_s} X_s(t)f(X_s(t)) dt.$$

Sia ora $\bar{s} > 0$ il punto di y^+ tale che l'orbita uscente da $(0, \bar{s})$ taglia l'asse x nel suo punto di intersezione $(\bar{x}, 0)$ con γ^+ (figura A.3b); per $f(x) = \frac{1}{3}x^3 - x$ risulta evidentemente $\bar{x} = \sqrt{3}$. E' chiaro che, se $s \leq \bar{s}$, allora si ha $f(x_s(t)) \leq 0$ per $0 \leq t \leq t_s$. In tali condizioni l'integrale in (A.10) risulta positivo, e questo basta per il punto (a).

Per il punto (b), la stima dell'integrale è poco più complicata: prendiamo una traiettoria uscente da $(0, s)$ con $s > \bar{s}$; dividiamola in tre curve \mathcal{C}_1 , \mathcal{C}_2 e \mathcal{C}_3 , a seconda che sia $X_s(t) < \bar{x}$, $X_s(t) \geq \bar{x}$, e poi ancora $X_s(t) < \bar{x}$ (figura A.3b), e scriviamo

$$\delta(s) = \int_{\mathcal{C}_1} \dot{W} dt + \int_{\mathcal{C}_2} \dot{W} dt + \int_{\mathcal{C}_3} \dot{W} dt.$$

Per il primo integrale è possibile sostituire x a t come variabile indipendente; usando $dt = dx/\dot{x} = (y - f(x))^{-1}$ si trova subito

$$\int_{\mathcal{C}_1} \dot{W}(X_s(t), Y_s(t)) dt = -\varepsilon^{-2} \int_0^{x_1} \frac{x f(x)}{\mathcal{Y}_s(x) - f(x)} dx,$$

ove si è indicata con $\mathcal{Y}_s(x)$ la curva corrispondente alla traiettoria $(X_s(t), Y_s(t))$. E' evidente da questa espressione che il contributo di \mathcal{C}_1 è positivo e decrescente al crescere di s (si osservi che $\mathcal{Y}_s(x)$ per ogni x è crescente con s). In modo del tutto analogo si vede che anche il contributo di \mathcal{C}_3 è positivo (attenzione al verso di percorrenza di \mathcal{C}_3), e decrescente con s . Per il secondo integrale possiamo invece usare y come variabile indipendente, e scrivere

$$\int_{\mathcal{C}_2} \dot{W} dt = - \int_{y'}^{y''} f(\mathcal{X}_s(y)) dy,$$

ove $\mathcal{X}_s(y)$ è la curva corrispondente alla traiettoria in esame, e gli estremi di integrazione sono ordinati in modo che sia $y'_s < y''_s$ (cioè in ordine opposto al verso di percorrenza; ciò spiega il segno meno). Si vede allora che il contributo di \mathcal{C}_2 è negativo, e crescente in modulo per s crescente: infatti, a fissato y , $\mathcal{X}_s(y)$ cresce con s (la traiettoria è più esterna), perciò anche f cresce, e inoltre il dominio di integrazione in y si allarga. Ciò basta per concludere che $\delta(s)$, per $s \geq \bar{s}$, è decrescente (strettamente), come asserito in (b).

Infine, poiché al crescere di s la traiettoria si sposta verso destra a piacere (per tutti i punti di γ^+ , con ascissa x comunque grande, passa una traiettoria che proviene da y^+ con s sufficientemente grande, e interseca l'asse delle ascisse a destra di x) è anche chiaro che $\delta(s) \rightarrow -\infty$ per $s \rightarrow +\infty$. ■

