

Numero 2 - 1991

RATIO MATHEMATICA

Atti del Convegno di
Matematica Applicata all'Economia e all'Ingegneria

Pescara 29, 27, 28 Gennaio 1989



a cura di

Franco Eugeni e Mario Gionfriddo

Comitato Scientifico

Ilio Adorisio, *Roma*

Albrecht Beutelspacher, *Giessen*

Franco Eugeni, *Pescara*

Mario Gionfriddo, *Catania*

Giovanni Melzi, *Milano*

Bruno Rizzi, *Napoli*

Aniello Russo Spena, *L'Aquila*

Romano Scozzafava, *Roma*

Numero 2 - 1991

RATIO MATHEMATICA

**Atti del Convegno di
Matematica Applicata all'Economia e all'Ingegneria**

Pescara 29, 27, 28 Gennaio 1989

a cura di

Franco Eugeni e Mario Gionfriddo

Comitato Scientifico

Ilio Adorisio, Roma

Giovanni Melzi, Milano

Albrecht Beutelispacher, Giessen

Bruno Rizzi, Napoli

Franco Eugeni, Pescara

Aniello Russo Spena, L'Aquila

Mario Gionfriddo, Catania

Romano Scozzafava, Roma

Direttore Responsabile: Prof. FRANCO EUGENI

Autorizzazione n. 9/90 del 10-07-1990

Fascicolo realizzato con il contributo del Dipartimento di Scienze, e Storia dell'Architettura e Restauro. Hanno collaborato alla realizzazione di questo fascicolo: il Prof. Antonio Maturo e gli ingegneri G. Di Biase, D. Marconi e G. Mataloni.

Fotocomposizione e stampa:



microPRINT - Lungomare Matteotti, 111 - Pescara

Indice del fascicolo

Un problema tridimensionale di trasporto fuori di una sfera riflettente. (L. Bassi e P. Gamba)	pag.	1
Applicazioni della matematica in problemi di comunicazione. (L. Beradi)	pag.	13
Il problema della protezione dell'informazione II. La gestione e lo sviluppo dei sistemi informativi: sicurezza contro costo? (A. Bonisoli e F. Eugeni)	pag.	29
I metodi evolutivi nella dinamica delle strutture: una rivista critica. Parte I - Fondamenti teorici. (F. Brancaleoni, L. Cavini e C. Gavarini)	pag.	57
Analisi della bontà di alcuni generatori di numeri pseudocausali per la cifratura dei messaggi e la simulazione. (N. Cera e A. Maturo)	pag.	83
Confronti fra inferenza bayesiana e test classici per l'analisi di successioni di numeri pseudocausali. (G. Di Biase e A. Maturo)	pag.	97
Un modello neokeynesiano di equilibrio. (M. Gallegati e L. Gardini)	pag.	111
Sulle coalizioni in teoria dei giochi. (S. Innamorati)	pag.	139
On irreducible blocking sets in projective planes. (S. Innamorati e A. Maturo)	pag.	151
Sistemi dinamici digitali e loro possibili applicazioni. (G. Melzi)	pag.	157
Una suggestiva analogia fra fenomeni biologici e fenomeni economici a soglia. (F. Mercanti)	pag.	161
Un problema di ottimizzazione non-smooth in statistica robusta. (L. Stefanini)	pag.	167

UN PROBLEMA TRIDIMENSIONALE DI TRASPORTO FUORI DI UNA SFERA RIFLETTENTE

L. Bassi, P. Gamba (*)

SOMMARIO. In questo lavoro studiamo l'equazione lineare monoenergetica di Boltzmann nello spazio esterno ad una sfera di raggio $a > 0$ con condizioni al contorno di riflessione speculare. Servendoci della teoria dei semigrupp di operatori lineari positivi in spazi di Banach dimostriamo che l'operatore di streaming è il generatore infinitesimale di un semigrupp fortemente continuo di contrazioni positive, di cui diamo la forma esplicita, e che l'operatore di Boltzmann è il generatore infinitesimale di semigrupp fortemente continuo e irriducibile. Inoltre studiamo in maniera completa lo spettro dell'operatore di streaming.

ABSTRACT. In this paper we study the linear monoenergetic Boltzmann equation in the space external to a sphere of radius $a > 0$ with specular reflection boundary conditions. By means of theory of linear positive operator semigroups in Banach spaces we show that streaming operator is the infinitesimal generator of a strongly continuous semigroup of positive contractions, of which we give the explicit form, and that Boltzmann operator is the infinitesimal generator of an irreducible strongly continuous semigroup. Moreover we study completely the spectrum of streaming operator.

1. INTRODUZIONE

L'equazione lineare del trasporto di Boltzmann fu usata all'inizio per lo studio del trasporto di energia radiante in atmosfere stellari o planetarie ([2], [9]) ed in seguito per descrivere il trasporto di neutroni e raggi gamma in sistemi nucleari ([1], [4], [5]): nel primo caso la funzione incognita rappresenta l'intensità di radiazione, nel secondo la distribuzione di particelle.

(*) Dipartimento di Matematica "Vito Volterra", Università degli Studi di Ancona.

Uno studio dettagliato dell'equazione lineare del trasporto è stato fatto da Voigt [13] ed alcuni dei più importanti risultati di questo studio sono stati usati nel nostro lavoro. Per l'equazione lineare del trasporto è importante considerare sia problemi interni, derivanti dalla teoria dei reattori nucleari, sia quelli esterni, derivanti dal trasporto di radiazione.

In questo lavoro ci serviamo della teoria dei semigrupp di operatori lineari in spazi di Banach L^p ($1 \leq p < \infty$).

Tale teoria si è sviluppata abbastanza rapidamente a partire dal teorema di generazione di Hille e Yosida del 1948 e da allora si è estesa con importanti applicazioni a molti campi dell'analisi grazie soprattutto a Ralph Phillips [7]. Tuttavia solo in tempi recenti la teoria ha avuto un'adeguata sistemazione essenzialmente basata sulle tecniche sviluppate per gli operatori positivi su spazi di Banach e combinando l'analisi funzionale con lo studio di problemi di Cauchy collegati in particolare con problemi concreti di evoluzione in teoria del trasporto, biomatematica e fisica ([3], [10], [11], [12]).

Sia X uno spazio di Banach ed A un operatore lineare da $\mathcal{D}(A) \subset X$ in X ; dato $x \in X$ il problema astratto di Cauchy per A con dato iniziale x consiste nel trovare una soluzione $f(t)$ del problema

$$\begin{cases} \frac{d f(t)}{d t} = A f(t), & t > 0 \\ f(0) = x \end{cases} \quad (*)$$

dove per soluzione si intende una funzione $f(t)$ a valori in X tale che $f(t)$ è continua per $t \geq 0$, derivabile con continuità, con $f(t) \in \mathcal{D}(A)$ per $t > 0$ e soddisfacente il problema (*). Se A è il generatore infinitesimale di un semigrupp fortemente continuo di operatori $T(t)$, il problema astratto di Cauchy per A ha soluzione unica $f(t) = T(t)x$, per ogni $x \in \mathcal{D}(A)$.

In questo lavoro, nel paragrafo 2 introduciamo le notazioni usate nel seguito e definiamo l'operatore di streaming T e l'operatore di Boltzmann B .

Nel paragrafo 3 studiamo l'operatore T , mostrando che esso è il generatore infinitesimale di un semigrupp fortemente continuo (teorema 1) di cui diamo la forma e stabiliamo le proprietà spettrali di T (teorema 2).

Nel paragrafo 4 esaminiamo l'operatore B , che, per il noto teorema di perturbazione limitata, risulta generatore di un semigrupp fortemente continuo di cui diamo (teorema 3) una importante proprietà mostrando che esso è irriducibile.

2. FORMULAZIONE DEL PROBLEMA

Consideriamo l'equazione di Boltzmann lineare monoenergetica nello

spazio D esterno ad una sfera di raggio $a > 0$, assumendo condizioni al contorno di riflessione speculare.

Supponiamo che lo spazio S delle velocità sia la sfera unitaria di $\mathbb{R}^3$ e che la sezione d'urto sia costante.

Pertanto esamineremo l'equazione

$$\frac{\partial f}{\partial t}(\vec{r}, \vec{\Omega}, t) = -\vec{\Omega} \cdot \nabla_{\vec{r}} f(\vec{r}, \vec{\Omega}, t) - f(\vec{r}, \vec{\Omega}, t) + \int_S k(\vec{r}, \vec{\Omega}', \vec{\Omega}) f(\vec{r}, \vec{\Omega}', t) d\vec{\Omega}', \quad (1)$$

insieme alle condizioni iniziali ed al contorno.

In (1) $f(\vec{r}, \vec{\Omega}, t)$ può essere considerata o come intensità specifica di radiazione o come funzione di distribuzione delle particelle nello spazio delle fasi $D \times S$ al tempo t ; $k: D \times S \times S \rightarrow \mathbb{R}$ rappresenta l'albedo per singolo scattering.

Introducendo su D la restrizione della misura di Lebesgue μ e su S la misura di superficie ω , supponiamo che k sia una funzione misurabile strettamente positiva.

Al fine di scrivere il nostro problema in forma astratta, sono necessarie alcune notazioni ed alcune definizioni.

Indichiamo con η la misura su $D \times S$, cioè $\eta = (\mu \times \omega)$. Denotiamo con $C_c^{1,0}(D \times S)$ l'insieme di tutte le funzioni complesse su $D \times S$ che sono continue, derivabili con continuità rispetto alle variabili spaziali e a supporto compatto; infine sia $X = L^p(D \times S)$ lo spazio di Banach complesso dotato della norma usuale.

Si definiscono gli operatori T_0 e K come segue:

$$\left\{ \begin{array}{l} \mathcal{D}(T_0) = \{f \in X : \exists g \in X \text{ tale che } \forall \varphi \in C_c^{1,0}(D \times S), \int [(-\vec{\Omega} \cdot \nabla_{\vec{r}} - 1) \varphi] f d\eta \\ \quad = - \int \varphi g d\eta\} \\ T_0 f = g \end{array} \right.$$

$$\left\{ \begin{array}{l} \mathcal{D}(K) = X \quad \mathfrak{R}(K) \subset X \\ (Kf) = \int_S k(\vec{r}, \vec{\Omega}', \vec{\Omega}) f(\vec{r}, \vec{\Omega}') d\vec{\Omega}'. \end{array} \right.$$

Osserviamo che se $f \in C_c^{1,0}$, allora è $T_0 f = -\vec{\Omega} \cdot \nabla_{\vec{r}} f - f$.

Detta σ la misura di superficie sulla frontiera ∂D di D , consideriamo il prodotto $|\vec{\Omega} \cdot \vec{n}(\vec{r})| \sigma \times \omega$ come misura su $\Sigma_{\pm} = \{(\vec{r}, \vec{\Omega}) \in \partial D \times S, |\vec{r}| = a; \pm \vec{\Omega} \cdot \vec{n}(\vec{r}) > 0\}$, essendo $\vec{n}(\vec{r})$ il versore normale in $\vec{r} \in \partial D$ entrante nello spazio D .

Possiamo così scrivere

$$\|f\|_{\pm} = \left\{ \int_{\Sigma_{\pm}} |f(\vec{r}, \vec{\Omega})|^p |\vec{\Omega} \cdot \vec{n}(\vec{r})| d(\sigma \times \omega) \right\}^{1/p}.$$

Si introducano ora le funzioni $t_{\pm} : D \times S \rightarrow [0, +\infty)$ definite come segue:

$$t_{\pm}(\vec{r}, \vec{\Omega}) = \inf \{t > 0 : \vec{r} \pm t \vec{\Omega} \notin D\}$$

che nel seguito saranno indicate semplicemente con $t_{\pm}$ qualora non sorgano ambiguità nel simbolo.

Se $f \in \mathcal{D}(T_0)$, esiste una funzione $\hat{f} \in \mathcal{D}(T_0)$ che soddisfa $\hat{f} = f$ quasi dappertutto (q.d.) e tale che la funzione

$$[-t_-, t_+] \ni t \rightarrow \hat{f}(\vec{r} + t \vec{\Omega}, \vec{\Omega}) \in \mathbb{R}$$

sia assolutamente continua per quasi ogni (q.o.) $(\vec{r}, \vec{\Omega}) \in D \times S$.

Se $-t_- \leq t' < t'' \leq t_+$, allora è

$$\int_{t'}^{t''} (-\vec{\Omega} \cdot \vec{\nabla}_{\vec{r}}) \hat{f}(\vec{r} + t \vec{\Omega}, \vec{\Omega}) dt = \hat{f}(\vec{r} + t'' \vec{\Omega}, \vec{\Omega}) - \hat{f}(\vec{r} + t' \vec{\Omega}, \vec{\Omega})$$

per q.o. $(\vec{r}, \vec{\Omega}) \in D \times S$.

Possiamo porre

$$(B_{\pm} f)(\vec{r}, \vec{\Omega}) = \lim_{t \downarrow 0} \hat{f}(\vec{r} \pm t \vec{\Omega}, \vec{\Omega}), \quad (\vec{r}, \vec{\Omega}) \in \Sigma_{\pm}$$

in quanto il limite esiste q.d. su $\Sigma_{\pm}$ e non dipende dalla funzione $\hat{f}$; le funzioni $B_{\pm} f$ sono misurabili [13].

Le condizioni al contorno sono definite dall'operatore J come segue

$$\begin{cases} \mathcal{D}(J) = L^p(\Sigma_-) & \mathfrak{R}(J) \subset L^p(\Sigma_+) \\ (Jf)(\vec{r}, \vec{\Omega}) = f(\tau(\vec{r}, \vec{\Omega})) & (\vec{r}, \vec{\Omega}) \in \Sigma_+ \end{cases}$$

dove è $\tau(\vec{r}, \vec{\Omega}) := [\vec{r}, \vec{\Omega} - 2[\vec{\Omega} \cdot \vec{n}(\vec{r})]\vec{n}(\vec{r})]$.

L'operatore J descrive la condizione al contorno di riflessione perfetta di tipo speculare; poiché l'applicazione $\tau : \Sigma_+ \rightarrow \Sigma_-$ è bigettiva e conserva la misura, J è un operatore positivo con $\|J\| = 1$.

Possiamo ora definire l'operatore T come segue:

$$\begin{cases} \mathcal{D}(T) = \{f \in \mathcal{D}(T_0) : B_{\pm} f \in L^p(\Sigma_{\pm}); B_+ f = JB_- f\} & \mathfrak{R}(T) \subset X \\ Tf := T_0 f \end{cases}$$

Allora stabiliamo il problema di Cauchy in forma astratta

$$df/dt = Bf \quad f(0) = f_0$$

dove l'operatore B è l'operatore di Boltzmann

$$\begin{cases} \mathcal{D}(B) = \mathcal{D}(T) \quad \mathfrak{R}(B) \subset X \\ (Bf)(\vec{r}, \vec{\Omega}) = [(T+K)f](\vec{r}, \vec{\Omega}). \end{cases}$$

Definiamo infine alcuni spazi ed alcune funzioni che saranno utili nel seguito [13]:

$$V^{\pm} = \{(\vec{r}, \vec{\Omega}) \in D_{\pm}S : t_{\pm} < +\infty; (\vec{r}_{\pm} t_{\pm} \vec{\Omega}) \cdot \vec{\Omega} \neq 0\}$$

$$V_0^{\pm} = \{(\vec{r}, \vec{\Omega}) \in D_{\pm}S : t_{\pm} < +\infty; (\vec{r}_{\pm} t_{\pm} \vec{\Omega}) \cdot \vec{\Omega} = 0\}$$

$$V_{\infty}^{\pm} = \{(\vec{r}, \vec{\Omega}) \in D_{\pm}S : t_{\pm} = +\infty\}.$$

E' facile riconoscere che: $D_{\pm}S = V^+ \cup V_0^+ \cup V_{\infty}^+ = V^- \cup V_0^- \cup V_{\infty}^-$; $V^{\pm}$ sono insiemi aperti in $D_{\pm}S$.

Inoltre osserviamo che $t_{\pm}$ sono funzioni continue su $V^{\pm}$.

Al fine di parametrizzare $V^{\pm}$, definiamo gli insiemi $\tilde{V}^{\pm}$ aperti in $\partial D_{\pm}S \times \mathbb{R}^+$ e le applicazioni $p^{\pm}$ come segue

$$\tilde{V}^{\pm} = \{(\vec{r}, \vec{\Omega}, t) \in \Sigma_{\pm} \times (0, +\infty) : 0 < t < t_{\pm} = +\infty\},$$

$$p^{\pm}: \tilde{V}^{\pm} \rightarrow V^{\pm}$$

$$p^{\pm}(\vec{r}, \vec{\Omega}, t) := (\vec{r}_{\pm} t \vec{\Omega}, \vec{\Omega}).$$

Se $q^{\pm}: V^{\pm} \rightarrow \tilde{V}^{\pm}$, date da $q^{\pm}(\vec{r}, \vec{\Omega}) = (\vec{r}_{\pm} t_{\pm} \vec{\Omega}, \vec{\Omega}, t_{\pm})$, sono le inverse delle $p^{\pm}$, osserviamo che $p^{\pm}$ e $q^{\pm}$ sono applicazioni bigettive continue.

Useremo su $\tilde{V}^{\pm}$ le misure $\tilde{\eta}_{\pm} = (|\vec{\Omega} \cdot \vec{n}(\vec{r})| \sigma(\omega) \times \mu_1)$ notando che le applicazioni $p^{\pm}$ e $q^{\pm}$ conservano la misura rispettivamente $\tilde{\eta}_{\pm}$ e η .

3. L'OPERATORE DI STREAMING T.

Per quanto concerne l'operatore di streaming T si dimostra il seguente

TEOREMA 1. *L'operatore T è il generatore infinitesimale del semigruppoo fortemente continuo $\{\mathcal{U}(t); t \geq 0\}$ dato da*

$$\mathcal{U}(t)h = \begin{cases} e^{-t} h(\vec{r} - t\vec{\Omega}, \vec{\Omega}) & \text{se } (\vec{r}, \vec{\Omega}) \in V_{\infty}^{+} \cup V_0^{+} \text{ o } (\vec{r}, \vec{\Omega}) \in V^{+} \text{ con } t_{-} > t \\ e^{-t} h[p_{-}(\tau(\vec{r} - t_{-}\vec{\Omega}, \vec{\Omega}), t - t_{-})] & \text{se } (\vec{r}, \vec{\Omega}) \in V^{+} \text{ con } t_{-} < t \end{cases} \quad (3)$$

che è semigruppoo di contrazioni positive.

Dim. Grazie al teorema 5.11 della monografia di Voigt ([13], p. 99), sappiamo che l'operatore $(T+I)$ è il generatore infinitesimale di un semigruppoo fortemente continuo di operatori lineari e limitati, la cui norma è minore o uguale al $\max\{1, \|I\|\}$. Poiché nel nostro caso è $\|I\| = 1$, $(T+I)$ è generatore infinitesimale di un semigruppoo di contrazioni. Da ciò e da un noto teorema di perturbazione ([5], p. 280), segue che T stesso genera un C_0 -semigruppoo $\{\mathcal{U}(t); t \geq 0\}$ che $\|\mathcal{U}(t)\| \leq e^{-t} \forall t \geq 0$. Per riconoscere la validità della (3), osserviamo che ogni numero complesso λ , con $\operatorname{Re} \lambda > -1$, appartiene all'insieme risolvente di T . Pertanto per tali λ e per ogni $h \in X$, il problema

$$(\lambda - T)f = h \quad (4)$$

ha un'unica soluzione $f \in \mathcal{D}(T)$, data da

$$f = R(\lambda; T)h = \int_0^{+\infty} e^{-\lambda t} \mathcal{U}(t)h \, dt \quad (5)$$

Ma possiamo anche risolvere il problema (4) usando lo stesso procedimento seguito da Voigt ([13], p. 100). Notiamo che l'operatore T^0 definito da

$$\begin{cases} \mathcal{D}(T^0) = \{f \in \mathcal{D}(T_0) : B_{+}f = 0\} \\ T^0f = T_0f \end{cases}$$

è il generatore infinitesimale del semigruppoo di contrazioni positive $\{\mathcal{U}_0(t); t \geq 0\}$:

$$\mathcal{U}_0(t)h = \begin{cases} e^{-t} h(\vec{r} - t\vec{\Omega}, \vec{\Omega}) & \text{per } (\vec{r}, \vec{\Omega}) \in V_{\infty}^{+} \cup V_0^{+} \\ & \text{o } (\vec{r}, \vec{\Omega}) \in V^{+} \text{ e } t_{-} > t \\ 0 & \text{altrimenti.} \end{cases}$$

Per ogni $\Phi \in L^p(\Sigma_{+})$ poniamo

$$L(\lambda; \Phi) = \begin{cases} e^{-(\lambda+1)t} \Phi(\vec{r}-t\vec{\Omega}, \vec{\Omega}) & \text{per } (\vec{r}, \vec{\Omega}) \in V^+ \\ 0 & \text{altrimenti.} \end{cases}$$

Come dimostrato da Voigt ([13], p. 100) è facile vedere che il problema (4) ha come soluzione la funzione

$$f = R(\lambda : T^0) h + L(\lambda; JB_R(\lambda : T^0) h) \quad \text{cioè}$$

$$f(\vec{r}, \vec{\Omega}) = \begin{cases} \int_0^t e^{-(\lambda+1)t} h(\vec{r}-t\vec{\Omega}, \vec{\Omega}) dt + \alpha \int_{t_-}^{+\infty} e^{-(\lambda+1)t} h(p^+(\tau(\vec{r}-t\vec{\Omega}, \vec{\Omega}), t-t_-)) dt, & \text{per } (\vec{r}, \vec{\Omega}) \in V^+ \\ \int_0^{+\infty} e^{-(\lambda+1)t} h(\vec{r}-t\vec{\Omega}, \vec{\Omega}) dt, & \text{per } (\vec{r}, \vec{\Omega}) \in V_{\infty}^+ \end{cases}$$

Confrontando questo risultato con la (5), deduciamo l'espressione di $\mathcal{U}(t)$.

Da questa espressione possiamo anche vedere che gli operatori $\mathcal{U}(t)$ ($t \geq 0$) sono tutti positivi.

Per completare lo studio dell'operatore T stabiliamo le sue proprietà nel seguente

TEOREMA 2. *Il piano spettrale di T è suddiviso nel seguente modo:*

- (i) $\rho(T) = \{\lambda \in \mathbb{C} : \operatorname{Re} \lambda \neq -1\}$
- (ii) $\sigma_p(T) = \emptyset$
- (iii) $\sigma_r(T) = \emptyset$
- (iv) $\sigma_c(T) = \{\lambda \in \mathbb{C} : \operatorname{Re} \lambda = -1\}$.

Dim.

- (i) Per il teorema 1, $\rho(T) \supset \{\lambda \in \mathbb{C} : \operatorname{Re} \lambda > -1\}$. Come nel teorema 1, possiamo facilmente vedere che l'operatore $-T$ è il generatore infinitesimale del C_0 -semigruppato $\{\mathcal{V}(t) : t \geq 0\}$ definito come segue:

$$\mathcal{V}(t)h = \begin{cases} e^t h(\vec{r} + t\vec{\Omega}, \vec{\Omega}) & \text{se } (\vec{r}, \vec{\Omega}) \in V_{\infty}^- \cup V_0^- \quad \text{o} \quad (\vec{r}, \vec{\Omega}) \in V^- \text{ con } t_+ > |t| \\ e^t h(p^+(\tau(\vec{r} + t_+ \vec{\Omega}, \vec{\Omega}), t - t_+)) & \text{se } (\vec{r}, \vec{\Omega}) \in V \text{ con } t_+ < |t| \end{cases}$$

Possiamo facilmente vedere che $\|\mathcal{V}(t)\| \leq e^t$ e quindi $\rho(-T) \supset \{\lambda \in \mathbb{C} : \operatorname{Re} \lambda > 1\}$ e pertanto $\rho(T) \supset \{\lambda \in \mathbb{C} : \operatorname{Re} \lambda < -1\}$.