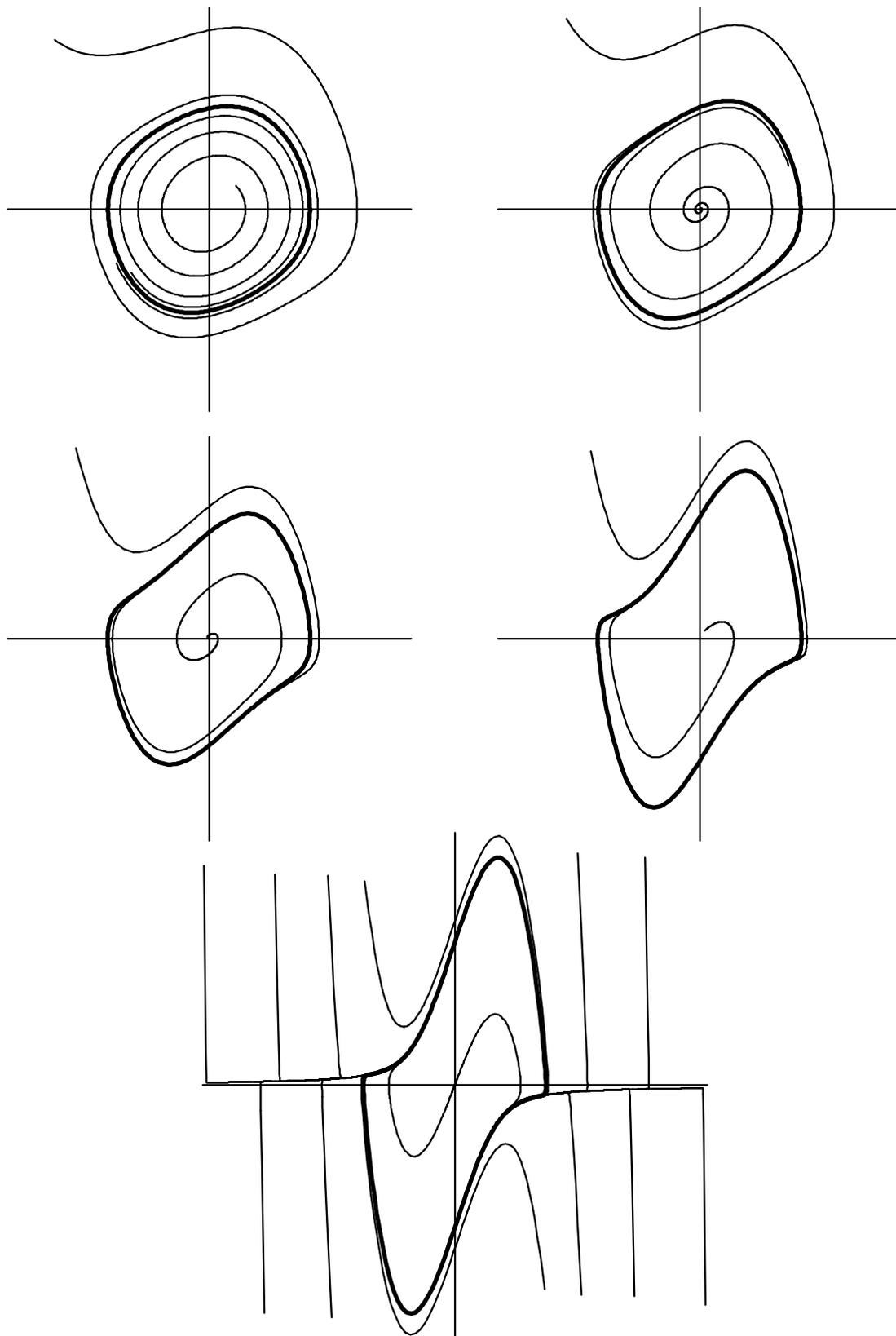


Figura A.4. Il ritratto in fase del sistema (A.2) per diversi valori di ϵ , nell'ordine: $\epsilon = 0.2, 0.4, 0.8, 1.6, 3$.

Dal lemma si deduce all'istante che esiste un unico zero s^* di δ in \mathbb{R}^+ , e dunque un'unica orbita periodica. Che poi tale orbita sia globalmente attrattiva è del tutto evidente: ad esempio per $0 < s < s^*$ si ha $\delta(s) > 0$, e dunque la successione $\sigma^k(s)$ è crescente; Il suo limite è necessariamente invariante (come segue dalla continuità di δ), e dunque coincide con s^* . In modo analogo si ragiona per $s > s^*$. ■

- **Esercizio.** L'espressione precisa di $f(x)$, si veda la (A.8), non è mai stata usata nella dimostrazione, che si basa interamente sulle proprietà qualitative della curva γ . Quali proprietà sono importanti?

La figura A.4 illustra il comportamento del sistema (A.2) per diversi valori di ε .

Una realizzazione elettrica

La figura A.5 mostra un circuito elettrico con in parallelo un'induttanza, un condensatore e un componente non lineare, che si suppone obbedire a una legge del tipo

$$I = f(V)/R , \tag{A.11}$$

avendo indicato con $V = V_b - V_a$ la differenza di potenziale ai suoi capi, con I la corrente che lo attraversa (positiva nel verso della freccia), e con R una costante; per $f(V) = V$ si ha eccezionalmente un componente lineare, precisamente un resistore di resistenza R .

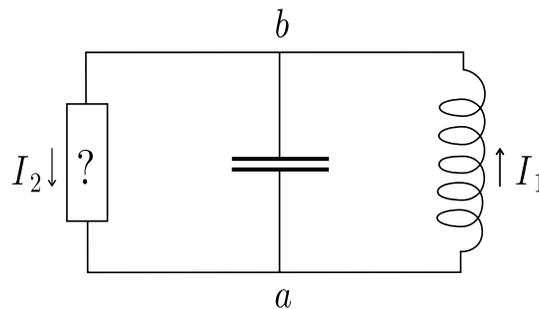


Figura A.5. Un circuito elettrico con un componente non lineare

Dalla (A.11) assieme alle consuete equazioni dei componenti lineari ($\dot{I}_1 = -V/L$ per l'induttanza, $\dot{V} = I_1 - I_2$ per il condensatore) si ricava senza difficoltà per V l'equazione

$$\ddot{V} + \mu f'(V)\dot{V} + \omega^2 V = 0 , \tag{A.12}$$

ove si è posto $\mu = 1/(RC)$, $\omega^2 = 1/(LC)$, e si è denotata con f' la derivata prima di f . Nel caso lineare $f(V) = V$ si ha un oscillatore armonico smorzato (con smorzamento, si osservi, inversamente proporzionale a R), e corrispondentemente per $t \rightarrow \infty$ si ha il comportamento asintotico banale $V(t) \rightarrow 0$. Lo stesso comportamento si trova con evidenza per ogni f monotona crescente (basta usare l'energia $\frac{1}{2}\dot{V}^2 + \frac{\omega^2}{2}V^2$ come funzione di Lyapunov). Per avere il fenomeno del ciclo limite occorre una f sostanzialmente diversa, in qualche modo simile alla cubica (A.8) che entra nell'equazione di Van der Pol: precisamente serve una f "a doppia gobba", con pendenza negativa per $|V|$ piccolo e positiva per $|V|$ grande (un

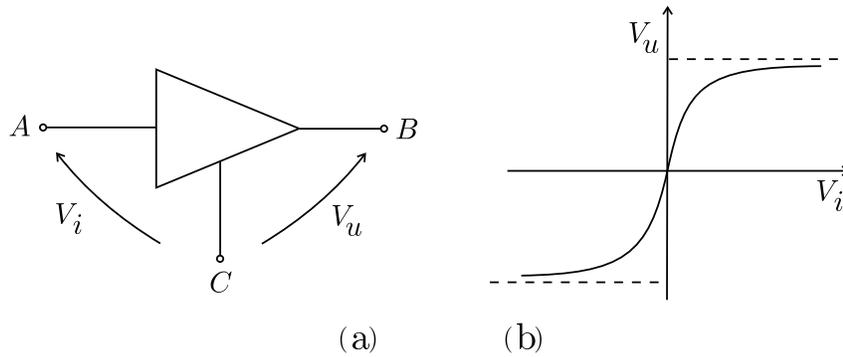


Figura A.6. (a) L'amplificatore ideale; (b) l'andamento di V_u come funzione di V_i .

attimo di riflessione mostra che solo queste proprietà di f sono importanti; la cubica (A.8) è solo il più semplice polinomio fatto a questo modo).

Il tratto a pendenza negativa attorno all'origine comporta il fatto curioso che la corrente scorre in direzione opposta alla differenza di potenziale. Ma in questo modo si immette energia nel circuito (cosa evidentemente necessaria per $|V|$ piccolo, se vogliamo un ciclo limite), e dunque il componente non lineare deve contenere un amplificatore. Con semplificazione drastica, intendiamo qui per amplificatore un elemento di circuito con tre connessioni esterne — un "ingresso" A , una "uscita" B , e una connessione comune C , si veda la figura A.6a — tale che, se indichiamo con $V_i = V_A - V_C$ e $V_u = V_B - V_C$ le tensioni in ingresso e in uscita,

- 1) V_u è funzione della sola V_i , $V_u = F(V_i)$, monotona crescente e della forma qualitativa indicata in figura A.6b; la pendenza $F'(V)$ si suppone maggiore di uno per per $|V|$ piccolo, mentre per $|V|$ grande F "satura" a due valori opposti $\pm V_{\max}$.
- 2) La corrente in ingresso è trascurabile.

Il numero $\alpha = F'(0) > 1$ è detto guadagno dell'amplificatore per piccoli segnali. L'alimentazione, pensata interna al dispositivo, non è rappresentata.

Con questo componente ideale costruiamo il componente non lineare a due connessioni esterne, da inserire nel circuito, nel modo indicato in figura A.7a; il circuito completo è disegnato in modo equivalente in figura A.7b. Poiché $V = V_i$, si ha $V_u = F(V)$, e dunque $I_2 = (V - F(V))/R$. Per confronto con la (A.11) si vede allora che la funzione f da inserire nella (A.12) è

$$f(V) = V - F(V) ,$$

e il suo grafico ha la forma voluta. In particolare si ha, come richiesto, $f'(0) = 1 - \alpha < 0$. La resistenza negativa che in questo modo si è ottenuta si spiega intuitivamente così: un piccolo valore ad esempio positivo di $V = V_i$ produce un valore V_u più grande di V , che fa scorrere una corrente contraria alla differenza di potenziale nel ramo di circuito ove il componente non lineare è inserito.

Una minima riflessione mostra che il parametro significativo che misura il termine non lineare nell'equazione (A.12) non è $\mu = \frac{1}{RC}$, ma $\tilde{\mu} = -f'(0)\mu = \frac{\alpha-1}{RC}$, e corrispondentemente il parametro adimensionale che caratterizza la (A.12) è

$$\beta = \frac{\alpha - 1}{R} \sqrt{\frac{L}{C}} .$$

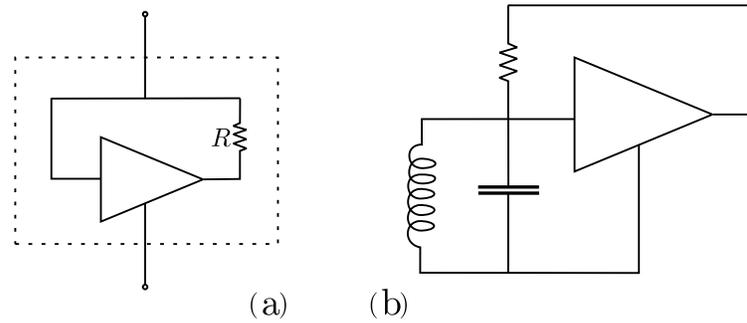


Figura A.7. (a) La costruzione del componente non lineare; (b) il circuito completo.

Ci si aspetta così un ciclo limite prossimo a una circonferenza, con oscillazioni di V quasi sinusoidali, per $\alpha - 1 \ll R\sqrt{L/C}$ (l'amplificatore lavora nella zona quasi lineare), e invece una forte distorsione per $\alpha - 1 \gg R\sqrt{L/C}$ (l'amplificatore lavora in saturazione).