Per $0 < \delta < 1/2$, siano u_δ le funzioni

$$u_\delta(\vec{r}, \vec{\Omega}) = \begin{cases} \frac{1}{\delta} e^{-\gamma(\vec{r}, \vec{\Omega})} e^{-i b \gamma(\vec{r}, \vec{\Omega})/\nu(\vec{r}, \vec{\Omega})} & \text{se } (\vec{r}, \vec{\Omega}) \in V^+ \text{ con } \delta^2 \leq \nu \leq \delta \\ 0 & \text{altrimenti} \end{cases}$$

$$\text{dove } \gamma(\vec{r}, \vec{\Omega}) := \frac{1}{a} \vec{r} \cdot (\vec{r} - t_- \vec{\Omega}) - a, \quad \nu(\vec{r}, \vec{\Omega}) := \frac{1}{a} \vec{\Omega} \cdot (\vec{r} - t_- \vec{\Omega}).$$

Con semplici calcoli, mostriamo che, se $\lambda = -1 + ib$, le funzioni u_δ soddisfano le seguenti condizioni:

$$a) \quad \|u_\delta\|_p^p \geq \frac{\pi^2 a^2 2^{p+1}}{p} > 0$$

$$b) \quad \|(\lambda - T)u_\delta\|_p^p = \|\nu u_\delta\|_p^p =$$

$$= \frac{8\pi^2 a^2 (1 - \delta^{p+1})\delta}{p(p+1)} \longrightarrow 0 \quad \text{per } \delta \rightarrow 0.$$

Pertanto $\sigma(T) = \{\lambda \in \mathbb{C} : \operatorname{Re} \lambda = -1\}$ e $\rho(T) = \{\lambda \in \mathbb{C} : \operatorname{Re} \lambda \neq -1\}$ [8].

(ii) Si dimostra che per ogni $\lambda \in \mathbb{C}$, l'unica soluzione dell'equazione $(\lambda - T)f = 0$ è la funzione nulla q.d.. Poiché la funzione $t \rightarrow f(\vec{r} - t\vec{\Omega}, \vec{\Omega})$ è assolutamente continua per q.o. $(\vec{r}, \vec{\Omega}) \in D \times S$, essa è q.d. derivabile, e si ha

$$(-\vec{\Omega} \cdot \nabla_{\vec{r}} f)(\vec{r} - t\vec{\Omega}, \vec{\Omega}) = \frac{\partial}{\partial t} f(\vec{r} - t\vec{\Omega}, \vec{\Omega})$$

e l'equazione $(\lambda - T)f = 0$ diventa

$$\frac{\partial}{\partial t} f(\vec{r} - t\vec{\Omega}, \vec{\Omega}) = (\lambda + 1) f(\vec{r} - t\vec{\Omega}, \vec{\Omega})$$

la cui soluzione ha la seguente forma:

$$f(\vec{r}, \vec{\Omega}) = \begin{cases} e^{-(\lambda+1)t_-} f(\vec{r} - t_- \vec{\Omega}, \vec{\Omega}) & \text{per } (\vec{r}, \vec{\Omega}) \in V^+ \\ e^{(\lambda+1)t_+} f(\vec{r} + t_+ \vec{\Omega}, \vec{\Omega}) & \text{per } (\vec{r}, \vec{\Omega}) \in V^- \\ e^{-(\lambda+1)t} f(\vec{r} - t\vec{\Omega}, \vec{\Omega}) & \text{per } (\vec{r}, \vec{\Omega}) \in V_\infty^+ \cap V_\infty^- \end{cases}$$

Sia $\pi_{\vec{\Omega}}$ il piano equatoriale perpendicolare ad $\vec{\Omega}$ per ogni $\vec{\Omega} \in S$, allora

$$V_{\infty}^{+} \cap V_{\infty}^{-} = \left\{ (\vec{r}, \vec{\Omega}) ; \vec{r} = \vec{r}' + t\vec{\Omega}, \vec{r}' \in \pi_{\vec{\Omega}}, |\vec{r}'| \geq a, t \in \mathbb{R} \right\}$$

e pertanto

$$\begin{aligned} \|f\|_p^p &= \int_{V_{\infty}^{+} \cap V_{\infty}^{-}} |e^{-(\lambda+1)t} f(\vec{r}-t\vec{\Omega}, \vec{\Omega})|^p d\eta + \\ &+ \int_{V^{+}} |e^{-(\lambda+1)t} f(\vec{r}-t\vec{\Omega}, \vec{\Omega})|^p d\eta + \int_{V^{-}} |e^{(\lambda+1)t} f(\vec{r}+t\vec{\Omega}, \vec{\Omega})|^p d\eta = \\ &= \int_S d\omega \int_{\pi_{\vec{\Omega}} \cap \{\vec{r}': |\vec{r}'| \geq a\}} d(\mu|\pi_{\vec{\Omega}}) \int_{-\infty}^{+\infty} e^{-tpRe(\lambda+1)} |f(\vec{r}', \vec{\Omega})|^p dt + \\ &+ \int_0^{+\infty} e^{-tpRe(\lambda+1)} \int_{\Sigma_+} |B_+ f|^p |\vec{\Omega} \cdot \vec{n}(\vec{r})| d(\sigma \times \omega) dt + \\ &+ \int_0^{+\infty} e^{tpRe(\lambda+1)} \int_{\Sigma_-} |B_- f|^p |\vec{\Omega} \cdot \vec{n}(\vec{r})| d(\sigma \times \omega) dt. \end{aligned}$$

Il primo integrale al secondo membro converge se e solo se $f(\vec{r}, \vec{\Omega})=0$ per q.o. $(\vec{r}, \vec{\Omega}) \in V_{\infty}^{+} \cap V_{\infty}^{-}$. Per quanto riguarda gli altri integrali, vediamo che

$$\begin{aligned} &\int_0^{+\infty} e^{-tpRe(\lambda+1)} \int_{\Sigma_+} |B_+ f|^p |\vec{\Omega} \cdot \vec{n}(\vec{r})| d(\sigma \times \omega) dt + \\ &+ \int_0^{+\infty} e^{tpRe(\lambda+1)} \int_{\Sigma_-} |B_- f|^p |\vec{\Omega} \cdot \vec{n}(\vec{r})| d(\sigma \times \omega) dt = \\ &= \int_0^{+\infty} e^{-tpRe(\lambda+1)} \|B_+ f\|^p dt + \int_0^{+\infty} e^{tpRe(\lambda+1)} \|B_- f\|^p dt = \end{aligned}$$

$$= \int_0^{+\infty} \left(e^{-t p R_*(\lambda+1)} + e^{t p R_*(\lambda+1)} \right) \| B_- f \| dt$$

che converge se e solo se $B_- f = 0$ per q.o. $(r, \Omega) \in \Sigma_-$. Ricordando che $\|B_+ f\| = \|B_- f\|$, anche $B_+ f$ è nulla q.d. su Σ_+ . Pertanto $(\lambda - T)f = 0$ se e solo se $f = 0$ quasi dappertutto. Resta così dimostrato che $\sigma_p = \emptyset$.

(iii) L'operatore $T: L^p(D \times S) \rightarrow L^p(D \times S)$ è un operatore lineare chiuso definito denso e perciò $\sigma_p(T) = \sigma_p(T')$ ([4], prop. 2.2, pg. 64), essendo T' l'operatore aggiunto definito da

$$\begin{cases} \mathcal{D}(T') = \{ \phi \in L^q(D \times S), \frac{1}{p} + \frac{1}{q} = 1 : \Omega \cdot \nabla \phi - \phi \in L^q(D \times S); B_+ \phi = J B_- \phi \} \\ T' \phi = \Omega \cdot \nabla \phi - \phi. \end{cases}$$

Se $p=1$, l'operatore aggiunto T' è definito da

$$\begin{cases} \mathcal{D}(T') = \{ \phi \in L^\infty(D \times S) : \Omega \cdot \nabla \phi - \phi \in L^\infty(D \times S); B_+ \phi = J B_- \phi \} \\ T' \phi = \Omega \cdot \nabla \phi - \phi. \end{cases}$$

otteniamo $\sigma_p(T') = \emptyset$ con lo stesso procedimento seguito in (ii).

(iv) Da (i), (ii) e (iii) segue immediatamente che $\sigma_p(T) = \{ \lambda \in \mathbb{C}; \operatorname{Re} \lambda = -1 \}$.

4. IL SEMIGRUPPO GENERATO DA B.

Poiché T è il generatore infinitesimale del semigruppato di contrazioni $\mathcal{U}(t)$ e K è operatore limitato in X , l'operatore $B := T + K$ è il generatore infinitesimale del semigruppato fortemente continuo $\{\mathcal{W}(t); t \geq 0\}$ definito come segue:

$$\mathcal{W}(t) = \sum_{j=0}^{\infty} \mathcal{U}_j(t), \quad \mathcal{U}_0(t) = \mathcal{U}(t), \quad \mathcal{U}_j(t) = \int_0^t \mathcal{U}(s) K \mathcal{U}_{j-1}(t-s) ds.$$

Essendo poi K e $\mathcal{U}(t)$ operatori positivi, $\{\mathcal{W}(t); t \geq 0\}$ risulta essere un semigruppato di operatori positivi.

Dalla disuguaglianza $\|\mathcal{W}(t)\| \leq e^{-(1-\|K\|)t}$, segue che $\{\mathcal{W}(t); t \geq 0\}$ è semigruppato di contrazioni se $\|K\| \leq 1$.

Pertanto sono vere le seguenti affermazioni:

(i) per ogni $f_0 \in \mathcal{D}(T)$ il problema astratto di Cauchy (2) ha soluzione unica data da

$$f(t) = \mathcal{W}(t)f_0;$$

- (ii) se il dato iniziale è non negativo, la soluzione forte del problema (2) è anche essa non negativa;
- (iii) l'applicazione $t \rightarrow \|f(t)\|$ è non crescente se $\|K\| \leq 1$;
- (iv) se $1 - \|K\| > 0$ è $\|f(t)\| \leq e^{-(1-\|K\|)t} \|f_0\|$ e quindi $\lim_{t \rightarrow \infty} f(t) = 0$.

Per quanto riguarda il semigruppato $\{\mathcal{W}(t); t \geq 0\}$ generato da B , si dimostra una ulteriore proprietà enunciata nel seguente

TEOREMA 3. *Il semigruppato $\{\mathcal{W}(t)\}$ è semigruppato irriducibile di operatori positivi.*

Dim. Ricordiamo che un semigruppato di operatori positivi $\{\mathcal{W}(t); t \geq 0\}$ su un lattice di Banach X si dice irriducibile se gli unici ideali chiusi invarianti rispetto a $\{\mathcal{W}(t); t \geq 0\}$ sono $\{0\}$ e X .

Gli ideali chiusi di X hanno la forma:

$$\mathcal{I}_E = \{f \in X : f(\vec{r}, \vec{\Omega}) = 0 \text{ per q.o. } (\vec{r}, \vec{\Omega}) \in D \times S \setminus E\}.$$

dove E è un sottoinsieme misurabile di $D \times S$.

Un ideale chiuso è $\mathcal{W}$ -invariante se e solo se lo è sia per $\mathcal{U}$ che per K ([12], p. 307 prop. 3.3).

Pertanto è sufficiente mostrare che se $\mathcal{I}_E$, con $\eta(E) > 0$, è invariante sia per $\mathcal{U}$ che per K , segue necessariamente $E = D \times S$.

Poiché $k(\vec{r}, \vec{\Omega}, \vec{\Omega}) > 0$, $\mathcal{I}_E$ sarà invariante per l'operatore K solo se $E = D_0 \times S$, essendo D_0 un sottoinsieme misurabile di D con $\mu(D_0) > 0$.

Inoltre $\mathcal{I}_E$ è invariante per ogni $\mathcal{U}(t)$ solo se E è invariante per ogni traslazione

$$(\vec{r}, \vec{\Omega}) \rightarrow (\vec{r} - t\vec{\Omega}, \vec{\Omega}) \text{ con } 0 \leq t < t_-.$$

Pertanto D_0 deve coincidere con D , cioè $\mathcal{I}_E = X$.

BIBLIOGRAFIA

1. K.M. Case and P.F. Zweifel, "Linear transport theory", Massachussets 1967.
2. S. Chandrasekhar, "Radiative transfer", Dover, New York 1960.
3. E.B. Davies, "One-parameter Semigroups", Academic Press, London 1980.
4. B. Davison, "Neutron transport theory", Oxford University Press, London 1957.

5. J.J. Duderstadt and W.R. Martin, "*Transport theory*", Wiley Interscience, 1979.
6. J.A. Goldstein, "*Semigroups of linear operators and applications*", Oxford University Press, 1985.
7. E. Hille and R.S. Phillips, "*Functional analysis and semigroups*", A.M.S. Coll. Publ., Providence, R. I., 1957.
8. J. Lechner e M.G. Wing, *Comm. Pure and Appl. Math.*, 8, 217-234, (1955).
9. E.A. Milne, "*Handbuch der Astrophysik*", vol. 3, Springer Verlag, Berlin 1930.
10. R. Nagel, "*One-Parameter Semigroups of Positive Operators*", L. N. n° 1184, Springer Verlag, Berlin 1986.
11. A. Pazy, "*Semigroups of linear operators and applications to partial differential equations*", Springer Verlag, 1983.
12. J.A. Van Casteren, "*Generators of strongly continuous semigroups*", Research Notes in Mathematics vol. 115, Pitman, 1985.
13. J. Voigt, "*Functional analytic treatment of the initial boundary value problem for collisionless gases*", Habilitationsschrift, München 1980.

APPLICAZIONI DELLA MATEMATICA IN PROBLEMI DI COMUNICAZIONE

Luigia Berardi

Dipartimento di Ingegneria Elettrica, Università, L'Aquila.

1. INTRODUZIONE

Sono molto lieta di poter intervenire in questo Convegno, che del resto ho contribuito ad organizzare, e sono lieta di poter parlare a tutti gli intervenuti di questo argomento riassunto nel titolo del mio intervento: le applicazioni della Matematica nei problemi di comunicazioni. Questo argomento è alla base della Crittologia, che è uno degli argomenti del nostro gruppo di ricerca. Inoltre la crittologia e i problemi di comunicazione in genere, sono di interesse non solo nell'ambiente degli Ingegneri ma anche in ambiente Economico. Ciò naturalmente è conseguenza della profonda rivoluzione che l'Elettronica e l'Informatica hanno provocato nel trattamento dell'Informazione e quindi in tutte le transazioni finanziarie del mondo di oggi.

Si pensi che soltanto un centinaio di anni fa avvenne una prima significativa rivoluzione in relazione al nostro problema.

Il primo francobollo postale è del 1839 (Inghilterra), il servizio "Pony Express", ormai mitico e leggendario, è del 1860. Nello stesso periodo nasce il telegrafo via cavo (Morse-1850) e nascono le prime macchine da scrivere (Sholes-1870), il telefono (Bell e Meucci -1876), il cinema (F.lli Lumiere-1895) e il fonografo (Edison-1899). L'aereo è successivo, è del 1918 l'inizio di un costosissimo servizio postale aereo.

(*) Lavoro eseguito nell'ambito del gruppo 40% MPI-Comitato 1.3: Modelli economici, dinamici, discreti, ottimizzazione, crittologia. Il lavoro è stato anche oggetto di una Conferenza durante il Convegno Nazionale della "Mathesis" tenutosi a Gioia del Colle (BA) nel Maggio 1989.

Fin dal 1895 Guglielmo Marconi riusciva a trasmettere, telegraficamente, senza cavo a distanze sempre maggiori. Con l'invenzione della valvola termoionica (Fleming- 1907) nasce la vera possibilità tecnica di una "radio-comunicazione". Nel 1930 Marconi, dal suo panfilo "Elettra", ancorato nel porto di Genova, inaugura l'Esposizione mondiale di "Sidney" parlando agli Elettrotecnici riuniti a congresso e accendendo luci augurali mediante impulsi radio.

La televisione, da una intuizione dell'abate Caselli (1860) risale al 1930, quando fu presentato il primo brevetto da Farnsworth. Essa viene inaugurata in Italia sulla fine degli anni '50.

L'elaboratore elettronico ha enormemente facilitato lo sviluppo delle comunicazioni. Il primo vero elaboratore ENIAC venne realizzato attorno al 1945, esso era dotato di ben diciottomila valvole. La sempre maggiore necessità di migliori comunicazioni condusse nel 1964 alla costituzione di un Consorzio INTELSAT (International Telecommunication Satellite Consortium) che si occupasse della realizzazione e gestione di satelliti artificiali che consentissero collegamenti telefonici e televisivi in tempo reale. Satelliti che percorrono in un giorno tutta l'orbita terrestre ad una distanza di 36.000 Km dalla Terra, tre di essi sono in grado di gestire le comunicazioni di tutto il nostro pianeta.

L'avvento dell'elettronica, la scoperta dei microprocessori che sono stati realizzati negli Stati Uniti, grazie alla corsa allo "spazio", durante gli anni '60 sono il fenomeno forse più significativo in questo contesto. Oggi in quasi ogni casa del mondo, ad alto tenore economico, esiste un Personal Computer, che per potenza e capacità elaborative è di gran lunga superiore a un gigantesco computer che ieri occupava un palazzo di più piani.

Siamo di fatto in grado e ciò sarà realizzato certamente in pochi anni, di metterci in rete con il resto del mondo. Oggi infatti un gran numero di messaggi viaggiano su canali elettronici: posta elettronica, telex, moneta elettronica, ordini e transazioni bancarie. Un grande pericolo incombe: una massa di informazioni che cada in mano di personaggi non autorizzati può dar luogo ad un sottobosco di illeciti di vario genere quale ad esempio lo spionaggio industriale con il relativo commercio delle informazioni ovvero la contraffazione di transazioni ovvero della posta elettronica ovvero l'appropriazione del denaro elettronico. Nasce quindi la necessità di *proteggere l'informazione*.

Si pensi soltanto ai rischi che si possono correre se un "bandito elettronico" acceda a reti o archivi quali quelli di Banche, Enti assicurativi, di Società ovvero di Enti governativi. Ci si potrebbe appropriare quindi non solo di denaro ma anche di informazioni su movimenti contabili, situazioni fiscali, piani di produzione e sviluppo, informazioni militari etc.etc., un operatore finanziario potrebbe disconoscere l'ordine elettronico di un cliente e viceversa.

Un rimedio è certamente quello di disporre di buoni prodotti crittografici. La crittografia, che oggi si presenta ben più sofisticata che non ieri quando il suo uso era riservato al solo ambito militare, è oggi una disciplina nella quale

si osserva un grosso fermento. Convegni internazionali, europei, nazionali, gruppi di ricerca, conferenze, tesi di Laurea e Corsi Universitari.

Lo scopo di questa Conferenza è quello di accostare alla problematica della Crittografia ed alla Matematica che in essa si adopera. Attraverso una panoramica di tipo anche storico ci si vuole accostare a Modelli Matematici concreti, nei quali è utilizzata una Matematica di estrema semplicità. Gli argomenti sono presentati in forma problematica ed euristica. Si parte da situazioni concrete, anche sotto forma di gioco, ovvero dal racconto di fatti storici dai quali è facile evincere la visione di una Crittografia interagente con la Matematica in una continua e stimolante evoluzione.

Tratteremo essenzialmente i seguenti tre problemi:

- I. Come “viaggia” un messaggio attraverso un “filo conduttore” o, meglio “via cavo”? Come stabiliamo se il messaggio ricevuto è quello trasmesso?
- II. Come si costruiscono le sequenze di numeri e/o di simboli che quotidianamente vediamo su prodotti commerciali e a cosa servono?
- III. Come si può trasmettere un messaggio segreto sotto gli occhi di tutti?

Naturalmente approfondendo in altro ambito queste problematiche si incontra anche matematica più sofisticata. Nei problemi di autenticazione ad esempio si fa grande uso della Matematica Discreta, che nella Crittografia ha trovato una delle sedi più naturali per le Applicazioni.

Per chi volesse approfondire i problemi connessi con l'autenticazione di un messaggio elettronico, ovvero con la firma elettronica consiglio i miei articoli indicati in bibliografia con i numeri [1] e [2]. Un interessante libro, esauriente e vivace in pari tempo è quello di Andrea Sgarro [8]. Un classico con una bibliografia poderosa è il Manuale scritto nel 1948 dal Generale Sacco [7]. Infine un articolo che indichi problematiche complementari a quelle che andremo a trattare ora, scritto da me e dall'amico Prof. Bruno Rizzi, è apparso sul Periodico di Matematiche n.4 (1987), 1-22.

2. UNA APPLICAZIONE DEL SISTEMA DI NUMERAZIONE BINARIO

Dopo aver trattato i vari sistemi di numerazione, ed in particolare quello in base 2, si può porre il problema della trasmissione di un messaggio attraverso un filo conduttore. Costruiamo

“UN CODICE NATURALE”

I PASSO

Ad ogni lettera dell'alfabeto associamo il suo numero d'ordine, rappresentato in base 2. Si ha:

A	→ 1	→ 2^0	→	1
B	→ 2	→ 2^1	→	10
C	→ 3	→ $2^1 + 2^0$	→	11
.....				
Z	→ 26	→ $2^4 + 2^3 + 2^1$	→	11010

Se ai simboli 1 e 0 associamo due “segnali” diversi, ad esempio ad 1 una tensione positiva ed a 0 una tensione negativa, oppure ad 1 corrisponde il passaggio di corrente ed a 0 il non passaggio, la trasmissione è possibile.

Ma, all’arrivo, decodificare la sequenza di 0 ed 1 non è possibile perché ogni lettera non è rappresentata da una sequenza di lunghezza costante. Quindi se ad esempio arriva 11010, non si può stabilire se è Z oppure FS (essendo $F=110$ e $S=10$).

II PASSO

Rappresentiamo ogni lettera con una sequenza di 5 simboli (tanti quanti ne occorrono per l’ultima lettera dell’alfabeto), anteponendo in ogni sequenza tanti zeri quanti ne servono.

A	→ 1	→ 2^0	→	00001
B	→ 2	→ 2^1	→	00010
C	→ 3	→ $2^1 + 2^0$	→	00011
.....				
Z	→ 26	→ $2^4 + 2^3 + 2^1$	→	11010

Ulteriori sequenze possono essere usate, volendo, per codificare anche i simboli di interpunzione (ad esempio, = 11011 etc.)

A questo punto è semplice trasmettere e ricevere, se ogni segnale trasmesso arriva in uscita senza “ombre”.

Ma ciò in genere non accade nei canali reali di comunicazione, nei quali sono presenti dei disturbi, che sono causa di *errori*, nel senso che qualche “segnale” ricevuto può essere diverso da quello corrispondente trasmesso. Così, ad esempio, se trasmettiamo C e l’ultimo 1 diventa 0 durante la trasmissione, riceviamo B, ed il nostro codice non *rivela* l’errore, nel senso che non ci possiamo accorgere che era stata trasmessa C e non S.

I canali di trasmissione sono buoni abbastanza per poter supporre che in ogni sequenza non entri più di un errore.

III PASSO

Costruiamo un codice rivelatore di un errore.

In ogni sequenza rappresentativa di una lettera, aggiungiamo un simbolo

scelto tra 0 ed 1 in modo che nella sequenza stessa compaia un numero pari di 1 (Codice di controllo di parità), (ovvero un numero dispari di 1 (Codice di controllo di disparità)).

A → 00001 → 000011
 B → 00010 → 000101
 C → 00011 → 000110

L'ultimo simbolo, quello sottolineato, è detto *simbolo di controllo di parità*. Ogni blocco di questo nuovo codice è lungo 6. Se in un blocco in uscita compare un numero dispari di 1, vuol dire che in quel blocco c'è un errore. Dunque abbiamo costruito un codice binario *rivelatore di un errore*.