B. Dimensione frattale

Studiando il comportamento asintotico dei sistemi dissipativi ci siamo imbattuti in sottoinsiemi di M , come l'insieme iperbolico Λ del ferro di cavallo o l'attrattore X del solenoide, che sono molto importanti per caratterizzare il comportamento del sistema, e tuttavia appaiono topologicamente mal fatti o "strani". Una delle proprietà evidenti di questi insiemi è la auto-similarità:⁴⁰⁾ ingrandendo una parte della figura si riottiene la figura intera. Questa proprietà caratterizza una classe interessante di insiemi "strani", cui di recente⁴¹⁾ (Mandelbrot, anni '70) è stato dato il nome di *frattali* (da *fractus*, spezzato). Esempi elementari di frattali sono l'insieme ternario di Cantor, la curva di Koch — una curva continua ma non differenziabile in nessun punto — la cui costruzione è illustrata in figura B.1, e il "setaccio di Sierpinski", la cui costruzione è illustrata in figura B.2. Le figure limite sono caratterizzate da una esatta auto-similarità, illustrata in figura B.3.

Agli insiemi frattali si può assegnare una dimensione, generalizzazione della consueta nozione di dimensione di una varietà, che però generalmente risulta non intera. La nozione di varietà differenziabile, ricordiamo, è quella di un insieme che localmente assomiglia a un aperto di \mathbb{R}^m , per un qualche m . Globalmente la varietà può essere complicata, ma osservandone una porzione abbastanza piccola essa appare liscia e si confonde col suo spazio

⁴⁰⁾ Negli esempi elementari che abbiamo visto l'auto-similarità è esatta; in altri casi è solo asintotica.

⁴¹⁾ Lo studio matematico di questi oggetti risale agli studi di Hausdorff, Julia e Fatou (1920 circa). Da Mandelbrot in poi si è riconosciuta l'importanza cruciale degli oggetti frattali come modelli per diverse situazioni fisiche (dall'accrescimento dei fiocchi di neve alla tortuosità dei polimeri al frastagliamento delle coste alla struttura degli alveoli polmonari), ed è ripreso in pieno il loro studio sia in ambito fisico sia anche, più di recente e più timidamente, in ambito matematico.

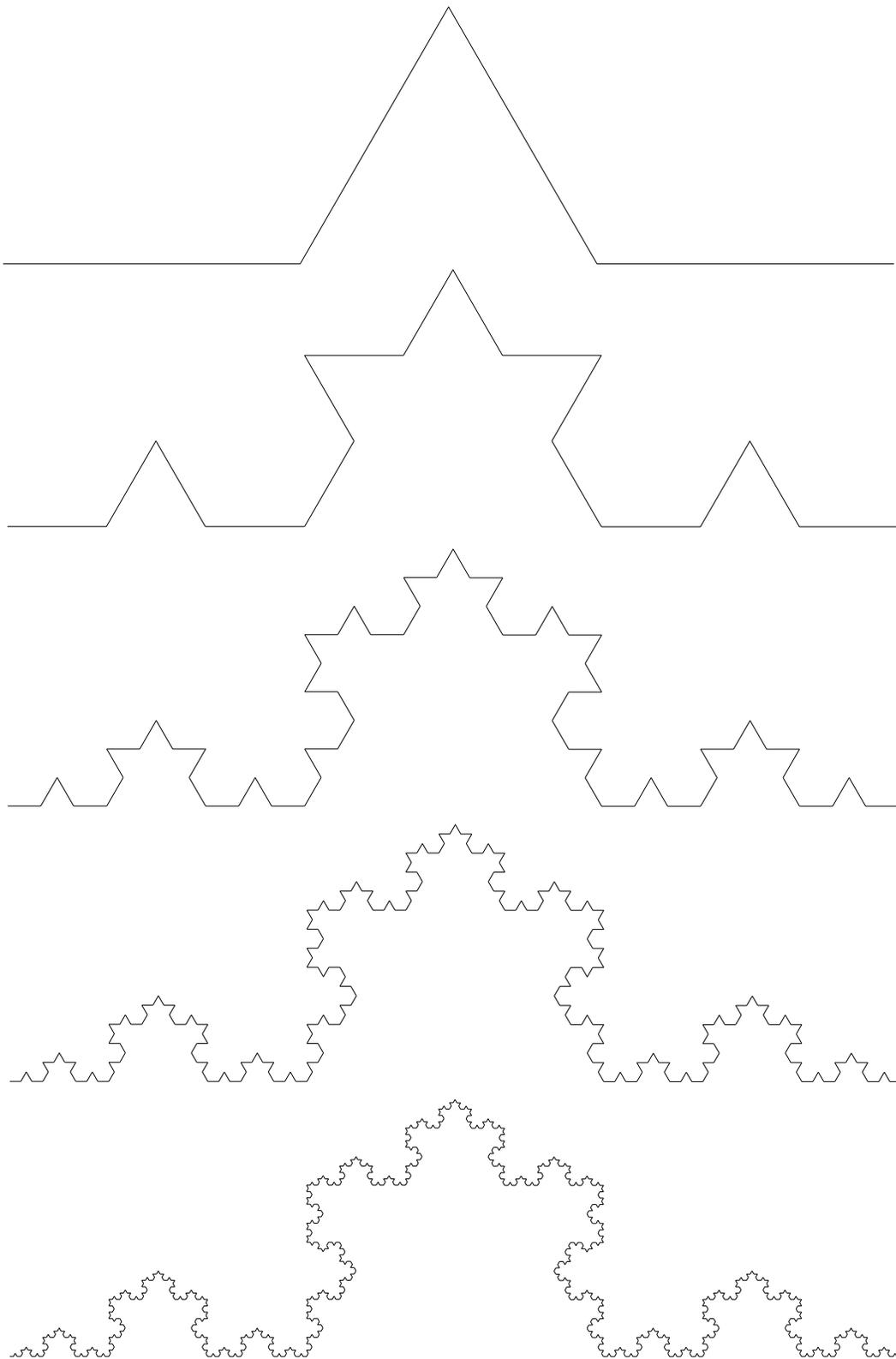


Figura B.1. La costruzione della curva di von Koch.

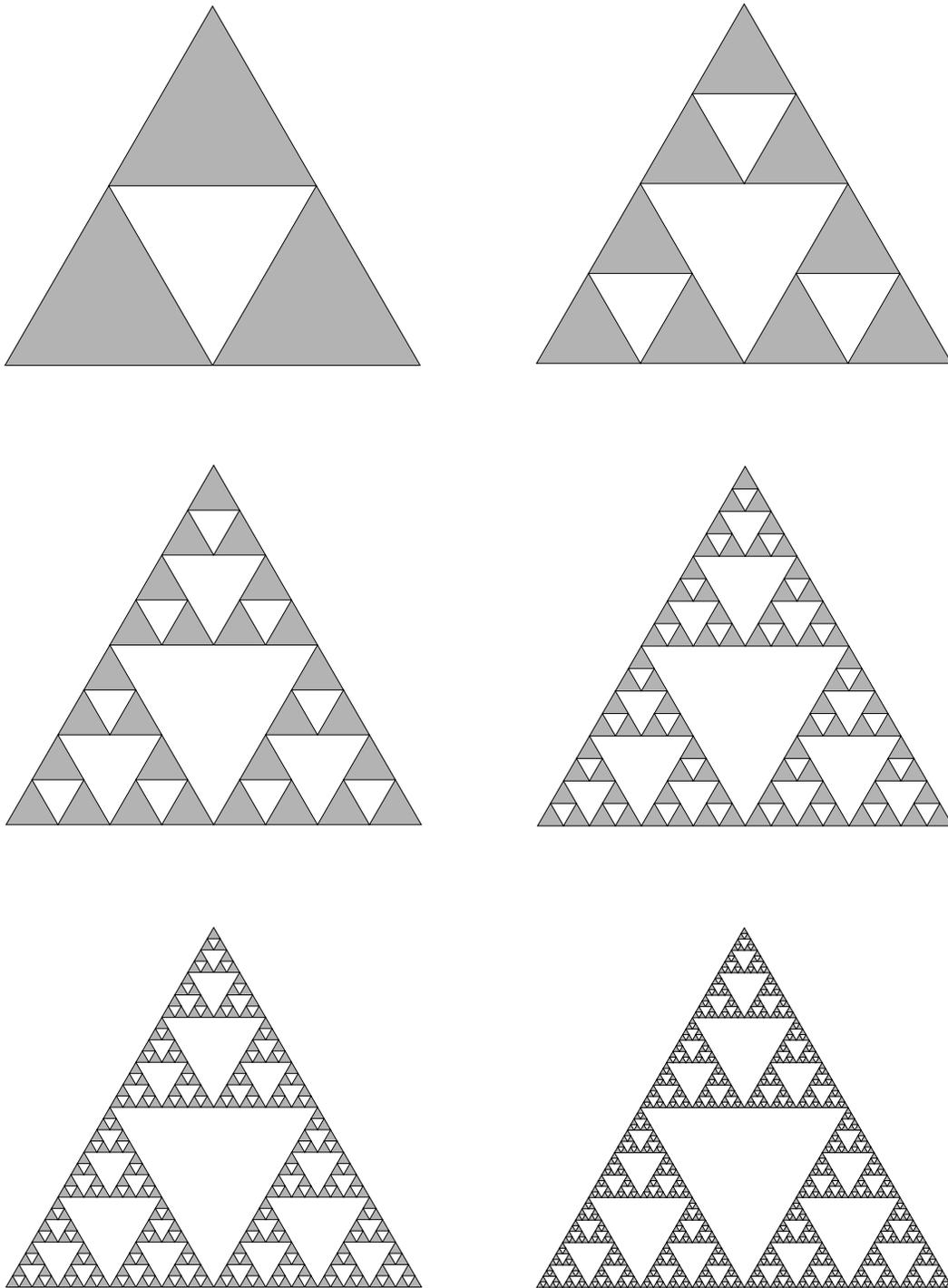


Figura B.2. La costruzione del setaccio di Sierpinski.