Il problema della correzione di un errore è più complicato dal punto di vista tecnico, ma vale la pena di introdurre almeno la problematica, per completezza.

IV PASSO

Esistono codici che, oltre a rivelare errori, sono in grado di correggere un certo numero di errori, essi prendono il nome di *codici correttori* di errori. Relativamente ad essi diciamo soltanto che se vogliamo un codice correttore di 1 errore, i blocchi rappresentativi delle lettere devono avere a due a due almeno tre simboli diversi.

Infatti, se due qualsiasi blocchi hanno almeno tre simboli diversi, ed entra un solo errore in un blocco, allora nel codice c'è un solo blocco che ha un solo simbolo diverso dal blocco in uscita, mentre tutti gli altri differiscono dal blocco in uscita per almeno 2 simboli.

ESEMPIO.- Supponiamo che un codice sia costituito dalle due parole:

0000 1110

(Le due sequenze potrebbero significare NO e SI, oppure FALSO e VERO oppure NERO e BIANCO, etc.). Trasmettiamo 0000 e supponiamo che entri 1 errore, cioè uno dei quattro simboli diventa 1, allora possiamo avere una delle sequenze:

(*) 1000 0100 0010 0001

Poiché l'altra parola del codice, cioè 1110, ha almeno due simboli diversi da ogni sequenza (*), è possibile correggere la sequenza in uscita sostituendola con la parola del codice che differisce da essa per un solo simbolo. Il principio applicato si chiama *principio di massima somiglianza*.

3. UNA APPLICAZIONE DELLA DIVISIBILITA' PER 11

Nella vita di tutti i giorni capita di vedere impresse sulle confezioni di prodotti commerciali delle sequenze di numeri e/o simboli, che caratterizzano

il prodotto.

Tali sequenze sono costruite mediante *codici rivelatori di errore*. Esempi sono dati dal codice EAN (**E**uropean **A**rticle **N**umber), che identifica i prodotti alimentari (vedasi l'articolo di Ippoliti et altri su Ratio 1), dal codice ISBN (**I**nternational **S**tandard **B**ook **N**umber) che identifica i libri, dal Codice Fiscale Italiano, che identifica ogni persona Italiana o che ha rapporti di lavoro con l'Italia. Il significato dei vari simboli che compaiono in una sequenza costruita tramite uno di questi codici scaturiscono da un preciso procedimento matematico. Inoltre è interessante vedere come viene costruito il simbolo di controllo che permette di rilevare un errore. Allo scopo si presta molto bene il codice ISBN, nel quale il simbolo di controllo viene stabilito in base alla divisibilità di un numero per 11.

Leggiamo su un libro :

ISBN 88 08 03858 0

10 numeri in fila, ognuno con un significato. Vogliamo controllare se il "numero del libro" è giusto, indipendentemente dal suo significato. Moltiplichiamo da sinistra verso destra ogni simbolo per i numeri 10, 9, 8, 7, 6, 5, 4, 3, 2, 1 e sommiamo. Si ottiene:

$$10 \cdot 8 + 9 \cdot 8 + 8 \cdot 0 + \dots + 2 \cdot 8 + 1 \cdot 0 = 286$$

ed essendo la somma pesata calcolata, divisibile per 11, il numero ISBN è giusto! Vogliamo ora capire.

Un numero ISBN è una sequenza di 10 simboli, anche ripetuti, presi nell'insieme $\{0, 1, 2, \dots, 9, X\}$. La prima parte, costituita da 1 o 2 simboli, indica la Nazione della Casa Editrice. La seconda parte è il numero della Casa Editrice, la terza individua il libro, l'ultima parte, costituita da un solo simbolo è un simbolo di controllo. Il simbolo X, che sta per 10, può comparire solo come simbolo di controllo. Vediamo come si calcola il simbolo di controllo. Sia $a_{10}, a_9, a_8, \dots, a_2, a_1$ un numero ISBN nel quale tutti i numeri sono noti tranne a_1 . Calcoliamo la somma

$$S = 10 a_{10} + 9 a_9 + 8 a_8 + \dots + 2 a_2$$

Il simbolo a_1 , per definizione, è il più piccolo intero positivo tale che $S + a_1$ sia un multiplo di 11. E' chiaro che a_1 può prendere un valore tra 0 ed $X=10$.

Tale codice rivela un errore. Infatti supponiamo che il numero ISBN

$$a_{10} a_9 a_8 \dots a_2 a_1$$

presenti un errore, ad esempio nel primo simbolo. Allora esso diventa

$$a'_{10} a_9 a_8 \dots a_2 a_1$$

La somma $S = 10 a_{10} + 9 a_9 + 8 a_8 + \dots + 2 a_2 + a_1$ è divisibile per 11.
Supponiamo che anche la somma

$$S' = 10 a'_{10} + 9 a_9 + 8 a_8 + \dots + 2 a_2 + a_1$$

sia divisibile per 11. Allora anche la loro differenza è divisibile per 11, cioè

$$S - S' = 10 a_{10} - 10 a'_{10} = 10 (a_{10} - a'_{10})$$

è divisibile per 11. Essendo 11 un primo, segue che $a_{10} - a'_{10}$ è divisibile per 11.
Poiché $a_{10} \leq 9$ ed $a'_{10} \leq 9$ segue che:

$$-9 \leq a_{10} - a'_{10} \leq 9$$

L'unica possibilità è allora data da $a_{10} = a'_{10}$.
Quindi se vi è un errore il numero ISBN (errato) non è divisibile per 11.
Naturalmente non è faticoso inventare dei "codici ISBN" con 4, 6, 12 cifre legati ai numeri primi 5, 7, 13 ed ai rispettivi criteri di divisibilità.

4. LA SEMPLICITA' DELLA MATEMATICA NEI CODICI SEGRETI: LORO EVOLUZIONE STORICA

Il problema trattato nel paragrafo 2 è quello della trasmissione di un messaggio usando un codice rivelatore di almeno un errore.

Non abbiamo trasmesso messaggi segreti, e quindi tutti potevano decodificare e conoscere i messaggi trasmessi.

In questo paragrafo il problema è ben diverso. Il problema è quello di trasmettere *messaggi segreti*, cioè trasmettere in un canale pubblico un messaggio in modo che esso sia *comprensibile solo al mittente ed al destinatario*. Questo problema è molto antico. Un metodo per raggiungere lo scopo, secondo Plutarco, è stato usato già nel IX secolo a.C. Tale metodo è noto come "Scitala Lacedemonica" e consiste in un cilindro su cui viene avvolta ad elica una striscia (a quei tempi di cuoio), sulla quale viene scritto un messaggio per righe longitudinali. Tolto il cilindro, ... è realizzato il caos delle lettere. Il destinatario riavvolge la striscia su un cilindro identico al precedente e ... il messaggio è chiaro. La chiave segreta, cioè il segreto, è il diametro del cilindro.

Vogliamo ora "nascondere" un messaggio letterale, sostituendo ad ogni lettera che in esso compare un'altra lettera. L'arte di "nascondere" i messaggi si chiama *Crittografia*.

Scriviamo due alfabeti su due righe successive in modo che le lettere del secondo alfabeto siano slittate di un certo numero di posti rispetto a quelle del primo. Chiamiamo il primo alfabeto *alfabeto in chiaro* ed il secondo *alfabeto*

cifrante.

ESEMPIO

CHIARO : A B C D E F G H I L M N O P Q R S T U V Z

cifrante : f g h i l m n o p q r s t u v z a b c d e

Sostituiamo ogni lettera del messaggio da trasmettere con la lettera corrispondente ad essa sull'alfabeto cifrante. Ad esempio si ha:

TUTTI ASPETTIAMO LA PRIMAVERA
bcbbp faulbbpftrt qf uzprfdlzf

E' un gioco "tradurre il messaggio" cioè cifrare. Se il destinatario conosce il numero dei posti di cui è stato slittato l'alfabeto (tale numero è la chiave segreta del codice), è ancora un gioco "ritradurre" cioè decifrare il messaggio ricevuto.

Una persona sa che un messaggio in italiano è stato cifrato (o *criptato*) con questo sistema, ma non conosce la chiave segreta usata, vede il messaggio e ne vuole conoscere il contenuto, anche se non è autorizzato, vuole cioè *forzare* il codice.

Tale arte si chiama *Crittoanalisi*.

Prima di tutto per forzare il codice, il messaggio deve essere "abbastanza lungo". In tale caso la crittoanalisi di un testo è basata sul fatto che in ogni lingua ogni lettera si presenta con una frequenza sua propria. Esistono tabelle delle frequenze nelle lingue più usate.

Frequenza delle lettere di una lingua.

Tedesco	Inglese	Francese	Italiano	Spagnolo	Portoghese
A 5	A 7.81	A 9.42	A 11.74	A 12.69	A 13.5
B 2.5	B 1.28	B 1.02	B .92	B 1.41	B .5
C 1.5	C 2.93	C 2.64	C 4.50	C 3.93	C 3.5
D 5	D 4.11	D 3.38	D 3.73	D 5.58	D 5
E 18.5	E 13.05	E 15.87	E 11.79	E 13.15	E 13
F 1.5	F 2.88	F .95	F .95	F .46	F 1
G 4	G 1.39	G 1.04	G 1.64	G 1.12	G 1
H 4	H 5.85	H .77	H 1.54	H 1.24	H 1
I 8	I 6.77	I 8.41	I 11.28	I 6.25	I 6
J ..	J .23	J .89	J ..	J .56	J .5
K 1	K .42	K ..	K ..	K ..	K ..
L 3	L 3.60	L 5.34	L 6.51	L 5.94	L 3.5
M 2.5	M 2.62	M 3.24	M 2.51	M 2.65	M 4.5
N 11.5	N 7.28	N 7.15	N 6.88	N 6.95	N 5.5

O	3.5	O	8.21	O	5.14	O	9.83	O	9.49	O	11.5
P	.5	P	2.15	P	2.86	P	3.05	P	2.43	P	3
Q	..	Q	.14	Q	1.06	Q	.61	Q	1.16	Q	1.5
R	7	R	6.64	R	6.46	R	6.37	R	6.25	R	7.5
S	7	S	6.46	S	7.90	S	4.98	S	7.60	S	7.5
T	5	T	9.02	T	7.26	T	5.62	T	3.91	T	4.5
U	5	U	2.77	U	6.24	U	3.01	U	4.36	U	4
V	1	V	1.00	V	2.15	V	2.10	V	1.07	V	1.5
W	1.5	W	1.49	W	..	W	..	W	..	W	..
X	..	X	.30	X	.30	X	..	X	.13	X	.2
Y	..	Y	1.51	Y	.24	Y	..	Y	1.06	Y	..
Z	1.5	Z	.09	Z	.32	Z	.49	Z	.35	Z	.3

Si studia la frequenza di tutte le lettere che compaiono nel testo criptato. Probabilmente la lettera che ha una maggiore frequenza corrisponde nell'alfabeto in chiaro alla lettera che ha una maggiore frequenza in quella lingua.

E' chiaro che quanto più il testo è lungo, tanto più lo studio delle frequenze porta al testo reale.

Il codice descritto è molto antico ed è noto come Codice di Cesare, in quanto usato da Giulio Cesare nella guerra in Gallia e nella corrispondenza epistolare con i familiari. Esso è un esempio di codice monoalfabetico. (Si chiamano monoalfabetici i codici in cui per criptare viene usato un solo alfabeto).

Si potrebbe ragionare diversamente per criptare un messaggio. Dividiamo il testo del messaggio in parti di una lunghezza fissata 1, ogni parte sarà chiamata *blocco*.

In ogni blocco "riordiniamo" le lettere in un modo fissato. Come si può fissare un "qualche modo"? Usiamo ancora una permutazione di 1 oggetti. Ad esempio se 1 è 4 e se fissiamo la permutazione che porta 1 2 3 4 in 3 1 4 2 si ha:

TESTO IN CHIARO:	ARRI	VERE	MOMA	RTED	IORE	OTTO	1 2 3 4
testo cifrato	: rair	rvee	mmao	erdt	rieo	toot	3 1 4 2

cioè in ogni blocco scambiamo le lettere usando la permutazione fissata. In effetti facciamo un anagramma di ogni blocco nel "modo" fissato.

I codici segreti di questo tipo si chiamano codici a trasposizione.

Da un punto di vista storico, possiamo dire che nel Medio Evo, per le condizioni politiche, i codici segreti vengono usati poco ed in genere solo per nascondere nomi di personaggi importanti. Verso la fine del Medio Evo, con l'inizio delle relazioni diplomatiche tra i vari stati, i codici segreti diventano una necessità.

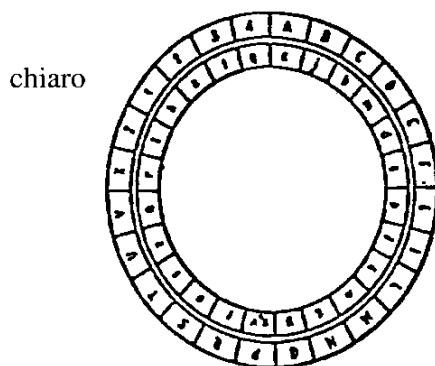
Secondo ricerche storiche dovute a Meister (1902) l'uso sistematico dei

codici segreti ebbe inizio nella Corte Papale, nelle Repubbliche e Signorie Italiane, a partire dal 1300.

E' in questo periodo che la *Crittografia* ha una grossa evoluzione. Troviamo un cambiamento radicale, infatti nascono i primi Codici segreti che non usano un solo alfabeto cifrante, ma molti alfabeti cifranti. Tali codici si chiamano *codici polialfabetici*.

Il primo codice polialfabetico è dovuto ad un nome italiano illustre: Leon Battista Alberti. Egli, su commissione di un segretario pontificio, Leonardo Dato, compilò un codice polialfabetico che fu messo a punto intorno al 1466.

Descriviamo tale codice. Esso è realizzato tramite una coppia di cerchi concentrici. Nel disco più esterno, in cui c'è l'alfabeto in chiaro troviamo 24 caselle, 20 delle quali contengono lettere (mancano per ragioni di sicurezza crittografica, le lettere che si presentano con minore frequenza, cioè J, K, Y, W, Q, H) e le rimanenti quattro i numeri 1, 2, 3, 4.



Il disco interno contiene 24 lettere (manca la lettera w ed è u=v, inoltre tali lettere sono in disordine) formanti l'alfabeto cifrante. Esso può ruotare rispetto al primo disco.

Si fissa, prima di cominciare a criptare il messaggio, una lettera dell'alfabeto in chiaro detta *indice* del codice. Poiché c'è una corrispondenza biunivoca tra le caselle dei due dischi, allora alla lettera scelta come indice del codice, sia ad esempio D, corrisponde una ed una sola lettera del disco interno. Come prima lettera del testo cifrato si scrive la lettera corrispondente a D e poi ogni lettera del messaggio viene sostituita con la corrispondente sempre del disco più interno. Supponiamo ora che dopo un certo numero di lettere (anche uno solo) si desideri cambiare alfabeto per rendere più difficile una possibile decrittazione. Si scelga allora uno dei numeri 1, 2, 3, 4 che sono a disposizione, ad esempio il numero 2, che va pensato come inserito nel testo in chiaro. A questo numero si sostituisce la lettera che gli corrisponde nella corrispondenza

dato dai due dischi. Fatto ciò, si ruota il disco finché la lettera corrispondente al numero scelto non si vada a situare esattamente sotto l'indice del codice, la lettera D nel nostro caso. Questa operazione cambia la biezione tra l'alfabeto in chiaro e quello cifrante, quindi otteniamo un altro alfabeto. fino a che non decidiamo di fissare un nuovo numero e così via.

Il codice dell'Alberti non ha avuto molto successo. La tendenza dominante nel periodo e quella di semplificare il modo di criptare. Troviamo vari altri codici polialfabetici (ricordiamo, ad esempio, quelli di G. Cardano e di Bellaso) e viene apportata un'altra innovazione: l'introduzione della *parola chiave*. Anche questo è dovuto ad un italiano Giambattista Della Porta (1563), l'inventore della camera oscura.

Vediamo come si usa la parola chiave in un codice polialfabetico.

1. Si fissa una parola del tutto arbitraria ma contenente lettere tutte distinte (ciò perché ad ogni parola corrisponderà un diverso alfabeto).

2. Si scrive tale parola il numero di volte necessario per avere una lunghezza pari a quella del messaggio.

ESEMPIO. Sia PORTA la parola chiave e si voglia inviare un messaggio:

testo in chiaro : NOI ABBIAMO UN SEGRETO
parola chiave : por taporta po rtaport

Cosa ne facciamo della parola chiave? Ogni lettera di essa, ad esempio p, ci dice quale alfabeto dobbiamo usare per criptare la lettera del messaggio corrispondente. Nel caso generale quindi si deve dare una permutazione dell'alfabeto per ogni lettera della parola chiave. La costruzione di queste permutazioni costituisce il codice stesso. Come esempio vediamo il *sistema Porta* che si assegna mediante le seguenti *tavole*, dette di Della Porta.

Tavola Della Porta

ab	A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
cd	A B C D E F G H I J K L M Z N O P Q R S T U V W X Y
ef	A B C D E F G H I J K L M Y Z N O P Q R S T U V W X
gh	A B C D E F G H I J K L M X Y Z N O P Q R S T U V W
ij	A B C D E F G H I J K L M W X Y Z N O P Q R S T U V
kl	A B C D E F G H I J K L M V W X Y Z N O P Q R S T U
mn	A B C D E F G H I J K L M U V W X Y Z N O P Q R S T
op	A B C D E F G H I J K L M T U V W X Y Z N O P Q R S
qr	A B C D E F G H I J K L M S T U V W X Y Z N O P Q R
st	A B C D E F G H I J K L M R S T U V W X Y Z N O P Q
uv	A B C D E F G H I J K L M Q R S T U V W X Y Z N O P
wx	A B C D E F G H I J K L M P Q R S T U V W X Y Z N O
yz	A B C D E F G H I J K L M O P Q R S T U V W X Y Z N

In questo esempio le lettere minuscole scritte in testa danno il nome all'alfabeto di quella riga, che è ottenuto dividendo l'alfabeto in due parti di 13 lettere ognuna e stabilendo una biezione tra i due insiemi di 13 elementi. Allora se la lettera della parola chiave che stiamo considerando è una a oppure una b, si cripta la lettera del testo corrispondente con l'alfabeto di nome ab, e così via.

parola chiave : **por taporta po rtaport**
 testo in chiaro : **NOI ABBIAMO UN SEGRETO**
 testo cifrato : **hin rouosqb bh avtlxbk**

Un codice polialfabetico, che ha raggiunto maggiore notorietà, è quello dovuto al francese Blaise de Vigenère. Egli nel 1586 pubblica il suo codice nel quale fa uso di una tavola quadrata, già introdotta dall'abate Tritemio e nota come *tavola di Vigenère*, sulla quale vi è veramente molto da dire.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
B	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A
C	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B
D	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C
E	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D
F	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E
G	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F
H	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G
I	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H
J	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I
K	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J
L	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K
M	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L
N	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M
O	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N
P	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
Q	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
R	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
S	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
T	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
U	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
V	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
W	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
X	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
Y	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
Z	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y

Ogni riga della tavola è un alfabeto di Cesare, a cui diamo il nome della lettera posta a sinistra. Per criptare un messaggio si usano tanti alfabeti di Cesare quante sono le lettere della parola chiave.

ESEMPIO

parola chiave : **por taporta po rtaport**
 testo in chiaro : **NOI ABBIAMO UN SEGRETO**
 testo cifrato : **ddc tbqzrgomc.....**

La tavola è un quadrato latino cioè una matrice in ogni riga e colonna della quale vi è una permutazione della prima riga o colonna. Si tratta anche della tabella additiva delle classi resto modulo 26, quindi un esempio di gruppo finito. Se prendo 24 ho la matematica dell'orologio e se prendo la moltiplicazione modulo 26 ho un caso di divisori dello zero. Si può anche scrivere 00,01,...,25 al posto di A,B,...,Z e allora il passaggio dal testo in chiaro a quello cifrato è la somma modulo 26. Insomma gli spunti sono tanti.

Il segreto di questo codice è tutto nella parola chiave. Quindi, il destinatario del messaggio (e si spera solo lui, oltre il mittente) conosce la parola chiave. Allora è in grado di decifrare il messaggio in arrivo usando il procedimento al contrario (ovvero la sottrazione modulo 26).

Il problema è molto più complicato che nei codici di Cesare o comunque monoalfabetici. Si pensi che il codice di Vigénère è stato usato da eserciti di diversa nazionalità ed ha resistito a tutti gli "attacchi" dei crittoanalisti per circa tre secoli.

Doveva arrivare un ufficiale prussiano di nome Kasiski (1863) per forzare il codice di Vigénère. Il metodo da lui usato è noto come *test di Kasiski* e si può applicare se il testo cifrato è abbastanza lungo. Il test di Kasiski è basato sui seguenti due punti.

1. Trovare la lunghezza della parola chiave.
2. Trovare le lettere della parola chiave.

Trattiamo il punto 1. Si cercano nel messaggio cifrato tutte le sequenze (cioè i gruppi) di tre o più lettere consecutive che si ripetono. *E' molto probabile* che a sequenze uguali del cifrato corrispondano sequenze uguali del testo in chiaro. Se così è vuol dire che le prime lettere di sequenze uguali sono state criptate con lo stesso alfabeto della tavola di Vigénère, analogamente le seconde, le terze, ... Da ciò segue che alle prime lettere delle sequenze uguali corrisponde la stessa lettera della parola chiave, analogamente alle seconde lettere e così via. Ma allora la distanza (= numero delle lettere) tra due sequenze uguali è un multiplo della lunghezza della parola chiave.

Quanto è lunga la parola chiave? *Molto probabilmente* la sua lunghezza è pari al M.C.D delle distanze tra le sequenze uguali tra loro, che si ripetono.

Chiaramente una distanza “strana” tra due sequenze uguali, nel senso che non ha divisori comuni con le altre distanze, deve essere scartata; questo è il motivo per cui abbiamo detto “molto probabilmente la sua lunghezza ...”. Ciò accade quando due sequenze uguali non corrispondono a due sequenze uguali del testo in chiaro.

Il problema al punto 1 è risolto e la lunghezza 1 è nota.

Trattiamo il punto 2. Conosciamo la lunghezza 1 della parola chiave. Il problema è scoprire le lettere della parola chiave. Notiamo che, nel testo cifrato, alla prima lettera, alla $(1+1)$ - ma lettera, alla $(2l+1)$ - ma lettera ... corrisponde la stessa lettera della parola chiave e quindi tutte queste lettere sono state criptate con uno stesso codice di Cesare (cioè con una stessa riga del quadrato di Vigénère).

Ripetendo il ragionamento, si ha che la seconda, $(1+2)$ - ma riga, $(2l+2)$ - ma riga, ..., sono state criptate con lo stesso codice di Cesare. In definitiva si ha:

$$\begin{array}{l} \{ \quad I, \quad (1+1) - \text{ma}, \quad (2l+1) - \text{ma}, \quad \dots \} \\ \{ \quad II, \quad (1+2) - \text{ma}, \quad (2l+2) - \text{ma}, \quad \dots \} \\ \{ \quad III, \quad (1+3) - \text{ma}, \quad (2l+3) - \text{ma}, \quad \dots \} \\ \dots\dots\dots \\ \{ 1 - \text{ma}, \quad 2l - \text{ma}, \quad 3l - \text{ma}, \quad \dots \} \end{array}$$

Le lettere di ognuno di questi insiemi sono state criptate con lo stesso codice di Cesare. Allora studiamo la frequenza delle lettere in ognuno di essi, con lo stesso metodo usato nei codici di Cesare. Scoperta una lettera, è noto il codice di Cesare usato, e quindi anche il nome dell'alfabeto (cioè la lettera che compare in testa), che ci dà la lettera della parola chiave. Allora il codice di Vigénère è *forzato* e quindi perde il suo interesse.

E' possibile “complicare” il codice di Vigénère per avere un codice sicuro? (sicuro significa capace di resistere agli attacchi).

SI

Un codice completamente sicuro è il *codice di Vernam* (1926). Questo è un codice di Vigénère (cioè si usa la tavola di Vigénère) nel quale si usa una parola chiave avente una lunghezza pari alla lunghezza del messaggio.

Sembra che il *telefono rosso*, esistente tra la Casa Bianca e il Cremlino, usi un codice di Vernam per comunicare.

Questo codice di Vernam è in un certo qual senso il padre dei codici costruiti mediante messaggi scritti con sequenze binarie e chiavi con sequenze pseudocasuali; codifica e decodifica sono la somma modulo 2. Siamo nel pieno della Teoria di Shannon ed il poco spazio ci impedisce di approfondire. Notiamo solo che la parola chiave, in un codice di Vernam, non è detto che debba essere di senso compiuto (come nel codice di Vigénère, poiché ai tempi di Vigénère il messaggero aveva necessità di ricordare a memoria la parola chiave), ed essa può essere una sequenza, lunga come il messaggio, di simboli

qualsiasi, ad esempio di 0 ed 1 casuali ovvero pseudocasuali. Vi è oggi una grande richiesta di sequenze pseudocasuali che, dal punto di intuitivo, sono sequenze scritte con una legge e quindi facilmente riproducibili in un secondo computer, ma con proprietà statistiche molto vicine a quelle della casualità.

Riprendiamo i codici a sostituzione ed a trasposizione. Se mescoliamo questi codici applicando sia gli uni che gli altri otteniamo un codice composto.

Un codice molto usato oggi per trasmettere un messaggio segreto attraverso un canale pubblico è un codice costruito dalla IBM il cui algoritmo (ma non le chiavi che ognuno fissa come vuole) è stato addirittura pubblicato al Federal Register of Information Processing Standard con data 1 Agosto 1975. Tale codice si chiama Data Encryption Standard, in sigla DES.

Il DES è un codice binario. Per trasmettere un messaggio segreto attraverso un canale pubblico con il DES si opera nel modo seguente:

1. Si trasforma il messaggio in una sequenza di simboli binari.
2. Si suddivide tutta la sequenza in blocchi, ciascuno costituito da 64 simboli.
3. Ad ogni blocco si applicano per 16 volte codici a sostituzione e a trasposizione.

La chiave segreta K del DES è data da tutte le permutazioni scelte nei 16 passi.

Il DES è oggi uno dei codici segreti più usati. Ad esempio nella realizzazione di una carta a banda magnetica tipo BANCOMAT, la Banca di emissione usa il DES. Indaghiamo sull'argomento. Una carta Bancomat ha una zona a banda magnetica che può memorizzare alcuni dati. Quando una Banca consegna ad un suo utente una carta del Bancomat, insieme ad essa comunica al cliente un numero segreto che va sotto il nome di PIN (Personal Identification Number), composto da 5 cifre della numerazione decimale.

Come si calcola il PIN?

Il conto corrente dell'utente ha un certo numero N . La Banca cripta il numero N tramite il DES, cioè scrive N in forma binaria e poi applica a questa sequenza il DES, usando una chiave segreta K .

Sulla striscia magnetica viene memorizzato il numero di codice della Banca ed il numero N del conto dell'utente. Quando l'utente usa la carta del Bancomat in un terminale, il terminale legge sulla carta sia N che il numero di codice della Banca che ha emesso la carta. Dal codice della Banca è facile per il terminale desumere da un archivio in suo possesso la chiave K usata dalla Banca, quindi applicando il DES ad N e a K è in grado di ricalcolare il PIN della carta. A questo punto chiede il PIN all'utente che lo deve digitare e solo se i due numeri coincidono il terminale fornisce l'assenso all'operazione. Tutto dura pochi istanti, naturalmente. Se una persona vuole scoprire il PIN di una carta, deve fare mediamente 50.000 tentativi, cioè 200 ore di sportello. Naturalmente ci sono varie altre possibilità, non ultimo leggere il PIN con un binocolo!

L'esempio del PIN, abbastanza semplice, fa capire come può essere usato

un codice segreto per proteggere un messaggio. Oggi con il DES si riescono a realizzare procedimenti molto più sofisticati del semplice calcolo di un PIN, quali ad esempio l'autentifica di un documento inviato elettronicamente o addirittura la sua "firma elettronica", cfr. [1].

BIBLIOGRAFIA

1. L.BERARDI, *Some remarks about an electronic signature derived from a generalized RSA-code*, J.of Information & Opti.Sci. 1 (1990), 189-194.
2. L.BERARDI-A.BEUTELSPACHER, *I buoni angeli custodi, ovvero i protettori di un messaggio*, Archimede 2-3 (1988), 129-140.
3. L.BERARDI-M.DI FONZO, *Protezione delle informazioni su personal computer*, Atti del Primo Simposio Nazionale su "Stato e prospettive della Ricerca Crittografica in Italia", a cura della Fondazione "U.Bordoni", Roma, Ottobre 1987, 167-173.
4. L.BERARDI-F.EUGENI, *Blocking sets e teoria dei Giochi: origini e problematiche*, (dedicato al Prof.Renato Nardini per il suo 70-mo compleanno), Atti Sem.Mat.Fis.Univ.Modena, 34 (1988), 165-196.
5. A.BEUTELSPACHER, *La scuola elementare della teoria dei Codici*, Quaderno n.1, suppl.didatt., Semin. Geom. Combin. Univ. L'Aquila, 1985.
6. M.CERASOLI-F.EUGENI-M.PROTASI, *Elementi di Matematica Discreta*, Zanichelli, Bologna, 1988.
7. L.SACCO, *Manuale di Crittografia*, Roma, Litografia Covi, 1947.
8. A.SGARRO, *Crittografia*, Muzzio Editore, 1986.
9. C.E.SHANNON, *Communication theory of secrecy system*, BSTJ 28 (1949), 666-715.

I lettori interessati a proseguire le ricerche possono consultare i vari volumi della serie Springer Verlag "grigia", dedicata all'Informatica con titoli Crypto 19., Eurocrypto 19..

IL PROBLEMA DELLA PROTEZIONE DELL'INFORMAZIONE II. LA GESTIONE E LO SVILUPPO DEI SISTEMI INFORMATIVI: SICUREZZA CONTRO COSTO ? (*)

Arrigo Bonisoli e Franco Eugeni

Sunto. Nell'attuale scenario dei sistemi informativi pubblici e privati le problematiche economico-gestionali che si presentano sono enormi e svariate. Ad alcune di queste può, se vuole, accostarsi con profitto il matematico. La trattazione dell'informazione può riguardarlo da vicino allora che si pensi ai seguenti tre fondamentali problemi: protezione crittografica del messaggio-informazione, certificazione/firma elettronica dello stesso, rivelazione/correzione di errori in un canale disturbato.

Nella prima parte dell'articolo illustriamo alcuni prodotti crittografici oggi disponibili, i loro limiti e le problematiche che nascono nel loro contesto. Nella seconda parte ci occupiamo da vicino di una classe di codici correttori interessanti soprattutto per le proprietà che ne facilitano l'implementazione.

1. L'ATTUALE SCENARIO DEI SISTEMI INFORMATIVI

In questi ultimi anni abbiamo vissuto e stiamo vivendo una profonda e continua evoluzione dei sistemi informativi. Ci si rivolge ad un gruppo di ingegneri del software, si definisce chiaramente tutto ciò che il sistema deve fare, quindi si reperisce nell'ampio e ricco mercato quanto occorre. Le attuali tendenze nelle medie e piccole aziende sono tutte proiettate verso il miglioramento dell'integrazione uomo-computer e della gestione attraverso sistemi/strumenti computerizzati con utilizzo di chip sempre più potenti. La favola del

(*) lavoro svolto nell'ambito del GNSAGA del CNR con il contributo del MURST

computer delle vecchie generazioni miniaturizzato in un piccolo personal da tavolo è di fatto una realtà. Tra i vari strumenti sono da ricordare i numerosi pacchetti applicativi per specialisti (ingegneri, architetti, medici) o per la gestione (programmi per la contabilità, l'editing, i fogli elettronici di gestione finanziaria). Così per la documentazione, l'archiviazione e la consultazione si è in grado di fare ricerche velocissime e operazioni "taglia e cuci" a livelli veramente alti. E' inoltre possibile lo scambio in tempi reali di grosse moli di documenti tra sedi diversamente ubicate.

Un altro aspetto che ci interessa evidenziare è quello dell'avvento dei "sistemi esperti". Un tale sistema può essere definito come un programma in cui sono condensate le conoscenze di un enorme numero di esperti in un particolare argomento o settore e che in più ha capacità di fare uso di procedure di inferenza. In altre parole un sistema esperto oltre ad immagazzinare notizie che gli arrivano da tutto il mondo le elabora utilizzando "l'intelligenza" della macchina ad integrazione dell'esperienza umana. Un tale sistema ha capacità di scelte e decisioni nel senso che segue. Un sistema esperto—strutturato in un magazzino di conoscenze più una capacità elaborativa di tipo inferenziale—acquisendo informazioni su un problema e utilizzando la propria banca dati è in grado di pervenire a risposte finali, cioè decide e sceglie; inoltre è in grado di fornire all'utente tutti i dettagli sia sui dati utilizzati che sulle procedure applicate. Sistemi esperti sono stati progettati per dedurre situazioni da dati provenienti da sensori, prevedere conseguenze da una serie di azioni, formulare diagnosi a partire da dati rilevati, configurare oggetti dal punto di vista del design tenuto conto di vincoli e richieste opportune e addirittura per gestire risorse finanziarie.

Nella gestione delle aziende moderne, quasi come in un organismo militare, sono ancora di importanza crescente i modelli strategici per la produzione, i piani di sviluppo e della globale gestione delle risorse. Certo è ancora costume imperante per molti quadri d'impresa operare scelte e strategie "a caldo" basate su intuizioni, su sensibilità di mercato ed esperienza. Tuttavia è sempre più importante, a causa della sempre crescente internazionalizzazione del mercato, per la competizione derivante dalla tendenza in atto ad una maggiore omogeneizzazione dei consumi nei più diversi paesi, operare in modo più consapevole. Nasce così in questo mondo economico sempre in fermento il marketing strategico, basato su analisi rigorose e complete. Alla base di queste analisi vi è la banca dati della nostra azienda, le accurate statistiche per le analisi di mercato, l'utilizzo dei sistemi esperti per la gestione delle informazioni che via via vengono immagazzinate per l'ottimizzazione della pianificazione e delle funzioni di produzione. Così la strategia che la grande azienda sviluppa ha alla base un patrimonio divenuto sempre più importante: l'analisi, la gestione, la manipolazione, la trasmissione delle informazioni che possiede.

Nell'attuale scenario vediamo dunque l'informazione al centro dei sistemi economico-gestionali viaggiare da un angolo all'altro del mondo a velocità

incredibili ed ai più diversi livelli. L'informazione si muove all'interno di una direzione aziendale o entro un gruppo di progettazione, da un edificio ad un altro, tra città, regioni, stati; inoltre essa viaggia ed interconnette vari campi dello scibile: per l'informazione non vi è il senso del confine, della distanza, del tempo. I sistemi di posta elettronica sono sempre più efficienti ed affidabili: ciascun utente immette la sua "password" di identificazione e invia i propri messaggi, i quali, codificati in modo più o meno robusto, viaggiano lungo i cavi del telefono e raggiungono rapidamente un qualunque nodo della rete.

Anche la moneta è oggi considerata o almeno trattata alla stregua di un'informazione. Si stanno effettuando studi sempre più numerosi sul problema delle transazioni di tipo economico per via elettronica. Naturalmente prima che la moneta elettronica possa avere l'intero diritto di cittadinanza occorre garantire l'utente con alcuni standard di sicurezza e con una legislazione adeguata. Si pensi ad un messaggio adeguatamente protetto dalle frodi e/o manipolazioni, autenticato, anzi, firmato elettronicamente (chi lo riceve è in grado di provare ad un terzo l'identità del mittente). Un tale messaggio potrebbe essere una transazione finanziaria fatta come davanti ad un notaio, ma in tempi di secondi.

La nuova moneta elettronica è una carta dello stesso aspetto del classico Bancomat ma contenente al suo interno un microprocessore ed una memoria da 8 kilobyte (8.000 caratteri per intenderci). La carta ha capacità elaborative e non può essere contraffatta. Ben nota in Francia (carte a memoire) e nei paesi anglosassoni (smart card), recentemente è stata introdotta anche in Italia dalla Eni Data, come libretto Universitario alle Università di Roma e Bologna. Per maggiori dettagli si veda [9]. Si pensi all'uso di punti di vendita elettronici, i cosiddetti POS (Point Of Sale), con i quali ogni transazione finanziaria viene registrata. Controllo su tutto, dalle tangenti ai ricatti. Potrebbe essere la fine del denaro sporco! Ma certo inventerebbero qualche cosa di nuovo.

Un obiettivo fondamentale nei moderni canali di comunicazione è il trasferimento di informazioni con elevato livello di sicurezza sia ai fini della loro corretta ricezione sia ai fini delle intercettazioni non autorizzate. Si pensi ai canali utilizzati dai satelliti artificiali, dalla posta elettronica, dai sistemi di telecontrollo e di trasferimento elettronico di fondi.

L'impiego di tecniche crittografiche è uno dei metodi più pratici ed economici per la protezione dei dati archiviati e trasmessi. In generale nella progettazione dei sistemi informatici è riconosciuta importanza notevole al problema della sicurezza dei dati e della loro protezione. Deve essere naturalmente essere tenuto in conto che la trasmissione, memorizzazione e l'utilizzo di un sistema crittografico non deve presentare grandi difficoltà ed occorre una attenta analisi costi-benefici per la scelta di algoritmi di protezione, di controllo, di rivelazione di errori o di correzione degli stessi.

Spesso il fallimento di un sistema dipende da cattiva gestione da parte degli utenti, ad esempio uso di una stessa chiave per tempi più lunghi di quanto

raccomandato dai progettisti, riutilizzo di vecchie chiavi, trasmissione di testi chiari dei quali vi sia in giro una versione crittografata.

Naturalmente ogni azienda dovrà classificare il proprio materiale da proteggere, dovrà valutare i rischi connessi con la distruzione dei dati, la loro modificazione e l'eventuale divulgazione. Inoltre dovrà anche ottimizzare certe situazioni, ad esempio minimizzando il numero di persone che negli anni possano acquisire la conoscenza completa di materiale riservato.

Non possiamo concludere questo paragrafo, illustrante appunto lo scenario nel quale si elabora l'informazione, senza parlare di un antico problema di garanzia per un qualsiasi prodotto commerciale, cioè della creazione di standard. Le esigenze che sorgono con la necessità di proteggere dati trattati elettronicamente conducono anche alla loro standardizzazione, almeno per quei sistemi per i quali un hardware oppure un software è di dominio pubblico. Per standard intendiamo quei prodotti che rispecchiano alcune norme tecniche emanate da enti nazionali od internazionali, o anche da privati che abbiano il potere politico o commerciale di farlo. Uno standard offre naturalmente sicurezza perché è sottoposto a verifiche e certificazioni ed anche economicità dividendosi tra molti i costi generali della sua produzione. L'ente internazionale che si occupa di questo problema è l'INTERNATIONAL ORGANIZATION OF STANDARDIZATION (ISO) che raggruppa più di 75 Nazioni. In Italia abbiamo l'ENTE NAZIONALE ITALIANO DI UNIFICAZIONE (UNI) che rappresenta l'Italia presso l'ISO. In ambito UNI l'ente che si occupa del settore informatico è l'UNIPREA presso il quale sono stati proposti vari standard per la protezione dell'informazione.

I due Paragrafi che seguono saranno dedicati allo studio di alcuni codici crittografici in uso oggi, in particolare i ben noti sistemi DES, che è uno standard della IBM, e il codice RSA basato sulla cattiva conoscenza che abbiamo dei numeri primi; quest'ultimo non è uno standard, ma è al centro di vaste ed interessanti polemiche interagenti perfino con la teoria dei numeri.

Lo schema generale della trasmissione di un'informazione sarà per noi quello dato dal seguente schema, del quale tratteremo i vari punti separatamente:

- Il trasmettitore prepara il **messaggio in chiaro**, lo crittografa trasformandolo in **testo cifrato** (eventualmente coincidente con il chiaro) ed appronta un secondo messaggio, che accompagnerà il primo: l'**autenticatore** (eventualmente vuoto).
- La coppia testo cifrato-autenticatore viene inviata nel canale.
- All'arrivo i due messaggi sono sottoposti ad un procedimento di correzione automatica degli errori.
- Utilizzando l'autenticatore ci si rende conto dell'identità del mittente, se l'autenticatore è una firma numerica si appronta una prova d'identità per un terzo.
- Infine si decifra il messaggio utilizzando il sistema crittografico pubblico o privato che sia.

2. CODICI CRITTOGRAFICI SEGRETI ANTICHI E MODERNI

Il significato della dicitura “protezione dell’informazione” è cambiato col passare del tempo. Infatti un tempo proteggere l’informazione significava soltanto impedire ad un non autorizzato la lettura di un messaggio, oggi vuol dire anche impedirne la manipolazione.

Per renderci conto di questa necessità è sufficiente pensare ad una transazione finanziaria che avviene a distanza, oppure ad un telecomando di un impianto strategico. La protezione dell’informazione oggi si attua con metodi crittografici.

Iniziamo questo paragrafo con uno sguardo al passato. Bisogna premettere che nella storia la parola codice è stata veramente usata con molti e diversi significati. Uno di questi è appunto il significato di codice segreto per inviare messaggi segreti.

Sul finire del Medio Evo, con l’inizio delle relazioni diplomatiche tra i vari stati, i codici segreti diventano una necessità. Sembra che l’uso sistematico dei codici segreti abbia inizio nella corte papale, nelle repubbliche e nelle signorie italiane a partire dal 1300. Si diffonde l’uso dei nomenclatori cioè di vocabolari con parole cammuffate e simboli nuovi rispetto alle lettere: è un codice efficiente ma abbastanza primitivo. La difficoltà principale è il trasporto del nomenclatore e nasce l’idea di crittografare con regole e chiavi.

In realtà un codice segreto di questo tipo era di fatto passato alla storia: il codice di Cesare. Scriviamo due alfabeti su due righe successive in modo che le lettere del secondo alfabeto siano slittate di un certo numero di posti rispetto a quelle del primo. Chiamiamo il primo alfabeto **alfabeto in chiaro** ed il secondo **alfabeto cifrante**. Il metodo di cifratura è ovvio, così come la sua rottura. La chiave di un codice di Cesare è la permutazione circolare usata. Dunque le chiavi possibili sono 26. È sufficiente che una persona ipotizzi che un messaggio in italiano sia stato cifrato (o **crittato**) con il codice di Cesare e, pur non conoscendo la chiave segreta usata, riesce a **forzare** il codice. La crittoanalisi di un codice di Cesare può essere anche basata sullo studio della frequenza delle lettere. Infatti in ogni lingua ogni lettera compare con una sua frequenza propria. Segue che, se il messaggio è abbastanza lungo, la lettera che in esso compare con una maggiore (minore) frequenza probabilmente corrisponde alla lettera che nella lingua usata ha una maggiore (minore) frequenza.

Le cose si complicano notevolmente usando in luogo di una permutazione circolare, una permutazione qualsiasi. In tal caso le chiavi possibili sono $26!$. Un codice segreto in cui si usi una qualunque permutazione dell’alfabeto si chiama codice a sostituzione.

Esiste nella pratica un altro tipo di permutazione usata nei codici, la trasposizione. Si tratta di una permutazione a blocchi. Per costruire un tale codice si fissa un intero n ed una permutazione P dei primi n interi. Si suddivide

il testo in blocchi di lunghezza n ed in ognuno di essi si opera con la permutazione P . Il numero delle sue chiavi è maggiore di quello di un codice a sostituzione, quindi è più difficile trovare la chiave.

Il codice di Cesare è, come suole dirsi, monoalfabetico. E' sulla fine del medio evo che la **crittografia** ha una grossa evoluzione. Troviamo un cambiamento radicale, infatti nascono i primi codici segreti che non usano un solo alfabeto cifrante: i **codici polialfabetici**. Non abbiamo lo spazio per illustrare in dettaglio i primi due magnifici codici del tempo: quello costruito da Leon Battista Alberti su commissione di un segretario pontificio intorno al 1466 e quello di Giambattista Della Porta (1563), l'inventore della camera oscura, al quale si deve l'introduzione della **parola chiave**. Troviamo nel tempo vari altri codici polialfabetici come quelli di G. Cardano e di Bellaso. Per la descrizione dei codici suddetti rimandiamo ad un articolo di E. Ambrisi e F. Eugeni [1] e ci occupiamo ora di un codice importante ancora oggi.

Si tratta del codice costruito dal francese Blaise di Vigenère. Egli nel 1586, riciclando e ripulendo da marchiani errori un'idea dell'abate Tritemius, fa uso di una tavola quadrata riportata sotto, nota appunto come **tavola di Vigenère**. La tavola in sostanza è la tabella additiva del gruppo Z_{26} e la codifica e decodifica sono legate all'addizione modulo 26.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
B	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A
C	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B
D	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C
E	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D
F	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E
G	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F
H	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G
I	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H
J	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I
K	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J
L	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K
M	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L
N	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M
O	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N
P	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
Q	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
R	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
S	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
T	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
U	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
V	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
W	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
X	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
Y	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
Z	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y

Esempio di codifica:

parola chiave : **wernhei swern heiswer nh eiswernhei ...**
 testo in chiaro : **accordi italo sloveni in trattative ...**
 testo cifrato : **wgtbyhp ...**

Il segreto di questo codice è tutto nella parola chiave. Il destinatario del messaggio conosce la parola chiave. Allora è in grado di decifrare il messaggio in arrivo usando il procedimento al contrario (ovvero la sottrazione modulo 26).

Il codice di Vigénère è stato usato da eserciti e stati in lungo e in largo e ha resistito a tutti gli “attacchi” dei crittoanalisti per circa tre secoli. Fu decrittato da un ufficiale prussiano di nome Kasiski (1863) con un procedimento noto come **test di Kasiski**. Esso si può applicare se il testo cifrato è abbastanza lungo e la parola chiave abbastanza corta. L’idea di base è quella di scoprire la lunghezza ℓ della parola chiave. Si considerano poi la prima lettera, la $(\ell + 1)$ -esima lettera, la $(2\ell + 1)$ -esima lettera, e così via, ottenendo un messaggio parziale che, essendo codificato con la stesso alfabeto, cioè con la stessa riga della tavola di Vigénère, ovvero con un codice di Cesare, si può decrittare facilmente.

E’ possibile “complicare” il codice di Vigénère per avere un codice sicuro? Un codice completamente sicuro è il **codice di Vernam** (1926). Questo è un codice del tipo di Vigénère nel quale si usa una parola chiave avente una lunghezza pari alla lunghezza del messaggio. Oggi in tempi di calcolatori si usa di fatto un metodo noto come “one-time-pad”. Si prendono un messaggio binario ed una sequenza binaria completamente casuale: codifica e decodifica avvengono sommando modulo 2 le singole componenti. Siamo nel pieno della Teoria di Shannon [46].

Esempio di codifica:

parola chiave: 11001101011100101011101010111000010
 testo in chiaro: 001110010101110010101101010101010
 testo cifrato: 1111010000101110000101111101101000

Una sequenza casuale molto lunga può essere memorizzata ma non si trasmette molto bene. Spesso si usano sequenze pseudo-casuali cioè costruibili con una legge prefissata, cioè programmabili. Inoltre si possono scambiare tra gli utenti tabelle di chiavi o “vermi” lunghissimi e segnalare di volta in volta con qualche criterio la chiave usata o il punto del verme.