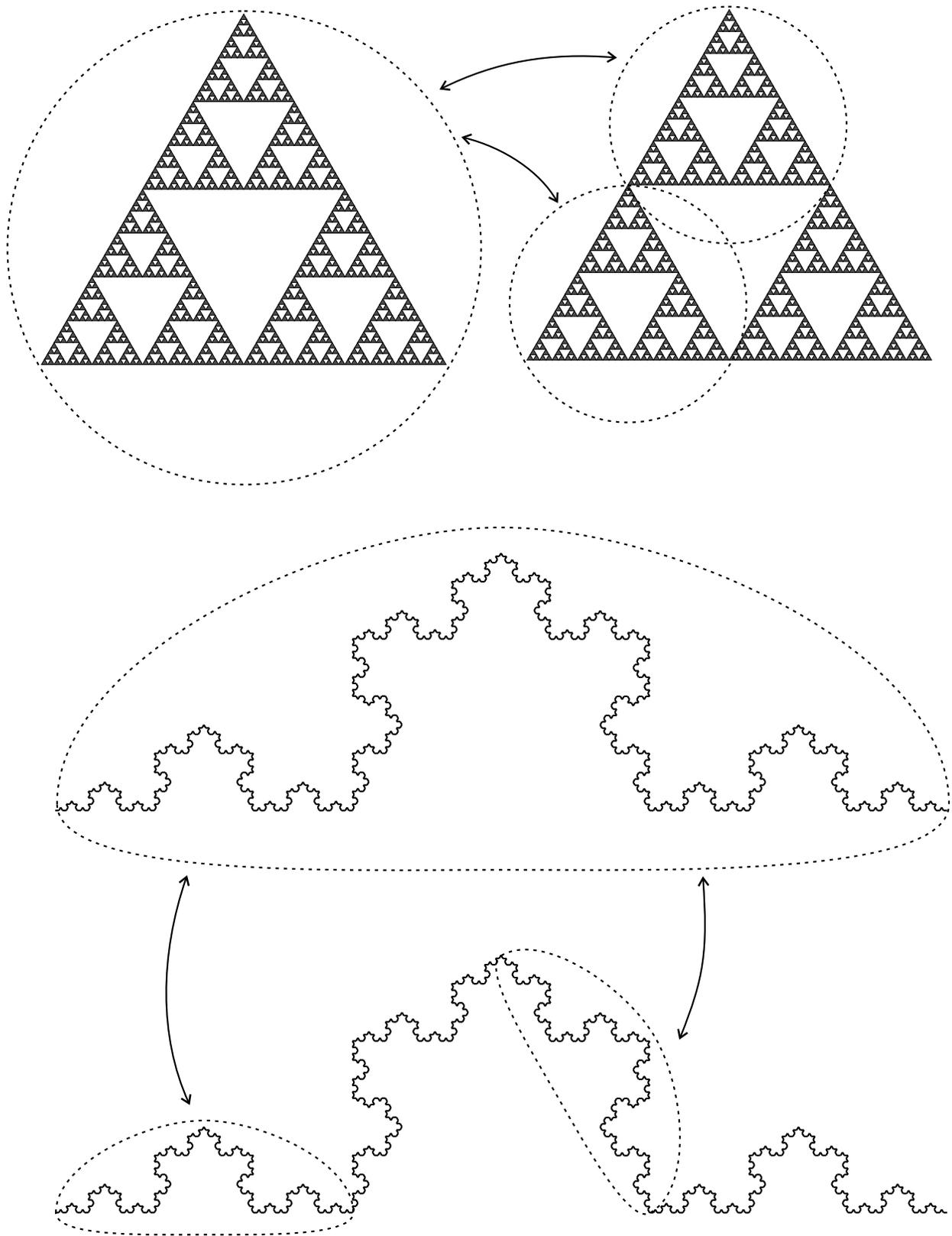


Figura B.3. L'esatta auto-similarità della curva di von Koch e del setaccio di Sierpinski

tangente. Non è così per gli oggetti strani cui abbiamo accennato sopra: comunque se ne ingrandisca una parte, si scopre nuovo dettaglio, e a nessuna scala si riesce ad approssimare questi oggetti con uno spazio euclideo. L'idea di Hausdorff per assegnare a questi insiemi una dimensione è quella di far leva sulla misura: precisamente di generalizzare l'idea elementare che, per una varietà m -dimensionale, una palla di diametro (piccolo) ρ ha misura che cresce come ρ^m .

B.1. La dimensione di Hausdorff d_H

Sia $F \subset \mathbb{R}^n$ limitato, $R = \{B_1, B_2, \dots\}$ un suo ricoprimento numerabile in sfere (aperte o chiuse: non avrà importanza) di diametro ρ_1, ρ_2, \dots , e sia \mathcal{R}_ρ la famiglia dei ricoprimenti di F tali che $\rho_i \leq \rho$ per ogni i . Per $s \in \mathbb{R}$ poniamo

$$m^s(R) = \sum_i \rho_i^s, \quad M_\rho^s(F) = \inf_{R \in \mathcal{R}_\rho} m^s(R). \quad (\text{B.1})$$

M_ρ^s è funzione monotona non crescente di ρ , perché al crescere di ρ si allarga l'insieme su cui si prende l'inferiore, perciò esiste il limite (eventualmente infinito)

$$M^s(F) = \lim_{\rho \rightarrow 0} M_\rho^s(F).$$

Per $\rho < 1$, M_ρ^s è funzione non crescente di s , e dunque lo è anche M^s ; più precisamente

Lemma B.1. Per $\rho < 1$ e $t > s$ si ha

$$M_\rho^t(F) \leq \rho^{t-s} M_\rho^s(F).$$

Dimostrazione. Banalmente, si ha

$$M_\rho^t(F) = \inf \sum_i \rho_i^{t-s} \rho_i^s \leq \rho^{t-s} \inf \sum_i \rho_i^s = \rho^{t-s} M_\rho^s(F).$$

■

Se ne deduce immediatamente che

Proposizione B.2. Esiste $d_H \geq 0$ tale che

$$M^s(F) = \infty \quad \text{per } s < d_H, \quad M^s(F) = 0 \quad \text{per } s > d_H.$$

Dimostrazione. Dal lemma segue subito che c'è al più un valore di s per cui $0 < M^s < \infty$. Non può essere sempre $M^s = 0$, perché si vede subito che $M^s = \infty$ per $s < 0$; non può essere sempre $M^s = \infty$, perché si vede facilmente che $M^s = 0$ per $s > n$. ■