Nel 1949 Shaunon [46] ha provato che se la parola chiave o verme è lungo come il testo ed è completamente casuale allora il sistema è indecifrabile. Questo spiega la sempre crescente ricerca di sequenze pseudocasuali, intendendo con questa locuzione sequenze generate al computer con qualche

sistema, quindi non casuali, ma soddisfacenti ad una serie di test statistici che le avvicinino abbastanza alla teorica casualità, il cui unico torto è appunto la non riproducibilità.

Sequenze pseudocasuali $(a_1, a_2, \dots, a_n)$ si costruiscono usando ricorrenze del tipo:

$$a_{n+1} = f(a_1, a_2, \dots, a_n)$$

presupponendo di solito che la legge f sia lineare. Una tecnica di questo tipo sta diventando tuttavia sempre meno sicura, così che vi è una certa tendenza a fare uso di ricorrenze non lineari, come ad esempio nella modernissima tecnica DES.

Ad un testo in chiaro si possono applicare successivamente svariate sostituzioni e trasposizioni. Naturalmente non a caso perché, come ben noto, non è detto che componendo due codici si abbia come risultato un codice più robusto. Un codice siffatto si dice appunto codice composto. I codici composti sono i tipici codici dei nostri giorni. Il DES (Data Encryption Standard) è un codice che è stato pubblicato dalla IBM nel 1977. Il DES è usato principalmente in campo bancario. Esso opera su blocchi di 64 bit. Ognuno di essi è suddiviso in 8 blocchi di 8 bit ciascuno (byte) in cui l'ottavo simbolo è un carattere di controllo. Dunque in ogni blocco ci sono 56 simboli significativi. Tale codice opera alternando trasposizioni e sostituzioni per 16 volte, ed ha 2^{56} chiavi possibili. L'algoritmo DES è stato pubblicato: ne segue che tutta la sicurezza/segretezza del DES risiede nella chiave, scelta su un totale di 2^{56} chiavi. Veramente tante!

Il DES è inoltre uno standard di mercato, esso quindi possiede i vantaggi di economicità e di compatibilità di un codice universale. Lo svantaggio di un metodo standard sta nel fatto che tutti i crittoanalisti stanno cercando di forzare il DES e in caso di riuscita essi potrebbero venire a conoscenza di segreti importanti. In tal caso ogni utente del DES dovrebbe cambiare l'intero sistema protettivo con enormi spese.

Ciò nonostante il DES è oggi uno dei codici più usati. Ad esempio nelle carte a banda magnetica (tipo Bancomat) si usa il DES. Una carta Bancomat ha una striscia (banda magnetica) in cui si possono memorizzare alcuni dati (pochi). Quando una banca consegna ad un suo cliente una carta Bancomat, comunica al cliente anche un numero segreto: il PIN (Personal Identification Number), composto da 5 cifre in numerazione decimale. La banca associa al cliente una stringa N formata dal numero di conto corrente del cliente assieme ad un proprio numero di codice. Il numero N è scritto sulla carta. La banca critta il numero N con il DES, cioè scrive N in forma binaria e poi applica a questa sequenza il DES, usando una chiave segreta K ; l'algoritmo produce un numero a cinque cifre, il PIN = $f(N, K)$. Quando il cliente introduce la carta nel terminale, nel breve tempo di "attendere prego, operazioni in corso ..." il terminale

legge N sulla carta identificando sia il cliente che il codice della banca che ha emesso la carta. Dal codice della banca è facile per il terminale desumere da un archivio in suo possesso la chiave K usata dalla banca, quindi applicando il DES ad N e a K è in grado di ricalcolare il PIN della carta e controllare se coincide con quello che deve digitare l'utente. Se una persona volesse scoprire il PIN di una carta, dovrebbe fare mediamente 50.000 tentativi, cioè 200 ore di sportello. I punti deboli sono il codice segreto, che non si dovrebbe mostrare a nessuno, e l'archivio on-line al quale qualcuno potrebbe riuscire ad accedere. Protezioni del Bancomat: cambiare spesso i numeri e limitare la somma da ritirare giornalmente.

3. CODICI PER L'AUTENTICA DEL MITTENTE E CODICI A CHIAVE PUBBLICA

In genere sono molti i significati che acquista la parola "codice". In questo paragrafo prescindiamo dalla realtà del soggetto che costituisce l'informazione, essa può essere musicale, scritta, parlata, gestuale, mentre il codice è qualcosa di più concreto: può ad esempio essere costituito da simboli su carta, da impulsi elettrici traducibili in voce al telefono o scritti da macchine stampanti, oppure da segnali di vario tipo quali i Morse, le bandierine, e così via. Assumiamo che Mr. Bonisoli trasmetta dati a Mr. Eugeni. Mr. E vuole essere certo che i dati ricevuti siano proprio quelli spediti da Mr. B e non quelli che proditoriamente o per fare uno scherzo ha voluto spedire un certo Mr. H (probabilmente il Prof. Heise). Si tratta allora di usare un codice che garantisca il messaggio. Un codice di questo tipo si può ottenere accludendo al messaggio un certo numero di simboli ridondanti, costruiti come funzione dei simboli del messaggio. Tale funzione deve essere naturalmente un segreto, anzi il segreto del sistema. E' dunque importante che il malintenzionato o scherzoso Mr. H non possa alterare il messaggio in suo favore. Ad impedire che ciò accada i nostri Mr. B e Mr. E si accordano su una chiave segreta C ed un algoritmo F . Ora Mr. B spedisce un messaggio m assieme ad un **autenticatore** che altro non è che il messaggio ridondante e crittografato $a = F(m, C)$. Cosa succede a Mr. E? Egli riceve un messaggio costituito dalla coppia (m', a') , che può essere la modifica di (m, a) escogitata da Mr. H. Il nostro Mr. E forma l'espressione $a^* = F(m', C)$; solo se a^* coincide con a' egli considera autentico il messaggio. Cosa può fare Mr. H? Non conoscendo F e C egli può solo fare dei tentativi per rompere il codice. Le sue possibilità di indovinare la coppia giusta sono espresse dal seguente teorema di Gilbert-MacWilliams-Sloane [28]:

TEOREMA. *Supponiamo equiprobabili i messaggi e le chiavi possibili. Allora in ogni codice di autenticazione la probabilità di vittoria di un estraneo è almeno pari al reciproco della radice quadrata del numero totale delle chiavi.*

Uno schema di autenticazione nel quale la probabilità di successo dell'estraneo sia uguale al valore minimo è detto **perfetto**. Un codice in cui la probabilità suddetta è dello stesso ordine di grandezza del valore minimo si dice **praticamente perfetto**.

Vediamo un esempio di codice perfetto di autenticazione dovuto a Gilbert, MacWilliams e Sloane [28]. Fissiamo il piano proiettivo finito $PG(2, q)$ di ordine q sul campo $GF(q)$. Fissiamo una retta L . Definiamo come **messaggi** i $q+1$ punti della retta L , le **chiavi** sono i q^2 punti esterni alla retta L , inoltre l'**autenticatore** di un dato messaggio m in una fissata chiave è la retta a che li congiunge. Segue che un decrittatore che dall'esterno veda passare la coppia (m, a) e la voglia alterare in (m^*, a^*) può solo tentare di scegliere un punto di $A \setminus L$ sperando che sia proprio la chiave. Ciò può essere fatto allora in esattamente q modi. Essendo q la radice del numero delle chiavi segue che lo schema è perfetto.

Vi sono alcuni esempi per lo più dovuti a Beutelspacher [15], [16] e anche a Tallini e Zanella [20]. Recentemente un semplicissimo schema perfetto basato sui disegni divisibili è stato costruito da Berardi ed Eugeni [9].

Gli schemi di autenticazione più interessanti sono quelli derivanti dai **crittosistemi a chiave pubblica**, dei quali vogliamo ora occuparci.

L'idea nuova del codice a chiave pubblica è che la chiave non è segretamente inviata ma è nota, come su un elenco del telefono. Alla base di questi codici vi sono le **funzioni unidirezionali**. Si tratta di funzioni calcolabili in tempi rapidi, teoricamente invertibili ma con funzione inversa calcolabile in tempi tanto lunghi da risultare proibitivi. Sia C una funzione unidirezionale facilmente calcolabile e sia P la sua inversa, calcolabile in tempi lunghissimi. Mr. **B** e Mr. **E** fanno ora parte di un sistema di scambi gestionali e in un elenco sono riportate le rispettive funzioni C_B e C_E , note a tutti. Ognuno di essi ha un segreto: l'inversa P_B è il segreto di Mr. **B**. Quando Mr. **B** vuole mandare il messaggio m a Mr. **E** gli manda non m ma $C_E(m)$. Può farlo perché legge C_E nell'elenco dei codici che è pubblico. Ma solo Mr. **E** conosce la funzione P_E , e solo lui può decifrare.

Una funzione unidirezionale può essere usata anche come **firma numerica non riutilizzabile** nel modo seguente. Il nostro Mr. **B** concorda una frase breve con un certo numero di persone: dichiara ad esempio che firmerà con la parola $A = ARRIGO$. Invia dunque il messaggio m crittografato o no ad un secondo personaggio assieme alla frase $P_B(A)$. Egli solo conosce la funzione P_B , difficile a chiunque da trovare, ed egli solo è in grado di costruire $P_B(A)$. Ma chiunque abbia accesso all'elenco conosce C_B e se questi conosce anche A è in grado di calcolare $C_B[P_B(A)]$ e vedere se coincide con la parola A prescelta, la firma numerica non riutilizzabile di Mr. **B**, che per un nuovo messaggio cambierà parola. Esempi di funzioni unidirezionali sono il **logaritmo discreto** e la **funzione prodotto di numeri primi**, naturalmente per valori elevati dei numeri che consideriamo. La funzione prodotto di primi viene utilizzata nel codice RSA, costruito da Rivest, Shamir ed Adelman nel 1978, utilizzato per scopi

bancari, essenzialmente come firma o per messaggi corti. E' un sistema molto costoso.

Per descrivere tale codice si consideri un sistema di utenti (organizzazioni civili o private) che partecipino al codice. Ogni utente sceglie due numeri N ed H che pubblica su un elenco ed un terzo numero D che rimane segreto. I numeri sono costruiti con i seguenti criteri. Il numero N è il prodotto di due primi p e q . Allora l'utente sa calcolare il numero di Eulero $E = (p-1)(q-1)$ e può scegliere un numero D primo con E . Ora può prendere un intero H tale che $(DH-1)/E$ sia un intero e renderlo pubblico. Un messaggio ora è una sequenza di numeri. Sia m il generico numero della sequenza e vediamo come si critta.

Se vogliamo mandare un messaggio m ad un utente, ci procuriamo i suoi numeri N ed H e costruiamo il crittogramma c di m , ponendo:

$$c = \min \{X \mid (X - m^H) / N \text{ è un intero}\};$$

l'utente calcola

$$m' = \min \{Y \mid (Y - c^D) / N \text{ è un intero}\}$$

ed è $m' = m$.

Le cifre che si usano attualmente per i primi che formano N sono dell'ordine delle duecento cifre. Recentemente è stata pubblicata a riguardo una notizia che è sembrata un grido d'allarme per il codice RSA. Precisamente si è saputo che un gruppo di matematici guidati da Mark Manasse del Centro di ricerca della Digital Equipment Corporation di Palo Alto e Arjen Lenstra della Chicago University facendo lavorare 400 calcolatori localizzati negli Stati Uniti, in Australia e in Olanda per circa un mese ed in contemporanea hanno scomposto un numero di cento cifre. Esso è risultato il prodotto di due primi rispettivamente di 41 e 60 cifre. Questo risultato sembrerebbe intaccare la sicurezza del codice RSA. Invero, sempre a proposito di questa notizia, non molto preoccupato si è dichiarato il Prof. Silvio Micali del MIT, uno dei big della crittografia internazionale. Avrebbe rilasciato la seguente significativa dichiarazione: "Un pericolo reale potrà venire solo da una nuova idea matematica e non dalla forza bruta di molti computer". D'altro canto il precedente primato mondiale di decomposizione di un numero era stato la decomposizione di un numero di 93 cifre decimali. Tale primato stabilito nel 1988 da Riele, Lioen e Winter richiese 95 ore di calcolo su di un supercomputer. Pertanto l'aumento di complessità tra 93 a cento cifre osservato su un caso da Manasse e Lenstra ci sembra veramente notevole! Tuttavia una nuova generazione di computer superveloci potrebbe spostare il problema molto avanti: quindi in guardia con la tecnologia!

4. STRATEGIE NEL MONDO REALE

Consideriamo un qualsiasi ambiente EDP (Electronic Data Processing). I problemi concreti che si pongono sono molti: impedire il contatto da parte di estranei con il materiale da proteggere, impedire di copiare ed elaborare in qualsiasi modo, ma anche il non permettere ad estranei di vedere o poter fare congetture plausibili. Naturalmente tutto questo era più facile da fare nel mondo militare e lo è un po' meno nel mondo civile. Ovviamente alla base di ogni progetto crittografico vi deve essere una protezione fisica del sistema che si presuppone sempre progettata ed organizzata a monte. E' impensabile ragionare sulla complessità di un sistema crittografico se poi le chiavi si mettono a disposizione di chiunque, oppure proteggere da falsificazioni senza avere da qualche parte una copia dell'originale, e questi sono soltanto esempi banali.

Un ulteriore aspetto da considerare è quello tecnologico. I nostri messaggi possono e devono essere pensati come oggetti di natura elettromagnetica, quindi altamente miniaturizzati e capaci di viaggiare a velocità altissime, con trasmissioni in tempo reale. Per comprendere la natura e la mole fisica di questi messaggi pensiamo che 10 cartelle dattiloscritte, per un totale di 5.000 caratteri circa, sono registrate su un tratto di nastro magnetico di circa 6 centimetri e che bastano 30-40 secondi di registrazione da un normale telefono per ottenerne una copia. L'attacco ad un centro di archiviazione ed elaborazione dati è oggi dunque più facile di ieri. Si pensi all'estensione che può avere una rete di terminali e ai molteplici punti in cui l'ascolto è possibile. Si pensi alla possibilità per un ascoltatore di captare brani di messaggi di cui conosce il corrispondente testo in chiaro, alla possibilità di ottenere statistiche, correlazioni, informazioni di ogni genere. Un'attenta raccolta e le giuste elaborazioni, come l'esperienza prova, possono dare risultati insperati.

Un ostacolo fino ad oggi è stata di sicuro l'incompatibilità di alcuni sistemi, la continua modifica degli hardware, la mancanza di standard, non ultima la buona retribuzione del poco personale altamente specializzato. Ma ragioni legali di certificazione dei sistemi di sicurezza impongono l'uso degli standard. La tecnologia attuale sembra tenda a stabilizzarsi (a parte i futuribili laser-computer che potrebbero dar luogo a nuove rivoluzioni per la loro enorme velocità), il numero dei competenti tende a crescere, insomma tutto fa presagire una crescente necessità di dure protezioni crittografiche.

Da quanto ci risulta, per la protezione di messaggi trasmessi in linea in entrambi i versi sono in uso macchine-hardware autonome, diciamo collegate con i modem e capaci di crittare e decrittare, usualmente con chiave segreta. Altri sistemi proteggono archivi e funzionano con chiavi in possesso di personaggi autorizzati, quali ad esempio programmi applicativi ovvero sistemi operativi.

5. LA CODIFICAZIONE A CORREZIONE NELLA TRASMISSIONE DEI MESSAGGI

La teoria dei codici correttori studia il problema della spedizione di messaggi lungo una linea disturbata, facendo in modo che il messaggio ricevuto coincida per quanto più possibile con quello originario, [2], [32], [33], [35], [36], [45]. Lo scopo è in fondo proprio l'opposto di quello della crittografia. Si pensi al telegrafo, al telefono, alla comunicazione con i satelliti e così via.

In una descrizione unitaria di questi fenomeni si è soliti rappresentare i messaggi da trasmettere come sequenze di simboli provenienti da un **alfabeto** F di cardinalità finita, tipicamente l'alfabeto binario $\{0, 1\}$. La linea di trasmissione prende il nome di **canale** ed è rappresentata da una matrice stocastica detta appunto **matrice di canale**; un elemento di questa matrice indica la probabilità condizionata $p(\beta/\alpha)$ che un certo simbolo β dell'alfabeto esca dal canale dopo che in ingresso era stato spedito un fissato simbolo α . L'ingresso nel canale di un simbolo α darà luogo a un **errore di trasmissione** (cioè all'emissione di un simbolo diverso dal simbolo originario) con probabilità $\sum_{\beta \in F \setminus \{\alpha\}} p(\beta/\alpha)$ che in generale è diversa da zero.

Per diminuire la probabilità di errori di trasmissione si ricorre alla **codificazione** che nella maggior parte dei casi può essere schematizzata come segue. I simboli in ingresso nel canale vengono suddivisi in blocchi, ciascuno costituito da k simboli. A ciascun blocco $x_1 x_2 \dots x_k$ di **simboli d'informazione** vengono aggiunti r **simboli di controllo** $y_1 y_2 \dots y_r$, mediante una opportuna regola; la sequenza o **parola** $c = x_1 x_2 \dots x_k y_1 y_2 \dots y_r$ deve appartenere a un sottoinsieme C di F^{k+r} che prende il nome di **codice**. Il meccanismo che realizza questo procedimento prende il nome di **codificatore**. Il codificatore spedisce i simboli d'informazione e i simboli di controllo lungo il canale il quale emette una parola $c' = x'_1 x'_2 \dots x'_k y'_1 y'_2 \dots y'_r$ che in generale non coinciderà con la parola c originariamente spedita. Scopo del **decodificatore** è di riuscire a ricostruire la parola originaria c a partire dalla parola ricevuta c' . In generale il decodificatore mediante una opportuna **regola di decodificazione** determinerà a partire da c' una parola $c'' = x''_1 x''_2 \dots x''_k y''_1 y''_2 \dots y''_r$ appartenente al codice C ; emetterà poi la sequenza $x''_1 x''_2 \dots x''_k$ come stima del blocco d'informazione originario. Di solito il decodificatore nella sua ricerca si basa su un criterio di **massima verosimiglianza**, cioè vuole che la parola c'' abbia la più alta probabilità "a posteriori" di essere stata immessa nel canale dopo che in uscita è stata osservata la parola c' . Per i canali più comunemente studiati — i cosiddetti canali "simmetrici" — ciò si attua facendo in modo che la parola di codice c'' differisca dalla parola ricevuta c' nel minor numero possibile di componenti; se chiamiamo **distanza di Hamming** di due parole a e b il numero $\rho(a, b)$ delle posizioni in cui esse hanno componenti diverse, allora possiamo dire che il decodificatore cerca una parola $c'' \in C$ per la quale risulti minima la distanza di Hamming $\rho(c', c'')$.

E' abbastanza chiaro a questo punto quali siano i parametri che misurano le qualità del nostro schema di codificazione-decodificazione. Se indichiamo con $n = k+r$ la **lunghezza del codice** C , il numero $R(C) = k/n$ si chiama **tasso d'informazione** del codice e misura il contenuto d'informazione utile che si trova mediamente in una parola di codice. Al contrario il rapporto r/n rappresenta la **ridondanza** del codice C , ossia l'informazione aggiuntiva che in un certo senso siamo stati costretti a sprecare con i simboli di controllo per migliorare la sicurezza della trasmissione. All'aumentare della ridondanza aumenta la necessità di memorizzare dati, diminuisce pertanto la velocità della trasmissione e quindi aumentano i costi. Se indichiamo con $d = d(C)$ la **distanza minima** del codice C , cioè il numero $d(C) = \min \{ \rho(c_1, c_2); c_1, c_2 \in C, c_1 \neq c_2 \}$, allora si definisce il **tasso di correzione** $\lambda(C)$ come il rapporto d/n . La regola di decodificazione precedentemente illustrata sarà in grado di correggere tanti più errori avvenuti nel canale quanto maggiore è la distanza minima del codice e quindi quanto maggiore è il suo tasso di correzione; quest'ultimo ci fornisce dunque una buona misura dell'affidabilità del sistema.

La contemporanea presenza di un alto tasso d'informazione e di un alto tasso di correzione in uno stesso codice è un evento che si verifica raramente: nelle famiglie di codici che si conoscono a tutt'oggi almeno uno dei due parametri tende a zero al divergere della lunghezza. Un famoso risultato di C.

Shannon [45] asserisce che in corrispondenza di un prefissato $\epsilon > 0$ è possibile trovare un intero $n(\epsilon)$ tale che per ogni $n > n(\epsilon)$ esiste un codice C di lunghezza n con il quale si ottiene, in presenza di un decodificatore a massima verosimiglianza, una probabilità di errori di trasmissione inferiore a ϵ . La natura di questo risultato è per l'appunto puramente esistenziale: a parte le informazioni sui suoi parametri, nulla conosciamo della struttura del codice C e quindi non sappiamo se sia effettivamente possibile trovare per esso delle regole di codificazione e decodificazione efficienti e realizzabili in pratica. Ci troviamo nella situazione conflittuale di dover garantire la sicurezza della trasmissione con apparecchiature dal costo accettabile.

Da questo punto di vista diventa allora importante avere a disposizione dei codici dotati di molte proprietà descrivibili in termini matematici. Un buon codice è generalmente ricco di "simmetrie" e ha una certa "omogeneità". Nei Paragrafi che seguono descriveremo una classe di codici lineari binari, i *codici di Reed-Muller*, che raggiunsero una certa fama all'inizio degli anni '70 quando furono utilizzati dalla NASA per equipaggiare le sue sonde spaziali. La *decodificazione a soglia o a maggioranza* introdotta per questi codici fu estesa ad altre classi di codici e ha il vantaggio di una semplice realizzazione strumentale, [37]. Ciò non significa necessariamente che le proprietà di questi codici utilizzate per la decodificazione siano semplici o comunque evidenti. In quest'ottica è molto istruttiva la lettura della prefazione del libro di Elwyn R. Berlekamp [12] che spiega come lo stesso fenomeno accada per un'altra importantissima classe di codici, i cosiddetti codici BCH.

6. LA SOMMA MISTA DI DUE CODICI LINEARI BINARI

Se n è un intero positivo indicheremo con $V_n(q)$ lo spazio vettoriale di tutte le n -uple a coefficienti nel campo finito $GF(q)$; ci interesserà particolarmente il caso binario $q = 2$: gli elementi di $GF(2)$ si chiamano **bit**. Se $\mathbf{x} \in V_n(q)$ è una parola arbitraria chiameremo **peso (di Hamming)** di $\mathbf{x}$ il numero $\gamma(\mathbf{x})$ delle componenti non nulle di $\mathbf{x}$. La distanza di Hamming su $V_n(q)$ è **invariante per traslazioni**, cioè per ogni $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V_n(q)$ si ha sempre $\rho(\mathbf{x}, \mathbf{y}) = \rho(\mathbf{x} + \mathbf{z}, \mathbf{y} + \mathbf{z})$. Ne segue che la distanza di Hamming $\rho(\mathbf{x}, \mathbf{y})$ di due qualunque parole $\mathbf{x}, \mathbf{y} \in V_n(q)$ è uguale al peso di Hamming $\gamma(\mathbf{x} - \mathbf{y})$ della loro differenza.

Un qualunque sottospazio vettoriale k -dimensionale C di $V_n(q)$ verrà chiamato (n, k) -**codice lineare su $GF(q)$** (nel caso $q = 2$ parleremo di codici lineari **binari**). Una matrice $k \times n$ a coefficienti in $GF(q)$ le cui righe formino una base del sottospazio vettoriale C si dirà una **matrice generatrice** di C . Si verifica subito che la distanza minima di un codice lineare è uguale al peso minimo delle sue parole non nulle.

Siano C_1 e C_2 due codici lineari binari di lunghezza n e di rispettive dimensioni k_1 e k_2 . La **somma mista** di C_1 e C_2 è definita da $C_1 \& C_2 := \{\mathbf{x}_1, \mathbf{x}_1 + \mathbf{x}_2; \mathbf{x}_1 \in C_1, \mathbf{x}_2 \in C_2\}$.

Se G_1 e G_2 sono rispettivamente matrici generatrici per C_1 e C_2 allora la matrice

$$G_1 \& G_2 = \left(\begin{array}{c|c} G_1 & G_1 \\ \hline 0 & G_2 \end{array} \right)$$

è una matrice generatrice della somma mista $C_1 \& C_2$. Non è difficile verificare che se C_1 ha distanza minima d_1 e C_2 ha distanza minima d_2 allora la somma mista $C_1 \& C_2$ è un $(2n, k_1 + k_2)$ -codice lineare binario con distanza minima $d := \min \{2d_1, d_2\}$.

Il tasso d'informazione $R(C_1 \& C_2) = (k_1 + k_2)/(2n)$ della somma mista $C_1 \& C_2$ si calcola come media aritmetica dei tassi d'informazione $R(C_1) = k_1/n$ di C_1 e $R(C_2) = k_2/n$ di C_2 . La somma mista è particolarmente promettente quando risulta $d_2 = 2d_1$. In tal caso infatti $C_1 \& C_2$ ha lo stesso tasso di correzione d_1/n di C_1 .

La somma mista si può naturalmente applicare anche a codici C_1 e C_2 di lunghezza diversa; in tal caso basta estendere il codice di lunghezza minore fino a raggiungere la lunghezza dell'altro, eventualmente aggiungendo brutalmente degli zeri nelle componenti mancanti.

7. DEFINIZIONE COMBINATORIA DEI CODICI DI REED-MULLER

Utilizziamo formalmente la stessa formula ricorsiva dei coefficienti binomiali abbinata alla somma mista per definire, in corrispondenza agli interi m, s con $0 \leq s \leq m$ il **codice** (binario) di **Reed-Muller** $RM(m, s)$ dell' s -esimo ordine e di lunghezza 2^m :

$$RM(m+1, s+1) := RM(m, s+1) \& RM(m, s) \quad \text{per } 0 \leq s \leq m.$$

Come valori iniziali per la ricorsione definiamo i codici di Reed-Muller

$$RM(m, 0) := \{00 \dots 0, 11 \dots 1\} \subseteq V_{2^m}(2), \quad RM(m, m) := V_{2^m}(2) \quad \text{per } m \geq 0.$$

I codici di Reed-Muller dello 0-esimo ordine sono codici ripetitivi, mentre il codice di Reed-Muller dell' m -esimo ordine e di lunghezza 2^m coincide con lo spazio vettoriale di tutte le 2^m -uple a componenti in $GF(2)$. Le distanze minime $d := d(RM(m, s))$ e le dimensioni $k := \dim(RM(m, s))$ si calcolano ricorsivamente a partire dai valori iniziali $d(RM(m, 0)) = \dim(RM(m, m)) = 2^m$, $d(RM(m, m)) = \dim(RM(m, 0)) = 1$, $m \geq 0$, mediante le formule ricorsive

$$d(RM(m+1, s+1)) = \min\{2d(RM(m, s+1)), d(RM(m, s))\}$$

e

$$\dim(RM(m+1, s+1)) = \dim(RM(m, s+1)) + \dim(RM(m, s))$$

per $0 \leq s \leq m$. Per induzione su m si riconoscono le relazioni

$$d = d(RM(m, s)) = 2^{m-s}, \quad k = \dim(RM(m, s)) = \sum_{i=0}^s \binom{m}{i}.$$

(Per la somma $\sum_{i=0}^s \binom{m}{i}$, $s < m$, non esiste un'espressione più semplice).

8. MATRICI GENERATRICI

Definiamo innanzitutto ricorsivamente per $m \geq 0$ una matrice generatrice da indicarsi con $G(m)$ del codice di Reed-Muller banale $RM(m, m) = V_n(2)$ dell' m -esimo ordine e lunghezza $n := 2^m$. L'unica matrice generatrice di $RM(0, 0)$ è la matrice $G(0) := (1)$ di tipo 1×1 . Per tutti gli $m \geq 0$ utilizziamo il metodo illustrato precedentemente per costruire ricorsivamente la matrice generatrice $G(m+1)$ del codice $RM(m+1, m+1) = RM(m, m) \& RM(m, m)$ a partire dalla matrice generatrice $G(m)$ del codice $RM(m, m)$:

$$G(m+1) := \left(\begin{array}{c|c} G(m) & G(m) \\ \hline 0 & G(m) \end{array} \right)$$

La matrice $G(m)$ è una matrice triangolare superiore di tipo $n \times n$. Associamo ora ordinatamente a ciascuna delle $n = 2^m$ righe di $G(m)$ uno degli n sottoinsiemi $\emptyset, \{0\}, \{1\}, \{1,0\}, \{2\}, \{2,0\}, \{2,1\}, \{2,1,0\}, \{3\}, \{3,0\}, \dots$ (ordinamento lessicografico) dell'insieme $M := \{0,1, \dots, m-1\}$; tale insieme verrà detto **l'insieme caratteristico** della riga. (Se non vi è pericolo di equivoci tralasciamo nella scrittura degli insiemi caratteristici le parentesi graffe “{” e “}” nonché le virgole, p.es. $\{3,2,0\} = 320$.) Se numeriamo le righe di $G(m)$ dall'alto verso il basso con gli interi $0, 1, \dots, n-1$, allora il numero assegnato alla riga avente l'insieme caratteristico $A \subseteq M = \{0,1, \dots, m-1\}$ risulta essere $\sum_{j \in A} 2^j$. Gli elementi di un insieme caratteristico indicano pertanto quelle posizioni nelle quali la rappresentazione binaria del numero di riga contiene un 1; p.es. numero di riga $13 = 2^3 + 2^2 + 2^0 = 1101$ (rapp. binaria), insieme caratteristico $\{3, 2, 0\}$. Per brevità indichiamo con $v(A)$ la riga cui è associato l'insieme caratteristico $A \subseteq M = \{0,1, \dots, m-1\}$, p.es. $v(320) = 000000000000101$.

Descriviamo le righe della matrice $G(m+1)$ ricorsivamente con l'aiuto delle righe della matrice $G(m)$, $m = 1, 2, \dots$; una riga di $G(m+1)$ il cui insieme caratteristico non contenga l'elemento m si ottiene scrivendo due volte di seguito la riga di $G(m)$ che pure possiede A come insieme caratteristico; una riga di $G(m+1)$ il cui insieme caratteristico contenga l'elemento m si ottiene scrivendo il vettore nullo 2^m -dimensionale seguito dalla riga di $G(m)$ avente l'insieme caratteristico $A \setminus \{m\}$.

Da questa descrizione ricorsiva si ricava subito la seguente descrizione non ricorsiva delle righe di $G(m)$. Determiniamo le componenti $x(0), x(1), \dots, x(n-1)$ della riga $v(A)$ di $G(m)$ di cui sia noto l'insieme caratteristico $A \subseteq M = \{0,1, \dots, m-1\}$. Per $j = m-1, m-2, \dots, 1, 0$ si ha che se $j \in A$ allora risulta

$$\begin{array}{llll} x(0) & = x(1) & = \dots = x(2^j - 1) & = 0 \\ x(2 \cdot 2^j) & = x(2 \cdot 2^j + 1) & = \dots = x(3 \cdot 2^j - 1) & = 0 \\ x(4 \cdot 2^j) & = x(4 \cdot 2^j + 1) & = \dots = x(5 \cdot 2^j - 1) & = 0 \\ x(6 \cdot 2^j) & = x(6 \cdot 2^j + 1) & = \dots = x(7 \cdot 2^j - 1) & = 0 \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \end{array}$$

Le componenti che alla fine non risultano ancora determinate hanno il valore 1. A parole: Se per un dato insieme caratteristico $A \subseteq M$ vogliamo determinare la corrispondente riga $v(A)$ di $G(m)$ procediamo come segue. Per ogni $j \in A$ poniamo una dopo l'altra uguali a zero le prime 2^j componenti, lasciamo indeterminate le successive 2^j componenti, poniamo di nuovo uguali a zero le

successive 2^j componenti, lasciamo indeterminate le successive 2^j componenti e così via. Alla fine poniamo uguali a 1 tutte le componenti che sono rimaste indeterminate.

Da questa descrizione non ricorsiva si riconosce che se $v(A)$ e $v(B)$ sono righe di $G(m)$, allora la riga $v(A \cup B)$ ha componente 1 esattamente nelle posizioni in cui sia $v(A)$ che $v(B)$ hanno componente 1. Grazie a questa semplice regola possiamo costruire una qualunque riga $v(A)$ di $G(m)$ a partire dalle righe $v(0)$, $v(1)$, ..., $v(m-1)$ mediante la relazione $v(A) = v(\cup_{j \in A} \{j\})$; pertanto la riga $v(320)$ ha componente 1 esattamente nelle posizioni in cui le righe $v(3)$, $v(2)$ e $v(0)$ hanno in comune la componente 1. Sfrutteremo questa proprietà per la caratterizzazione algebrica dei codici di Reed-Muller del Paragrafo 5.

Per $s = 0, 1, 2, \dots, m$ indichiamo ora con $G(m, s)$ la matrice che consiste di tutte le righe di $G(m)$ il cui insieme caratteristico ha al più s elementi. Dalla precedente descrizione ricorsiva delle righe di $G(m+1)$ possiamo dedurre una formula ricorsiva per le matrici $G(m, s)$: per tutti gli interi m, s con $0 \leq s < m$ risulta

$$G(m+1, s+1) := \left(\begin{array}{c|c} G(m, s+1) & G(m, s+1) \\ \hline 0 & G(m, s) \end{array} \right)$$

Per ogni intero $m \geq 0$ la matrice $G(m, m)$ coincide con la matrice generatrice $G(m)$ del codice $RM(m, m)$, mentre la matrice $G(m, 0)$ consiste del solo vettore 2^m -dimensionale $v(0) = \mathbf{1} = 11 \dots 1$ ed è pertanto una matrice generatrice del codice $RM(m, 0)$. Ne ricaviamo che $G(m+1, s+1)$ è una matrice generatrice della somma mista del codice generato da $G(m, s+1)$ e del codice generato da $G(m, s)$ (somma eseguita nell'ordine indicato) e quindi che la matrice $G(m, s)$ è una matrice generatrice del codice di Reed-Muller $RM(m, s)$ che chiameremo **matrice generatrice canonica** di $RM(m, s)$.

Richiamiamo l'attenzione su due proprietà dei codici di Reed-Muller $RM(m, s)$ che si ottengono direttamente dalla rappresentazione della loro matrice generatrice canonica come sottomatrice di $G(m)$ formata scegliendone opportunamente certe righe:

1. $RM(m, s) \subseteq RM(m, s+1)$ per $0 \leq s < m$,
2. $k = \dim(RM(m, s)) = \sum_{i=0}^s \binom{m}{i}$ per $0 \leq s \leq m$.

9. CARATTERIZZAZIONE ALGEBRICA

Siano m, s interi con $0 \leq s \leq m$ e sia $n := 2^m$. Definiamo in $V_n(2)$ il **prodotto**

per componenti

$$\mathbf{x}: \begin{cases} V_n(2) \times V_n(2) & \rightarrow & V_n(2); \\ (\mathbf{x} = x_1 x_2 \dots x_n, \mathbf{y} = y_1 y_2 \dots y_n) & \mapsto & \mathbf{x} \times \mathbf{y} = (x_1 y_1, x_2 y_2, \dots, x_n y_n) \end{cases}$$

Il prodotto per componenti $\mathbf{x} \times \mathbf{y}$ di due vettori $\mathbf{x}$ e $\mathbf{y}$ ha componente 1 esattamente nelle posizioni in cui sia $\mathbf{x}$ che $\mathbf{y}$ hanno componente 1. Per tutti i vettori $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V_n(2)$ e tutti gli scalari $\lambda \in GF(2)$ valgono banalmente le proprietà algebriche

$$\begin{aligned} (\mathbf{x} \times \mathbf{y}) \times \mathbf{z} &= \mathbf{x} \times (\mathbf{y} \times \mathbf{z}), & \mathbf{x} \times \mathbf{y} &= \mathbf{y} \times \mathbf{x}, & \mathbf{x} \times \mathbf{1} &= \mathbf{x}, \\ (\mathbf{x} + \mathbf{y}) \times \mathbf{z} &= \mathbf{x} \times \mathbf{z} + \mathbf{y} \times \mathbf{z}, & \lambda (\mathbf{x} \times \mathbf{y}) &= (\lambda \mathbf{x}) \times \mathbf{y} & \text{ e } & \mathbf{x} \times \mathbf{x} = \mathbf{x}. \end{aligned}$$

Dalla descrizione non ricorsiva delle righe di $G(m)$ del precedente Paragrafo ricaviamo che per ogni sottoinsieme $A = \{j_1, j_2, \dots, j_s\} \subseteq M$ con $|A| = s$ la riga di $G(m)$ avente A come insieme caratteristico può essere scritta come

$$\mathbf{v}(A) = \mathbf{v}(\cup_{i=1}^s \{j_i\}) = \mathbf{v}(j_1) \times \mathbf{v}(j_2) \times \dots \times \mathbf{v}(j_s),$$

p.es. $\mathbf{v}(320) = \mathbf{v}(3) \times \mathbf{v}(2) \times \mathbf{v}(0)$. Le righe di $G(m)$ aventi insiemi caratteristici con al più s elementi formano una base del codice di Reed-Muller $RM(m, s)$ dell' s -esimo ordine e lunghezza $n = 2^m$. Osserviamo che quando il prodotto per componenti è privo di fattori si ottiene il vettore $\mathbf{v}(0) = \mathbf{1} = 11\dots 1$. Abbiamo così la

CARATTERIZZAZIONE ALGEBRICA DEI CODICI DI REED-MULLER. *Il codice di Reed-Muller $RM(m, s)$ dell' s -esimo ordine consiste delle combinazioni lineari dei prodotti per componenti di al più s delle righe $\mathbf{v}(0), \mathbf{v}(1), \dots, \mathbf{v}(m-1)$ della matrice $G(m)$.*

Le righe $\mathbf{v}(0), \mathbf{v}(1), \dots, \mathbf{v}(m-1)$ formano una base di $RM(m, 1)$. Un semplice calcolo mostra che il prodotto per componenti di al più s vettori di $RM(m, 1)$ è una parola di $RM(m, s)$. Pertanto il codice $RM(m, s)$ si può caratterizzare anche come lo spazio delle combinazioni lineari dei prodotti per componenti di al più s parole di $RM(m, 1)$.

10. CARATTERIZZAZIONE GEOMETRICA

Per lo studio di certe proprietà dei codici di Reed-Muller è più appropriata una descrizione geometrica di quella combinatoria o quella algebrica. Ricapitoliamo innanzitutto alcuni concetti fondamentali della geometria.

Sia F un campo, sia m un intero positivo e sia $V_m(F)$ lo spazio vettoriale di tutte le m -uple a componenti in F . Per ogni vettore $\mathbf{a} \in V_m(F)$ la **traslazione**

$$\tau_{\mathbf{a}} : V_m(F) \rightarrow V_m(F); \quad \mathbf{x} \mapsto \tau_{\mathbf{a}}(\mathbf{x}) := \mathbf{x} + \mathbf{a}$$

è una biiezione. Le traslazioni di $V_m(F)$ rispetto alla composizione di applicazioni formano un gruppo isomorfo al gruppo additivo di $V_m(F)$. Sia ora $U \subseteq V_m(F)$ un sottospazio vettoriale di dimensione $s := \dim U$; un **traslato** di U è una classe laterale $\tau_{\mathbf{a}}(U) = \mathbf{a} + U$ dove $\mathbf{a}$ è un vettore di $V_m(F)$; i traslati di tutti i sottospazi vettoriali s -dimensionali di $V_m(F)$ si chiamano **sottospazi affini di dimensione s** di $V_m(F)$. I sottospazi affini di dimensione rispettivamente $0, 1, 2, m-2, m-1$ si chiamano rispettivamente **punti, rette, piani, iperrette, iperpiani**. Due sottospazi affini U, W della stessa dimensione si dicono **paralleli** e si scrive $U \parallel W$ se esiste una traslazione $\tau_{\mathbf{a}}$ di $V_m(F)$ tale che risulti $\tau_{\mathbf{a}}(U) = W$. Per ogni dato sottospazio vettoriale U e ogni punto $\mathbf{a} \in V_m(F)$ esiste esattamente un sottospazio affine W (della stessa dimensione di U) parallelo a U e tale che $\mathbf{a} \in W$, si tratta precisamente del sottospazio $W = \tau_{\mathbf{a}}(U) = \mathbf{a} + U$ (*postulato euclideo della parallela*). Due sottospazi affini paralleli o coincidono oppure sono disgiunti. I sottospazi affini paralleli a un dato sottospazio affine U formano una partizione di $V_m(F)$, cioè ricoprono tutto $V_m(F)$ senza sovrapporsi.

In un contesto geometrico, invece che dello spazio vettoriale $V_m(F)$ si parla dello **spazio affine** $AG_m(F)$ (se $F = GF(q)$ si scrive anche $AG_m(q)$) e si utilizzano terminologie geometriche come “un punto $\mathbf{a}$ è *incidente* un sottospazio U ” invece di “ $\mathbf{a} \in U$ ”, “due rette G e H si *intersecano* in un punto $\mathbf{a}$ ” invece di “ $\{\mathbf{a}\} = G \cap H$ ” e così via.

Comunque presi due punti distinti $\mathbf{a}, \mathbf{b} \in AG_m(F)$ esiste un'unica retta incidente entrambi, vale a dire $\{\mathbf{a} + \lambda(\mathbf{b} - \mathbf{a}); \lambda \in F\}$. Più in generale: comunque dati $s+1$ punti, tali che s qualunque di essi non giacciono mai in uno stesso sottospazio affine di dimensione $s-1$, esiste un unico sottospazio affine s -dimensionale incidente ciascuno di essi.

L'intersezione di una qualunque famiglia di sottospazi affini è ancora un sottospazio affine. L'intersezione di s iperpiani è vuota oppure è un sottospazio affine di dimensione almeno $(m-s)$. Ogni sottospazio affine $(m-s)$ -dimensionale si può rappresentare come intersezione di s opportuni iperpiani.

Un ben noto conteggio mostra che $V_m(q)$ contiene $\begin{bmatrix} m \\ s \end{bmatrix}_q := \prod_{i=0}^{s-1} (q^{m-i} - 1) / (q^{s-i} - 1)$ sottospazi vettoriali s -dimensionali; pertanto lo spazio affine $AG_m(q)$ contiene esattamente $q^{m-s} \begin{bmatrix} m \\ s \end{bmatrix}_q$ sottospazi affini s -dimensionali, in quanto ciascuno dei q^{m-s} traslati di un sottospazio vettoriale s -dimensionale di $V_m(q)$ è un sottospazio affine s -dimensionale di $AG_m(q)$.

Per la descrizione dei codici di Reed-Muller utilizziamo lo spazio affine m -dimensionale $AG_m(2)$ su $GF(2)$. Ogni sottospazio vettoriale $(m-1)$ -dimensionale di $V_m(2)$ consiste di 2^{m-1} vettori; l'insieme complementare di un iperpiano

in $AG_m(2)$ è dunque un iperpiano ad esso parallelo. La *differenza simmetrica* $(U \cup V) \setminus (U \cap V)$ di due iperpiani U e V di $AG_m(2)$ o è un iperpiano oppure è l'insieme vuoto (nel caso $U = V$) oppure è tutto lo spazio affine (nel caso $U \parallel W$ con $U \neq W$). In generale tuttavia la differenza simmetrica di due sottospazi affini non è un sottospazio affine.

Scriviamo ora in un ordine particolare gli $n = 2^m$ punti dello spazio affine $AG_m(2)$ come colonne $s(0), s(1), \dots, s(n-1)$ di una matrice $m \times n$ che denoteremo con $S(m)$: le righe di $S(m)$ devono risultare uguali ai complementi dei vettori $v(0), v(1), \dots, v(m-1)$ cioè i vettori $v^*(0) := v(0) + \mathbf{1}$, $v^*(1) := v(1) + \mathbf{1}$, $\dots$, $v^*(m-1) := v(m-1) + \mathbf{1}$. Fra l'altro per $j = 0, 1, \dots, n-1$ la colonna nr. j di $S(m)$ dà proprio la rappresentazione binaria del numero $n-j-1$.

Possiamo ora rappresentare un qualunque sottoinsieme X di $AG_m(2)$ mediante il suo **vettore d'incidenza** $x_0, x_1, \dots, x_{n-1} \in V_n(2)$: poniamo $x_j = 1$ se e soltanto se $s(j) \in X$. Se x, y sono rispettivamente i vettori d'incidenza dei sottoinsiemi X, Y di $AG_m(2)$ allora la somma $x + y$ e il prodotto per componenti $x \times y$ sono rispettivamente i vettori d'incidenza della differenza simmetrica $(X \cup Y) \setminus (X \cap Y)$ e dell'intersezione $X \cap Y$.

Per $i = 0, 1, \dots, m-1$ la riga $v(i)$ di $G(m)$ è proprio il vettore d'incidenza dell'insieme di tutti quei punti $s(j) = (s_{j,0}, s_{j,1}, \dots, s_{j,m-1})^T$ di $AG_m(2)$ per i quali risulti $s_{j,i} = 0$; in altre parole $v(i)$ è il vettore d'incidenza di un sottospazio vettoriale $(m-1)$ dimensionale $U(i)$ di $V_m(2)$, cioè di un iperpiano di $AG_m(2)$ contenente il vettore nullo $s(n-1) = \mathbf{0}^1$. Più in generale: per ogni insieme caratteristico $A \subseteq M := \{0, 1, \dots, m-1\}$ con $|A| = s$ la riga $v(A)$ di $G(m)$ è il vettore d'incidenza di quel sottospazio vettoriale $(m-s)$ -dimensionale $U(A)$ di $V_m(2)$ che contiene tutti i punti $s(j)$ tali che risulti $s_{j,i} = 0$ per ogni $i \in A$.

Comunque presi due insiemi caratteristici $A, B \subseteq M = \{0, 1, \dots, m-1\}$ risulta

- 1) $A \subseteq B \Leftrightarrow U(A) \supseteq U(B)$,
- 2) $U(A \cup B) = U(A) \cap U(B)$,
- 3) $U(A \cap B) = U(A) + U(B)$.

In particolare si ha $U(M) = \{\mathbf{0}\}$ e $U(\emptyset) = V_m(2)$. L'applicazione che ad ogni insieme caratteristico associa il sottospazio vettoriale $U(A)$ di $V_m(2)$ è dunque un anti-isomorfismo dal reticolo $\mathcal{P}(M)$ dei sottoinsiemi di M al sottoreticolo $\{U(A); A \subseteq M\}$ del reticolo dei sottospazi vettoriali di $V_m(2)$.

Il codice di Reed-Muller del prim'ordine $RM(m, 1)$ contiene assieme ai vettori d'incidenza $v(0), v(1), \dots, v(m-1)$ dei sottospazi vettoriali $(m-1)$ -dimensionali $U(0), U(1), \dots, U(m-1)$ di $V_m(2)$ anche i loro complementi cioè i vettori d'incidenza $v^*(0), v^*(1), \dots, v^*(m-1)$ degli iperpiani ad essi paralleli. A parte la parola $v(\emptyset) = \mathbf{1}$ e il vettore nullo $\mathbf{0}$, tutte le rimanenti $2(m-1)$ parole di $RM(m, 1)$, in quanto vettori d'incidenza di differenze simmetriche d'iperpiani, sono esse stesse vettori d'incidenza di iperpiani.