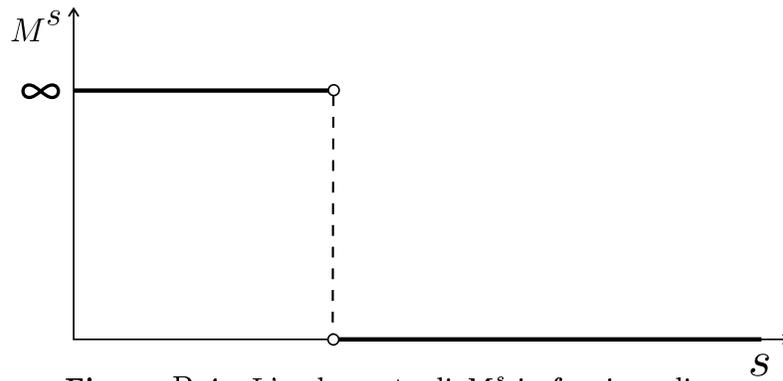


Figura B.4. L'andamento di M^s in funzione di s .

M^s , come funzione di s , ha dunque l'andamento indicato in figura B.4; per $s = d_H$, M^s può essere infinito, finito o nullo.

Definizione. Il numero d_H si dice *dimensione di Hausdorff* dell'insieme F ; il numero M^{d_H} si dice *misura di Hausdorff* di F .

In effetti si può dimostrare che M^{d_H} è una misura. Si verifica senza difficoltà (esercizio) che

Proposizione B.3. Il valore di d_H non cambia se si considerano, nella definizione di M^s_ρ , ricoprimenti con rettangoli o insiemi di altra forma prefissata, o anche con insiemi qualsiasi, di diametro non superiore a ρ .

Cambia però M^{d_H} ; per fissare le idee ci riferiremo nel seguito alle sfere. E' ancora un facile esercizio verificare che d_H riproduce le dimensioni consuete degli oggetti classici, in particolare:

- i. per un insieme finito si ha $d_H = 0$, e M^0 coincide con la cardinalità dell'insieme;
- ii. per un insieme numerabile si ha $d_H = 0$, $M^0 = \infty$;
- iii. per un segmento di lunghezza L si ha $d_H = 1$, $M^1 = L$; più in generale per un rettangolo m -dimensionale si ha $d_H = m$, e $M^m = c_m \text{Vol}^m(F)$, ove c_m è il volume della sfera m -dimensionale di diametro uno.

Proposizione B.4. Per l'insieme di Cantor ternario si trova

$$d_H = \frac{\log 2}{\log 3} .$$

Dimostrazione. La dimostrazione è facile anche se un po' laboriosa, e si basa sulla verifica che per $s = \frac{\log 2}{\log 3}$ si ha $0 < M^s < \infty$. La stima da sopra è immediata: per $\rho = 3^{-k}$, $k \geq 1$, prendiamo R "omogeneo", cioè costituito da 2^k palle (chiuse) di ugual diametro ρ ; segue subito

$$m^s(R) = 2^k 3^{-ks} = 1 ,$$

e dunque $M^s_\rho \leq 1$. Mandando ρ a 0 lungo la successione 3^{-k} si trova evidentemente $M^s \leq 1$. Per la stima da sotto, prendiamo un qualsiasi ricoprimento $R \in \mathcal{R}_\rho$ in palle

aperte, che ai fini del calcolo dell'inferiore possiamo pensare finito, e mostriamo che per $s = \frac{\log 2}{\log 3}$ esistono $k^* \in \mathbb{N}$ e un ricoprimento omogeneo $R^* \in \mathcal{R}_\rho$, composto da 2^{k^*} palle di ugual raggio 3^{-k^*} , tale che

$$m^s(R) \geq \frac{1}{2} m^s(R^*) = \frac{1}{2}. \quad (\text{B.2})$$

Prendendo allora l'inferiore su $R \in \mathcal{R}_\rho$ si trova $M_\rho^s \geq \frac{1}{2}$, e per $\rho \rightarrow 0$ segue la stima da sotto $M^s \geq \frac{1}{2}$.

Per ottenere la (B.2) ragioniamo così: per ogni palla B_i di R , detto ρ_i il suo diametro, definiamo k_i in modo che sia $3^{-k_i-1} \leq \rho_i < 3^{-k_i}$, e poi poniamo $k^*(R) = \max_i k_i$. Necessariamente la palla B_i ha intersezione non vuota con al più uno dei segmenti di lunghezza 3^{-k_i} non vuoti che intervengono nella costruzione dell'insieme di Cantor allo stadio k_i (la distanza minima tra questi segmenti è infatti 3^{-k_i}). Si ottiene allora sicuramente un ricoprimento $R' \in \mathcal{R}_\rho$ traslando ciascuna B_i al centro del segmento in questione, alzandone diametro a 3^{-k_i} e chiudendo la palla. Ciascun termine della sommatoria (B.1) per m^s è cresciuto al più di un fattore 3^s , pertanto

$$m^s(R) \geq 3^{-s} m^s(R').$$

Infine otteniamo il ricoprimento R^* sostituendo ciascuna palla di R' con $2^{k^*-k_i}$ palle di diametro 3^{-k^*} , disposte in modo da seguire la struttura dell'insieme di Cantor fino allo stadio di costruzione k^* , ed eliminando eventuali palle coincidenti. Per la scelta particolare $s = \frac{\log 2}{\log 3}$, $3^{-s} = 2^{-1}$, risulta $m^s(R') \geq \frac{1}{2} m^s(R^*)$, come richiesto. (La disuguaglianza da sotto non è ottimale; con qualche considerazione in più si riuscirebbe a ottenere $M^{d_H} = 1$). ■

In modo analogo si dimostra che la curva di Koch ha $d_H = \frac{\log 4}{\log 3}$, il setaccio di Sierpinski ha $d_H = \frac{\log 3}{\log 2}$. Per quanto riguarda l'attrattore X del solenoide, con considerazioni simili si calcola $d_H = \frac{\log 2}{\log 4} = \frac{1}{2}$ per la sezione (figura 4.2), $d_H = \frac{3}{2}$ per l'attrattore completo.

E' spontaneo chiedersi quali requisiti debba avere una quantità d per meritare il nome di dimensione. Un elenco ragionevole è il seguente:

- a. monotonicità: se $E \subset F$, allora $d(E) \leq d(F)$;
- b. stabilità finita: $d(E \cup F) \leq \max(d(E), d(F))$; più in generale,
- ḃ. stabilità numerabile: $d(\cup_{i \in \mathbb{Z}} F_i) \leq \sup_{i \in \mathbb{Z}} d(F_i)$;
- c. invarianza per traslazione, rotazione, dilatazione di F in \mathbb{R}^n , o più in generale, invarianza per trasformazioni regolari invertibili $\mathbb{R}^n \rightarrow \mathbb{R}^n$;
- d. compatibilità coi valori classici: $d(F) = 0$ per un insieme finito di punti (per un insieme numerabile, se vale ḃ), $d(F) = n$ per un aperto di \mathbb{R}^n , $d(F) = m$ per una sottovarietà m -dimensionale di \mathbb{R}^n .

Non è difficile dimostrare (esercizio) che la dimensione di Hausdorff d_H soddisfa effettivamente tutte queste proprietà. Ha anche altre proprietà che ci si aspetta da una dimensione: ad esempio, "genericamente" si ha, come per le sottovarietà di \mathbb{R}^n ,

$$d_H(E \cap F) = \max(0, d_H(E) + d_H(F) - n).$$

In questo però non ci addentriamo.

B.2. Varianti

E' di uso comune, specialmente tra i fisici, la cosiddetta *dimensione di capacità* (o semplicemente capacità) d_c , detta anche *dimensione di conteggio* (inglese *box counting*). E' un'idea un po' più rozza di quella di dimensione di Hausdorff, ma più pratica nell'uso concreto, in particolare negli studi sperimentali o numerici. La definizione è semplicissima: denotata con $N(\rho)$ è la cardinalità minima di un ricoprimento di F con sfere di ugual diametro ρ , si pone (se il limite esiste)

$$d_c(F) = \lim_{\rho \rightarrow 0} \frac{\log N(\rho)}{\log \rho^{-1}}.$$

Anche qui è indifferente prendere sfere o insiemi di altra forma (però fissata). Si riconosce facilmente che d_c è il valore che si ottiene seguendo la procedura con cui è stata definita d_H , ma restringendo \mathcal{R}_ρ ai soli ricoprimenti in palle di ugual diametro ρ ; se ne deduce subito $d_c(F) \geq d_H(F)$. Una maniera indipendente di definire d_c , che si dimostra essere equivalente, è la seguente: si considera un cubo di \mathbb{R}^n di lato L contenente F , e per ogni $k \geq 1$ si considera la decomposizione in cubetti di lato (ad esempio) $\rho = L/k$; si pone $n_k =$ numero di cubetti contenenti almeno un punto di F , e poi

$$d_c = \lim_{k \rightarrow \infty} \frac{\log n_k}{\log k}$$

(di qui la terminologia "*box counting*"). Questa variante della definizione fornisce un metodo assai pratico di calcolo empirico: basta contare le celle piene, riportare n_k in funzione di k in scala logaritmica, e valutare una pendenza asintotica.

La dimensione di capacità d_c esiste e coincide con d_H in molti casi interessanti, in particolare negli esempi modello sopra considerati. Quanto alle proprietà, si dimostra che d_c soddisfa tutti i requisiti sopra elencati, tranne la stabilità numerabile \tilde{b} . Ad esempio (esercizio) l'insieme $\{1/j, j \in \mathbb{N}\}$ ha dimensione di capacità $\frac{1}{2}$, anziché zero. Si vede anche che F e la sua chiusura hanno la stessa dimensione di capacità, il che è assai spiacevole: l'insieme dei razionali di un intervallo ha dimensione uno (superiore alla dimensione dell'insieme di Cantor, che pure è non numerabile). Nella letteratura soprattutto fisica si trovano ulteriori varianti alla nozione di dimensione frattale, nelle quali qui non entriamo.

La letteratura sui frattali è enorme. Segnaliamo:

- K. Falconer, *Fractal geometry*, Wiley 1990. Un raro esempio di libro sui frattali chiaro matematicamente e non difficile.
- B. Mandelbrot, *Oggetti frattali*, Einaudi, Torino 1987. Un testo classico, ove l'autore espone la sua tesi principale (che ha avuto molto seguito tra i fisici) secondo la quale i frattali forniscono modelli adatti a molti e diversi fenomeni fisici. Matematicamente è un po' povero.
- H.-O. Peitgen e P.H. Richter, *La bellezza dei frattali*, Bollati Boringhieri, Torino 1987. Interessante e con immagini splendide.
- E. Guyon e H.E. Stanley, *Fractal forms*, Elsevier 1991. Una bella collezione di fotografie di oggetti naturali, per cui è spontaneo cercare modelli nella geometria frattale.