Lo spazio affine $AG_m(2)$ possiede esattamente $2 \cdot \binom{m}{m-1}_2 = 2(n-1)$ iperpiani; ne segue che il codice di Reed-Muller $RM(m, 1)$ consiste dei vettori d'incidenza di tutti gli iperpiani di $AG_m(2)$, del vettore d'incidenza $\mathbf{v}(0) = \mathbf{1}$ di tutto lo spazio $AG_m(2)$ e del vettore d'incidenza $\mathbf{0}$ dell'insieme vuoto $\emptyset$. Da questa descrizione delle parole del codice di Reed-Muller $RM(m, 1)$ e tenendo in mente la caratterizzazione algebrica si deduce che il codice $RM(m, s)$ contiene i vettori d'incidenza di tutte le intersezioni di al più s iperpiani di $AG_m(2)$.

CARATTERIZZAZIONE GEOMETRICA DEI CODICI DI REED-MULLER.
Numerando opportunamente i punti di $AG_m(2)$, il codice di Reed-Muller $RM(m, s)$ consiste dei vettori d'incidenza di tutti i sottospazi affini almeno $(m-s)$ -dimensionali di $AG_m(2)$ e delle loro differenze simmetriche.

11. DECODIFICAZIONE A MAGGIORANZA

Presentiamo qui la decodificazione a maggioranza del codice di Reed-Muller $RM(m, s)$ dell' s -esimo ordine e di lunghezza $n := 2^m$, introdotta da I.S. Reed nel 1954.

Il codificatore prende il blocco d'informazione

$$\mathbf{a}(s) = a_0 a_1 a_{1,0} a_{2,0} a_{2,1} \dots a_{m-1, m-2, \dots, m-s}$$

costituito da $k := \sum_{i=0}^s \binom{m}{i}$ bit a_A (che indiciamo con gli insiemi caratteristici $A \subseteq M := \{0, 1, \dots, m-1\}$, $|A| \leq s$, ordinati lessicograficamente come nel Paragrafo 5) e lo codifica mediante la matrice generatrice canonica $G(m, s)$ nella parola

$$\mathbf{x}(s) := x_{s,0} x_{s,1} \dots x_{s,n-1} := \sum_{A \subseteq M: |A| \leq s} a_A \mathbf{v}(A)$$

di $RM(m, s)$ e la manda al canale. Qui la parola $\mathbf{x}(s)$ viene (eventualmente) disturbata in alcune componenti e viene trasformata in una parola $\mathbf{y}(s) = y_{s,0} y_{s,1} \dots y_{s,n-1}$ a partire dalla quale il decodificatore dovrebbe riuscire a ricostruire il blocco d'informazione originario $\mathbf{a}(s)$. Il decodificatore è strutturato in $s+1$ stadi concatenati. Il primo stadio si chiama stadio nr. s , il secondo si chiama stadio nr. $s-1$, ..., lo stadio $(s+1)$ -esimo si chiama stadio nr. 0. Nello stadio nr. $t = s, s-1, \dots, 0$ il decodificatore calcola per ogni insieme caratteristico $A \subseteq M$ con $|A| = t$ un bit b_A che rappresenta la sua stima del bit d'informazione originale a_A . Quando la procedura di decodificazione ha superato l'ultimo stadio nr. 0, il decodificatore consegna al ricevitore la sequenza di bit

$$\mathbf{b} = b_0 b_1 b_{1,0} b_2 b_{2,0} b_{2,1} \dots b_{m-1, m-2, \dots, m-s}$$

che costituisce la sua stima del blocco d'informazione $\mathbf{a}(s)$.

Ciascuno stadio nr. $t = s, s-1, \dots, 0$ eredita dallo stadio precedente nr. $t+1$ (il primo passo nr. $t = s$ eredita dal canale) una parola $\mathbf{y}(t) = y_{t,0} y_{t,1} \dots y_{t,n-1}$, calcola a partire dalle sue componenti i bit b_A con $A \subseteq M$, $|A| = t$ (come questo calcolo si svolga effettivamente verrà spiegato più avanti) e consegna la parola

$$\mathbf{y}(t-1) = y_{t-1,0} y_{t-1,1} \dots y_{t-1,n-1} := \mathbf{y}(t) - \sum_{\substack{A \subseteq M: \\ |A|=t}} b_A \mathbf{v}(A)$$

al successivo stadio nr. $t-1$ (con l'eccezione dell'ultimo stadio nr. $t=0$).

La parola $\mathbf{y}(s)$ è una versione disturbata della parola $\mathbf{x}(s)$ del codice $RM(m, s)$. Più in generale possiamo interpretare la parola $\mathbf{y}(t)$ per $t = s, s-1, \dots, 0$ come una versione disturbata della parola

$$\mathbf{x}(t) := \mathbf{x}(s) - \sum_{\substack{A \subseteq M: \\ t < |A| \leq s}} a_A \mathbf{v}(A) = \sum_{\substack{A \subseteq M: \\ |A| \leq t}} a_A \mathbf{v}(A)$$

del codice di Reed-Muller $RM(m, t)$ del t -esimo ordine; i bit b_A sono in effetti l'approssimazione dei bit d'informazione a_A . Possiamo scrivere la parola $\mathbf{x}(t) \in RM(m, t)$ nella forma $\mathbf{a}(t) \cdot G(m, t)$ dove il blocco $\mathbf{a}(t) = a_0 a_1 \dots a_{m-1, m-2, \dots, m-t}$ è una parte di lunghezza $\sum_{j=0}^t \binom{m}{j}$ del blocco d'informazione $\mathbf{a}(s)$. Il lavoro dello stadio nr. t del nostro $RM(m, s)$ -decodificatore non si differenzia per nulla dal lavoro del primo stadio (che è pure lo stadio nr. t) di un $RM(m, t)$ -decodificatore in un sistema di comunicazione che utilizzi il codice di Reed-Muller $RM(m, t)$ del t -esimo ordine e nel quale il codificatore codifichi il blocco d'informazione $\mathbf{a}(t)$ nella parola $\mathbf{x}(t) \in RM(m, t)$ ottenendo la parola $\mathbf{y}(t)$ in uscita dal canale.

Esaminiamo ora come vengono ricostruiti i bit b_A in un passo nr. t del decodificatore a partire dalle componenti $y_{t,0}, y_{t,1}, \dots, y_{t,n-1}$ della parola $\mathbf{y}(t)$.

Sia $A \subseteq M$ uno degli $\binom{m}{t}$ insiemi caratteristici con $|A| = t$, sia $A' = M \setminus A$ il suo insieme complementare e sia $U(A)$ risp. $U(A')$ il sottospazio vettoriale $(m-t)$ -dimensionale risp. t -dimensionale di $V_m(2)$ dato dal vettore d'incidenza $\mathbf{v}(A)$ risp. $\mathbf{v}(A')$. Indichiamo con $U_1(A'), U_2(A'), \dots, U_{2^{m-t}}(A')$ i 2^{m-t} sottospazi affini di $AG_m(2)$ paralleli a $U(A')$. Per $i = 1, 2, \dots, 2^{m-t}$ il decodificatore calcola la somma $b_A(i)$ di tutte le componenti $y_{t,j}$ di $\mathbf{y}(t)$ per le quali il punto $s(j)$ appartiene al sottospazio affine $U_i(A')$:

$$b_A(i) := \sum_{\substack{0 \leq j \leq n-1: \\ s(j) \in U_i(A')}} y_{t,j}; \quad i = 1, 2, \dots, 2^{m-t}.$$

Dopo aver calcolato tutte queste 2^{m-t} somme di controllo il decodificatore prende una *decisione di maggioranza*: quando il numero delle somme di controllo $b_A(i)$ che hanno il valore 0 raggiunge o supera il *valore di soglia* $2^{m-t}/2 = 2^{m-t-1}$ allora il decodificatore pone $b_A := 0$, altrimenti pone $b_A := 1$. (La decisione per il bit 0 nel caso di un pareggio è del tutto arbitraria; talora in questi casi di patta un decodificatore un po' più complicato è in grado di prendere decisioni migliori).

Il codice $RM(m, s)$ ha distanza minima $d = 2^{m-s}$. L'algoritmo di decodificazione di Reed qui descritto garantisce una decodificazione corretta, cioè $\mathbf{b} = \mathbf{a}(s)$, se la parola di codice $\mathbf{x}(s)$ è stata disturbata in meno di $d/2$ componenti, cioè se la distanza di Hamming $\rho(\mathbf{x}(s), \mathbf{y}(s))$ risulta minore di 2^{m-s-1} . A tale scopo dimostriamo innanzitutto che per ogni insieme caratteristico $A \subseteq M$ di cardinalità t e ogni indice $i = 1, 2, \dots, 2^{m-t}$ si ha sempre

$$a_A = \sum_{\substack{0 \leq j \leq n-1; \\ s(j) \in U_i(A)}} x_{t,j}.$$

In primo luogo abbiamo

$$\sum_{\substack{0 \leq j \leq n-1; \\ s(j) \in U_i(A)}} x_{t,j} = \sum_{\substack{0 \leq j \leq n-1; \\ s(j) \in U_i(A)}} \sum_{\substack{B \subseteq M; |B| \leq t; \\ s(j) \in U(B)}} a_B = \sum_{\substack{B \subseteq M; \\ |B| \leq t}} \sum_{\substack{0 \leq j \leq n-1 \\ s(j) \in U_i(A) \cap U(B)}} a_B.$$

Dal momento che stiamo calcolando in $GF(2)$, per tutti i sottoinsiemi caratteristici $B \subseteq M$ con $|B| \leq t$ per i quali il sottospazio affine $U_i(A) \cap U(B)$ contenga un numero pari di punti, cioè per i quali risulti $\dim(U_i(A) \cap U(B)) \neq 0$, abbiamo la relazione

$$\sum_{\substack{0 \leq j \leq n-1 \\ s(j) \in U_i(A) \cap U(B)}} a_B = 0$$

L'intersezione del sottospazio affine t -dimensionale $U_i(A)$ col sottospazio vettoriale almeno $(m-t)$ -dimensionale $U(B)$ consiste di un unico punto se e soltanto se i due insiemi caratteristici A e B sono l'uno il complementare dell'altro, cioè se risulta $A = B$. Dalla relazione

$$\sum_{\substack{0 \leq j \leq n-1 \\ s(j) \in U_i(A) \cap U(A)}} a_A = a_A$$

segue l'asserto.

Se nel canale la parola di codice $\mathbf{x}(s)$ è stata disturbata in un numero e di posizioni minore di 2^{m-s-1} , allora il decodificatore nel primo stadio nr. s del calcolo dei bit b_A con $|A| = s$ prende sempre la decisione esatta, cioè risulta $b_A = a_A$. Pertanto anche la parola $\mathbf{y}(s-1)$ differisce dalla parola di codice $\mathbf{x}(s-1)$ in esat-

tamente e posizioni. Poiché e a maggior ragione risulta minore del valore di soglia 2^{m-s} del secondo stadio nr. $s-1$, allo stesso modo il decodificatore calcolando i bit b_A con $|A| = s-1$ determina sempre correttamente i bit d'informazione $b_A = a_A$. E così di seguito.

12 . UN ESEMPIO

Illustriamo l'algoritmo di decodificazione di Reed nel caso del (32, 6)-codice di Reed-Muller RM (5, 1).

Il codificatore prende un blocco d'informazione $\mathbf{a}(1) = a_0 a_1 a_2 a_3 a_4$ costituito da sei bit e lo codifica nella parola di codice

$$\mathbf{x}(1) = x_{1,0} x_{1,1} \dots x_{1,31} = a_0 \mathbf{v}(\emptyset) + \sum_{j=0}^4 a_j \mathbf{v}(j).$$

Nel canale questa parola viene trasformata nella parola $\mathbf{y}(1) = y_{1,0} y_{1,1} \dots y_{1,31}$. Il decodificatore inizia il suo lavoro:

1.0 se almeno 8 delle 16 somme di controllo

$$y_{1,0} + y_{1,1}, \quad y_{1,2} + y_{1,3}, \quad \dots, \quad y_{1,30} + y_{1,31}$$

hanno il valore 0 si ponga $b_0 := 0$, altrimenti si ponga $b_0 := 1$;

1.1 se almeno 8 delle 16 somme di controllo

$$y_{1,0} + y_{1,2}, \quad y_{1,1} + y_{1,3}, \quad \dots, \quad y_{1,29} + y_{1,31}$$

hanno il valore 0 si ponga $b_1 := 0$, altrimenti si ponga $b_1 := 1$

1.2 se almeno 8 delle 16 somme di controllo

$$y_{1,0} + y_{1,4}, \quad y_{1,1} + y_{1,5}, \quad \dots, \quad y_{1,27} + y_{1,31}$$

hanno il valore 0 si ponga $b_2 := 0$, altrimenti si ponga $b_2 := 1$;

1.3 se almeno 8 delle 16 somme di controllo

$$y_{1,0} + y_{1,8}, \quad y_{1,1} + y_{1,9}, \quad \dots, \quad y_{1,23} + y_{1,31}$$

hanno il valore 0, si ponga $b_3 := 0$, altrimenti si ponga $b_3 := 1$;

1.4 se almeno 8 delle 16 somme di controllo

$$y_{1,0} + y_{1,16}, \quad y_{1,1} + y_{1,17}, \quad \dots, \quad y_{1,15} + y_{1,31}$$

hanno il valore 0 si ponga $b_4 := 0$, altrimenti si ponga $b_4 := 1$.

Si costruisca la parola $\mathbf{y}(0) := y_{0,0} y_{0,1} \dots y_{0,31} := \mathbf{y}(1) - \sum_{j=0}^4 b_j \mathbf{v}(j)$.

2. Se almeno 16 dei 32 bit

$$y_{0,0}, y_{0,1}, \dots, y_{0,31}$$

hanno il valore 0 si ponga $b_5 := 0$, altrimenti si ponga $b_5 := 1$.
Emettere il blocco $b_0 b_1 b_2 b_3 b_4$.

BIBLIOGRAFIA

1. E. Ambrisi, F. Eugeni, *Il problema della protezione dell'informazione I: cenni storici e metodi statistici; per la decrittazione di sistemi di cifratura classici*, Ratio Math. **1** (1990), 1-29.
2. E. Angelieri, "Trasmissione dati", Editoriale Delfino, Milano, 1972.
3. L. Berardi, *A new secret code based on Steiner system*, Preprint (1989).
4. L. Berardi, *Some remarks about an Electronic Signature Derived from a Generalized RSA-code*, J. of Info. & Opt. Sci. **11** (1990), 189-194.
5. L. Berardi, A. Beutelspacher, *I buoni angeli custodi, ovvero i protettori di un messaggio*, Archimede **2-3** (1989), 129-140.
6. L. Berardi, M. Di Fonso, *Protezione dell'informazione su personal computer*, Atti del I Simposio Nazionale su stato e prospettive della ricerca crittografica in Italia (SPRCI) (1987), Fondazione Bordini.
7. L. Berardi, M. Di Fonso, F. Eugeni, *Threshold Schemes based on Criss-Cross Block Designs*, J. of Info. & Opt. Sci. **11** (1990), 153-160.
8. L. Berardi, F. Eugeni, *Strutture geometriche, crittografia e sistemi di sicurezza richiedenti un quorum*, Atti del I Simposio Nazionale su stato e prospettive della ricerca crittografica in Italia, SPRCI (1987), 127-133, Fondazione Bordini.
9. L. Berardi, F. Eugeni, *Smart Cards and Authentication schemes*, 5-th Meeting Euro Working Group on Financial Modelling, Acireale (CT) 1989. Rend. di Mat., Roma. In corso di stampa.
10. L. Berardi, F. Eugeni, *Blocking sets e teoria dei giochi*, Atti Sem. Mat. Fis. Univ. Modena **34** (1988), 165-196.
11. L. Berardi, B. Rizzi, *Generalizziamo il codice RSA e la Funzione di Eulero*, Atti del I Simposio Nazionale su stato e prospettive della ricerca crittografica in Italia (SPRCI) (1987), Fondazione Bordini.
12. E.R. Berlekamp, "Algebraic Coding Theory", McGraw-Hill, New York, 1968.
13. A. Beutelspacher, *La scuola elementare della teoria dei codici*, Quad. n.1 (suppl. didatt.) Sem. Geom. Comb. L'Aquila (1985).

14. A. Beutelspacher, "Kryptologie", Vieweg-Verlag, Braunschweig & Wiesbaden, 1988.
15. A. Beutelspacher, *Enciphered Geometry. Some applications of Geometry to Cryptography*, Ann. Disc. Math. **37** (1988), 59-68.
16. A. Beutelspacher, *Perfect and essentially perfect authentication systems*, in "Advances in Cryptology - EUROCRYPT '87", Lecture Notes Comp. Science, 304, D. Chaum, W.L. Price (Eds.), Springer, Berlin, 1988, pp. 167-170.
17. A. Beutelspacher, *How to say "no"*, Preprint (1989).
18. A. Beutelspacher, F. Eugeni, *Geometrie finite e Crittosistemi: stato dell'arte e problematiche*, Atti del II Simposio Nazionale su stato e prospettive della ricerca crittografica in Italia (SPRCI) (1989), Fondazione Bordini.
19. A. Beutelspacher, U. Rosenbaum, *Geometric authentication systems*, Ratio Math. **1** (1990), 30-41.
20. A. Beutelspacher, G. Tallini, C. Zanella, *Examples of essentially s -fold secure geometric authentication system with large s* , Rend. di Mat., Roma. In corso di stampa.
21. A. Beutelspacher, K. Vedder, *Geometric structures as threshold schemes*, IMA, Proceedings Cirencester, Cambridge Univ. Press. In corso di stampa.
22. G.R. Blakley, *Safeguarding Cryptographic keys*, Proc. NCC **48** (1979), 313-317, AFIPS Press, Montvale, NJ.
23. O. Brugia, *Sistemi crittografici a chiave pubblica*, (Quad. n.4673 (1985), Scuola Superiore "G. Reiss Romoli", L'Aquila.
24. O. Brugia, S. Improta, W. Wolfowicz, *Segretezza ed autenticazione nelle moderne reti di comunicazione*, Rel. Int. 2B 63385 (1985), Fondazione Bordini, Roma.
25. M. Cerasoli, F. Eugeni, M. Protasi, "Matematica Discreta", Zanichelli, Bologna, 1988.
26. M. De Soete, K. Vedder, *Some new classes of geometric threshold schemes*, in "Advances in Cryptology-EUROCRYPT '88", Lecture Notes Comp. Science, Springer, Berlin.
27. W. Diffie, M.E. Hellman, *New directions in Cryptography*, IEEE Trans. Inform. Theory **22** (1976), 644-654.
28. E.N. Gilbert, F.J. MacWilliams, N.J.A. Sloane, *Codes which detect deception*, Bell Syst. Techn. J. **53** (1974), 405-424.
29. M. Gionfriddo, *Sui sistemi di Steiner*, Quad. n.9 Sem. Geom. Comb. Univ. L'Aquila (1986), 3-28.
30. M. Gionfriddo, F. Eugeni, *On the minimum number of blocks of a maximal partial spread in $STS(v)$ and $SQS(v)$* , J. of Geometry **36** (1989), 37-48.
31. W. Heise, P. Quattrocchi, "Informations- und Codierungstheorie," Springer, Berlin, 1988.
32. W. Heise, P. Quattrocchi, *Teoria algebrica dei codici*, Boll. Un. Mat. Ital. (7) **1-A** (1987), 313-331.

33. G. Longo, "Teoria dell'informazione," Boringhieri, Torino, 1980.
34. G. Longo, *Introduzione alla teoria dei codici algebrici*, Archimede **34-3** (1982), 99-107.
35. L. Lunelli, "Teoria dell'informazione e dei codici," Clup, Milano, 1981.
36. F.J. MacWilliams, N.J.A. Sloane, "The theory of error-correcting codes", North-Holland, Amsterdam, 1977.
37. J.L. Massey, "Threshold decoding", MIT Press, Cambridge, MA, 1963.
38. A. Maturo, *Messaggi cifrati per mezzo di numeri pseudo-causali ottenuti a partire da successioni in algebre di supporto*, Atti del I Simposio Nazionale su stato e prospettive della ricerca crittografica in Italia, SPRCI (1987), Fondazione Bordini.
39. D.E. Muller, *Application of Boolean algebra to switching circuit design and to error detection*, IEEE Trans. Computers **3** (1954), 6-12.
40. I.S. Reed, *A class of multiple-error-correcting codes and the decoding scheme*, IEEE Trans. Inform. Theory **4** (1954), 38-49.
41. A. Rizzi, *Alcune considerazioni sulla crittografia*, Atti del I Simposio Nazionale su stato e prospettive della ricerca crittografica in Italia, SPRCI (1987), Fondazione Bordini.
42. L. Sacco, "Manuale di Crittografia," 1947. Ristampa Anastatica.
43. A. Sgarro, "Crittografia", Muzzio, Padova, 1986.
44. A. Shamir, *How to share a secret*, Communications ACM **11** (1979), 612-613.
45. C.E. Shannon, *A Mathematical Theory of communication*, Bell Syst. Techn. J. **27** (1948), 379-423.
46. C.E. Shannon, *Communication Theory of Secrecy Systems*, Bell Syst. Techn. J. **28** (1949), 656-715.

A. Bonisoli, Istituto di Matematica, Università della Basilicata, via N. Sauro 85, I-85100 Potenza. F. Eugeni, Dipartimento di Scienze e Storia dell'Architettura, Università "G. D'Annunzio", viale Pindaro 42, I-65127 Pescara.

ANALISI DELLA BONTÀ DI ALCUNI GENERATORI DI NUMERI PSEUDOCASUALI PER LA CIFRATURA DEI MESSAGGI E LA SIMULAZIONE (*)

Nisida Cera (**) Antonio Maturo (**)

1. INTRODUZIONE

Indichiamo con il termine “messaggio” sia una comunicazione inviata da un individuo ad altri sia un promemoria messo da parte da una persona per poterlo consultare in un secondo tempo.

Utilizzando un computer, un messaggio del primo tipo consiste in un documento inviato attraverso un canale elettronico; un messaggio del secondo tipo, consiste in un file di dati immessi in una opportuna area di memoria e richiamabili per mezzo di un programma.

In entrambi i casi, un messaggio si presenta come una successione finita $M = \{m_i\}_{i \in \{1, 2, \dots, h\}}$ di cifre in una certa base b . Il numero h si dice **lunghezza**, in base b , del messaggio.

In genere si assume $b=2$ oppure $b=10$.

Si può porre il problema di voler rendere accessibile un messaggio solo a particolari individui e non ad altri.

A tale scopo, si considera una seconda successione finita $K = \{k_i\}_{i \in \{1, 2, \dots, h\}}$ di simboli in base b , anch'essa di lunghezza h , conosciuta solo dalle persone destinatarie del messaggio e, al posto di M , s'invia il messaggio $C=M+K$, definito come la successione di termine generale

(*) Lavoro svolto nell'ambito della ricerca M.P.I. 40% 1988 Modelli Economici Discreti, Ottimizzazione, Crittografia.

(**) Dipartimento di Scienze e Storia dell'Architettura, Università degli Studi “G. D'Annunzio” viale Pindaro 42, 65127 PESCARA.

$$c_i = (m_i + k_i) \mod b, \quad i=1,2,\dots,h.$$

Poiché $(c_i - k_i) \mod b = m_i$, un individuo che conosce **K** può risalire, dalla conoscenza di **C**, al messaggio **M**.

Il procedimento consistente nel sostituire **C** ad **M** si dice di **cifratura** o **codifica** del messaggio, quello che fa risalire ad **M** a partire da **C** si dice di **decodifica**.

La successione **K** si dice **chiave del messaggio**.

Due esigenze fondamentali a cui deve soddisfare la chiave **K** sono:

(1) **segretezza** o **sicurezza**: dev'essere massimizzata la difficoltà di risalire a **K** dalla conoscenza di **C**;

(2) **semplicità**: dev'essere massimizzata la semplicità nella formulazione della chiave e nei procedimenti di codifica e decodifica.

Le due esigenze sono solitamente contrastanti e la scelta della chiave da utilizzare va fatta tenendo conto dei problemi che ci si pongono e delle circostanze in cui si opera.

Si può capire ciò ricorrendo ad una analogia con la vita comune.

Alcuni oggetti che ci interessano, in casa o in ufficio, vengono chiusi a chiave, con una chiave comune, in un cassetto.

Il sistema chiave comune - cassetto non offre certamente un grado di sicurezza pari ad esempio a quello del sistema formato da una cassaforte.

Tuttavia si usa il sistema chiave comune - cassetto poiché, per l'importanza degli oggetti ed una valutazione obiettiva dei rischi, si ritiene sufficiente la sicurezza ottenuta e soddisfacente la semplicità di acquisizione e di uso di tale sistema di sicurezza.

Un procedimento di codifica che ha il massimo requisito di sicurezza è quello di porre **K** uguale ad una **successione casuale** di lunghezza h di cifre in una data base b .

Esso, tuttavia, non ha il requisito della semplicità, in quanto, se h è elevato, ad esempio superiore a 50, è certamente un'operazione molto lunga ottenere la successione **K** ed è difficoltoso trasmetterla e utilizzarla per operazioni di codifica e decodifica.

Si ottiene, invece, il massimo requisito di semplicità, paragonabile a quella del sistema chiave comune - cassetto, ponendo **K** uguale ad un segmento h di una **successione pseudocasuale** di cifre in base b .

Per quanto riguarda il requisito della sicurezza, esso dipende dal tipo della successione pseudocasuale utilizzata, precisamente dalle sue **proprietà matematiche e statistiche**.

2. SUCCESSIONI OTTENUTE DA GENERATORI CONGRUENZIALI

Sia m un intero positivo maggiore o uguale a 2 ed indichiamo con $I(m)$ l'insieme $\{0, 1, \dots, m-1\}$.

Introduciamo, in $I(m)$, le operazioni $(+)$ e $(\cdot)$ ponendo, per definizione, per ogni coppia (x, y) di elementi di $I(m)$,

$$x (+) y = (x + y) \bmod m, \quad x (\cdot) y = (x \cdot y) \bmod m.$$

È facile verificare che la terna $(I(m), (+), (\cdot))$ è un anello, isomorfo all'anello $(Z_m, +, \cdot)$ degli interi modulo m .

Sia n un intero positivo ed indichiamo con $[I(m)]^n$ l'insieme delle n -ple di elementi di $I(m)$. Chiamiamo, per semplicità, **vettori** tali n -ple e **scalari** gli elementi di $I(m)$.

Per ogni elemento x di $[I(m)]^n$, sia $(x)_i$ la sua componente i -esima.

Introduciamo, in $[I(m)]^n$, le operazioni $(+)$ e $*$ ponendo, per definizione, per ogni coppia (x, y) di elementi di $[I(m)]^n$ e per ogni $a \in I(m)$,

$$(x (+) y)_i = (x)_i + (y)_i, \quad (a * x)_i = a (\cdot) (x)_i.$$

È facile verificare che $([I(m)]^n, (+), *)$ è un modulo su $I(m)$.

È noto che, se m è primo, $I(m)$ è un campo e quindi $[I(m)]^n$ è uno **spazio vettoriale** su $I(m)$.

2.1 Definizione Siano n ed s interi positivi ed m un intero maggiore di 1. Diciamo **generatore congruenziale di dimensione n , ordine s e modulo m** ogni **funzione**

$$g: \{[I(m)]^n\}^s \longrightarrow I(m) \quad (2.1)$$

non costante rispetto al vettore prima coordinata. ***

Fissando s vettori $x_0, x_1, \dots, x_{s-1}$, a partire dalla (2.1), ponendo, per ogni $h \geq s$,

$$x_h = g(x_{h-s}, x_{h-s+1}, \dots, x_{h-1}), \quad (2.2)$$

si ottiene una successione $X = \{x_h\}$, con $h \in N_0$, di elementi di $[I(m)]^n$.

2.2 Definizione Diciamo **funzione di supporto** in $[I(m)]^n$ ogni **epimorfismo** Φ fra gli $I(m)$ -moduli $[I(m)]^n$ e $I(m)$. (Considerando $I(m)$ modulo su $I(m)$) ***

2.3 Definizione Diciamo **successione in $I(m)$** associata al generatore g , alla

funzione di supporto Φ e alla scelta dei vettori iniziali $x_0, x_1, \dots, x_{s-1}$, la successione Y di termine generale $y_h = \Phi(x_h)$, con $h \in \mathbb{N}_0$. ***

Se la successione Y soddisfa in maniera "sufficiente" ad alcuni requisiti matematici e statistici, essa può essere assunta come atta a simulare una successione casuale e si dice **successione pseudocasuale**.

Da una successione pseudocasuale si può ottenere una chiave K formata da k cifre in una data base b con vari criteri, di cui i più comuni sono i seguenti:

(1) caso $m = b$

Fissato un elemento y_a qualsiasi di Y si assume:

$$K = \{y_a, y_{a+1}, \dots, y_{a+k-1}\}.$$

(2) caso $m = b^s$, con s intero maggiore di 1.

Ogni elemento y_h si può scrivere con s cifre in base b .

Allora, fissato un elemento y_a qualsiasi di Y e, detti q ed r il quoziente ed il resto della divisione di k per s , si assume K uguale alla successione delle cifre, in base b , di $y_a, y_{a+1}, \dots, y_{a+q-1}$, nell'ordine, e delle prime r cifre, sempre in base b , y_{a+q} .

(3) caso $m = b^{1/s}$, con s intero maggiore di 1.

Ogni elemento di K si ottiene scrivendo in base b un numero avente come cifre, in base m , s elementi consecutivi di Y .

Fissando un elemento y_a qualsiasi di Y , si assume K uguale alla successione delle cifre in base b ottenute scrivendo, in base b , i numeri già scritti in base m nella forma

$$y_{a+rs} y_{a+rs+1} \dots y_{a+rs+s-1}$$

per $r = 0, 1, 2, \dots, k-1$.

(4) caso in cui non valgono le condizioni precedenti o comunque si vogliono ottenere numeri con s cifre in una data base fissata b e risulta $m > b^s$.

Sostituiamo ad ogni elemento y_h della successione Y il numero

$$u_h = [y_h b^s / m], \quad (2.3)$$

indicando, per ogni numero reale x , con $[x]$ la sua parte intera nella base b

fissata.

Ogni u_h si scrive con s cifre in base b , per cui ci si riconduce al caso (2), sostituendo le u_h alle y_h .

In pratica, specie nell'ultimo caso, è opportuno scegliere le basi $b=2$ oppure $b=10$. Per altre scelte della base, dato che i calcolatori usano o il sistema binario o quello decimale, la (2.3) può presentare errori di arrotondamento o troncamento, per cui per essere sicuri che sono verificate le necessarie proprietà matematiche e statistiche deve essere preso s in modo che sia $m \gg b^s$. Ciò però ha come conseguenza una notevole perdita nella velocità di acquisizione delle cifre.

3. PROPRIETÀ MATEMATICHE E STATISTICHE DEI GENERATORI CONGRUENZIALI LINEARI

Un generatore congruenziale g di dimensione n , ordine s e modulo m si dice **lineare** se la (2.2) assume la forma

$$x_h = a_0 * x_{h-s} (+) a_1 * x_{h-s+1} (+) \dots (+) a_{s-1} * x_{h-1} (+) b, \quad (3.1)$$

con $a_0 \neq 0$, $a_1, a_2, \dots, a_{s-1}$ elementi di $\mathbf{I}(m)$ e b elemento di $[\mathbf{I}(m)]^s$.

I generatori congruenziali più studiati nella letteratura e più usati sia per la loro semplicità e sia per la possibilità di individuarne con chiarezza alcune fondamentali proprietà matematiche e statistiche sono quelli lineari di dimensione 1 e di ordine 1. Essi assumono la forma

$$x_h = a * x_{h-1} (+) b, \quad (3.2)$$

con $a \neq 0$ e b elementi di $\mathbf{I}(m)$.

Limitiamoci a considerare il caso particolare dei generatori (3.2) con la funzione di supporto Φ uguale alla funzione identica in $\mathbf{I}(m)$. Allora la successione in $\mathbf{I}(m)$ associata al generatore (3.2) e alla Φ assume la forma:

$$y_h = (a y_{h-1} + b) \mod m. \quad (3.3)$$

Usando la terminologia classica, diversa da quella usata in questo lavoro, in cui vengono introdotte le funzioni g e Φ , chiamiamo, da ora in poi, “**generatore lineare**” senza ulteriori specificazioni il generatore (3.3) e “**generatore moltiplicativo**” un generatore lineare con $b=0$.

Rinviando ai lavori di A. Rizzi [10], A. Maturo [6], N. Cera e A. Maturo [7], [8] e altri (cfr. ad es. [1], [11], [12]), per uno studio dettagliato delle proprietà “a priori” dei generatori lineari e per le dimostrazioni, riportiamo di seguito

alcuni dei risultati più significativi.

Sia $\{y_h\}$, con $h \in \mathbb{N}_0$, la successione associata alla (3.3) e alla scelta di un fissato valore iniziale y_0 .

3.1 Risultato La successione in $\mathbf{I(m)}$ $\{y_h\}$, con $h \in \mathbb{N}_0$, è periodica. Se μ è l'antiperiodo e δ è il periodo, risulta $\mu + \delta \leq m$. ***

3.2 Risultato La successione in $\mathbf{I(m)}$ $\{y_h\}$, con $h \in \mathbb{N}_0$, ha il massimo periodo possibile $\delta = m$ se e solo se sono soddisfatte le seguenti condizioni:

(L1) M.C.D.(a,m) = 1;

(L2) p numero primo divisore di m $\Rightarrow a \bmod p = 1$;

(L3) 4 divisore di m $\Rightarrow a \bmod 4 = 1$;

(L4) M.C.D.(b,m) = 1. ***

Osserviamo che, se $\delta = m$, la successione di termine generale y_h è, a meno di una traslazione, indipendente dal valore iniziale y_0 . Inoltre, se $p_1, p_2, \dots, p_s$ sono i fattori primi distinti di m le (L1), (L2), (L3), si possono esprimere scrivendo:

$$a \bmod (p_1 \cdot p_2 \cdots p_s) = 1, \text{ se } 4 \text{ non divide } m; \quad (3.4)$$

$$a \bmod (2p_1 \cdot p_2 \cdots p_s) = 1, \text{ se } 4 \text{ divide } m. \quad (3.5)$$

La (L4) non è una condizione molto restrittiva. Invece le (L1), (L2), (L3), conducono a successioni molto "regolari" e quindi senza "apparenza di casualità" se m è primo oppure è il prodotto di 2^r , con $r \in \{0, 1, 2\}$ e di numeri primi dispari distinti. Infatti, in tali circostanze le nostre condizioni implicano che $a = 1$.

In vari lavori (ad es. [4], [6], [10]) è stata messa in evidenza l'opportunità che, ai fini di ottenere buone proprietà statistiche di casualità, nella scomposizione in fattori primi di m, compaiano non più di due fattori primi, ciascuno con un esponente "abbastanza grande" (ad es. non inferiore ad 8).

Si ritiene inoltre opportuno che tali fattori primi sia piccoli, preferibilmente formati da una sola cifra decimale.

Fra i moduli che soddisfano a tali condizioni vi sono $m=2^r$ e $m=10^r$, fondamentali per l'invio di messaggi con cifre, rispettivamente, in base 2 ed in base 10.

Il generatore più semplice è quello moltiplicativo. Esso, dato che non è soddisfatta la (L4), non dà luogo a successioni di periodo m. Tuttavia, nei lavori citati, è stato dimostrato che si possono ottenere, con tale generatore, successioni con periodi di lunghezze accettabili e con "apparenza di casualità" spesso superiore a quella dei generatori lineari di massimo periodo.

Alcuni dei risultati più significativi relativi al generatore moltiplicativo

sono i seguenti:

3.3 Risultato Sia $m = p^r$ con $r \geq 1$ e p numero primo. La successione $\{y_h\}$, $h \in \mathbb{N}_0$, in $\mathbf{I}(m)$, associata al generatore

$$y_h = a y_{h-1} \pmod{m} \quad (3.6)$$

è periodica e priva di antiperiodo. Si ottiene il massimo periodo $\delta = m(p-1)/p$ se e solo se sono verificate le seguenti condizioni:

(M1) a è radice primitiva modulo m ;

(M2) $\text{M.C.D.}(y, m) = 1$. ***

3.4 Risultato Sia $m = 2^r$, con $r > 3$. La successione in $\mathbf{I}(m)$ associata al generatore (3.6) è periodica e priva di anti-periodo. Il periodo massimo ottenibile è $\delta = m/4$ e si ha se e solo se valgono le seguenti condizioni:

(P1) $a \pmod{8} = 3$ oppure $a \pmod{8} = 5$;

(P2) y_0 è dispari. ***

Se m non è una potenza di un numero primo, le condizioni per avere il massimo periodo sono più complesse ed il massimo periodo ottenibile è molto inferiore ad m .

Per il caso in cui $m = 10^r$ si ha il seguente

3.5 Risultato Sia $m = 10^r$, con $r > 4$. La successione in $\mathbf{I}(m)$ associata al generatore (3.6) è periodica e priva di antiperiodo. Si ottiene il periodo massimo $\delta = m/20$ se e solo se valgono le seguenti condizioni:

(D1) $a \pmod{200}$ è uguale ad uno dei seguenti 32 numeri:

3, 11, 13, 19, 21, 27, 29, 37, 53, 59, 61, 67, 69, 77, 83, 91, 109, 117, 123, 131, 133, 139, 141, 147, 163, 171, 173, 179, 181, 187, 189, 197;

(D2) y_0 è primo con 10. ***

Osserviamo che, nei casi in cui $m = 2^r$ o $m = 10^r$ c'è una vistosa diminuzione del periodo massimo rispetto a quello dei generatori lineari non moltiplicativi. Tuttavia il periodo non è l'unico elemento da considerare, nè il più importante. Una volta che ci siamo assicurati che il periodo è sufficientemente lungo, almeno dell'ordine di 10^8 - 10^{10} , dobbiamo preoccuparci soprattutto che venga-

no soddisfatte le proprietà statistiche di “apparenza di casualità” e della velocità di elaborazione. Quest’ultima è massima se m è una potenza della base b in cui sono considerate le cifre nel computer utilizzato.

4. ANALISI STATISTICHE

Consideriamo una successione $\mathbf{Y} = \{y_h\}$, con $h \in \mathbb{N}_0$, di elementi di $\mathbf{I}(m)$ ottenuta da un qualsiasi generatore congruenziale modulo m .

Indichiamo con $\mathbf{Z} = \{z_h\}$, con $h \in \mathbb{N}_0$, la successione di termini generale $z_h = y_h/m$ e con $\mathbf{U} = \{u_h\}$, con $h \in \mathbb{N}_0$, una successione di cifre in una data base b ottenuta, con uno dei criteri descritti nel paragrafo 2, a partire dalla $\mathbf{Y}$.

Sia r un intero positivo assegnato. Si ammette che $\mathbf{Z}$ sia una **successione di numeri pseudocasuali rispetto ai campioni di ampiezza r e rispetto a dei fissati test statistici** se, per un generico segmento $\{z_{h+1}, z_{h+2}, \dots, z_{h+r}\}$, di lunghezza r della successione, i test statistici dati portano ad “accettare” o meglio a “non rifiutare” l’ipotesi nulla:

H_0 : $\{z_{h+1}, z_{h+2}, \dots, z_{h+r}\}$ è un campione casuale della variabile casuale UC, uniforme continua in $[0,1]$.

Si ammette che $\mathbf{U}$ sia una **successione di cifre pseudocasuali in base b rispetto ai campioni di ampiezza r e rispetto a dei fissati test statistici** se, per un generico segmento $\{u_{h+1}, u_{h+2}, \dots, u_{h+r}\}$, di lunghezza r della successione, i test statistici dati portano ad “accettare” l’ipotesi nulla:

H_1 : $\{u_{h+1}, u_{h+2}, \dots, u_{h+r}\}$ è un campione casuale della variabile casuale UD (b), uniforme discreta di parametro b .

A seconda del valore dell’ampiezza r possiamo distinguere le analisi statistiche per la verifica della casualità delle successioni $\mathbf{Z}$ ed $\mathbf{U}$ in

- (i) analisi “globali”, per r uguale al periodo δ della successione considerata;
- (ii) analisi campionarie, per r “molto minore” del periodo δ , ma “abbastanza grande” perché il campione sia significativo.

Abbiamo eseguito sia delle analisi di tipo (i), sia delle analisi di tipo (ii) per vari tipi di generatori.

Per quanto riguarda le analisi globali, basandoci sulle teorie svolte nei lavori di Coveyou-Macpherson [3], Knuth [5], Maturo [9], abbiamo elaborato un complesso programma in basic per la verifica della casualità delle succes-

sioni. Gli output più significativi sono, per ogni successione, dei numeri positivi σ_j , $j \in \{1, 2, \dots, 7\}$, indicanti il "grado di vicinanza" della distribuzione delle j -ple di elementi consecutivi della successione da una distribuzione uniforme j -dimensionale.

Precisamente, tenuto conto delle considerazioni svolte nei lavori citati, abbiamo considerato tale grado di vicinanza

- (a) buono, per $\sigma_j \geq 1$;
- (b) sufficiente, per $0.1 < \sigma_j < 1$;
- (c) insufficiente, per $\sigma_j \leq 0.1$.

Abbiamo inoltre giudicato l'importanza di avere un alto grado di vicinanza decrescente al crescere di j . In pratica, per $j > 4$ le informazioni date dai valori σ_j sono di modesta rilevanza.

Riportiamo i risultati ottenuti per 9 generatori particolari nella tabella della pagina seguente. In essa non sono riportati i valori di σ_1 , sempre buoni.

I 9 generatori sono stati scelti per confrontare le varie e spesso discordanti proposte di autorevoli matematici e statistici. Precisamente:

- (1) per i generatori 1, 5, 8, 9 si è fatto in modo che a sia dell'ordine di grandezza di $[(\sqrt{5} - 1)/8]m$;
- (2) per i generatori 2 e 6 si è seguita l'indicazione di Greenberger di porre $a \approx \sqrt{m}$;
- (3) per i generatori 3 e 7 sono stati sperimentati moltiplicatori vicini, rispettivamente, al doppio e alla metà di quelli relativi al primo criterio;
- (4) per il generatore 4 abbiamo voluto sperimentare il comportamento con il modulo 10^{10} di un moltiplicatore che aveva dato un buon esito con il modulo 2^{35} .

Tabella dei valori σ_j per $j=2, 3, \dots, 7$.

n° d'ordine	term. b	modulo m	multipl.a	σ_2	σ_3	σ_4	σ_5	σ_6	σ_7
(b,m)=1	1	10^{10}	1545084981	3.14	0.32	0.26	1.30	1.07	1.32
(b,m)=1	2	10^{10}	100001	3.14	6.2E-9	7.9E-9	1.7E-8	3.3E-8	6.0E-8
(b,m)=1	3	10^{10}	3141592421	2.68	0.34	0.54	0.59	1.42	0.75
(b,m)=1	4	10^{10}	5308871541	2.82	2.84	0.09	3.82	1.53	0.56
(b,m)=1	5	2^{35}	5308871541	3.30	2.93	0.69	1.39	2.44	2.54
(b,m)=1	6	2^{35}	185365	3.14	1.1E-4	0.01	1.24	3.27	1.69
(b,m)=1	7	2^{35}	2718281821	2.59	1.16	1.75	4.25	2.26	0.74
b=0	8	$10^{11}-23$	28 ⁷	0.74	1.02	0.41	0.66	2.28	0.05
b=0	9	2^{35}	5308871551	1.25	4.44	1.86	2.93	1.56	1.69

Dalla tabella della pagina precedente si vede che, dal punto di vista delle proprietà statistiche globali, i generatori 2 e 6 appaiono piuttosto scadenti, soprattutto per quanto riguarda il comportamento delle terne e delle quaterne. Dal punto di vista crittografico ciò fa pensare che si presentino delle “regolarità” che permettono ad un esperto di decodificare facilmente un messaggio che usa tali generatori come chiave.

Il generatore 4 presenta qualche debolezza solo per il comportamento delle quaterne, mentre gli altri appaiono tutti sufficienti. Particolarmente buoni appaiono i generatori 7 e 9, il primo con m uguale ad una potenza di 2 e il cui moltiplicatore a non è fra quelli consigliati nella letteratura, ma è uno di quelli sperimentati da noi con il criterio (3) ed il secondo moltiplicativo, con m numero primo.

Un “buon comportamento globale” di un generatore di numeri pseudocasuali non garantisce però che non ci siano delle “regolarità” nei segmenti brevi della successione. Calcolando una chiave ottenuta da 200 numeri per un messaggio relativamente breve, corrispondente (per i valori considerati di m) all’emissione di circa 6000-7000 cifre binarie ed una chiave di 1000 numeri pseudocasuali per un messaggio più lungo, abbiamo quindi ritenuto necessario eseguire delle analisi statistiche campionarie per $r=200$ ed $r=1000$ sui generatori risultati sufficienti nell’analisi globale.

A tale scopo abbiamo elaborato un programma in basic con il quale sottoporre ad 11 test statistici un qualsiasi generatore lineare ed abbiamo fatto girare il programma, per ognuno dei 9 generatori considerati, valutando i risultati, ai livelli di significatività del 10% e del 1% con 200 replicazioni per $r=200$ e con 100 replicazioni per $r=1000$.

I test utilizzati, la cui descrizione dettagliata si può trovare nei nostri lavori [6], [8] ed in [10], sono i seguenti

1. Test sulla media dei numeri;
2. Test sulla varianza dei numeri;
3. Test sulla frequenza dei numeri;
4. Test sulla frequenza delle cifre;
5. Test sulla frequenza delle coppie di numeri;
6. Test sulla frequenza delle coppie di cifre;
7. Test sulla correlazione fra i numeri;
8. Test sulla correlazione fra le cifre;
9. Run tests;
10. Gap tests;
11. Tests sulle sequenze complete.

In generale, la logica seguita per tali test è la seguente:
data una successione $W = \{w_n\}$ di numeri reali ed assegnata una variabile

casuale X , ci poniamo il problema di verificare se il segmento generico $\{w_h, w_{h+1}, \dots, w_{h+r-1}\}$ di lunghezza r di W soddisfa l'ipotesi nulla:

$H: \{w_h, w_{h+1}, \dots, w_{h+r-1}\}$ è un campione casuale di ampiezza r di X .

I metodi di controllo, non parametrici, considerati in questo lavoro si basano essenzialmente nel seguente tipo di procedimento:

- (1) si fissa una statistica $D = D(X_1, X_2, \dots, X_r)$ funzione del campione casuale $X = (X_1, X_2, \dots, X_r)$ di ampiezza r di X tale che $D \geq 0$. La D dev'essere costruita in modo da assumere, generalmente, valori tanto più piccoli, quanto più è "accettabile" l'ipotesi H e si dice **misura della discrepanza o discrepanza** fra un generico segmento di lunghezza r della successione W e l'ipotesi H ;
- (2) si fissa un numero reale $\delta > 0$ e, detto d il valore assunto da D per $X = (w_h, w_{h+1}, \dots, w_{h+r-1})$, se $d \geq \delta$ si ritiene eccessiva la discrepanza e quindi si giudica "non accettabile" H ; se, invece, $d < \delta$ si ritiene "accettabile" la discrepanza e quindi H .

Il numero δ è assegnato in modo da soddisfare una uguaglianza del tipo

$$\text{prob}(D \geq \delta) = \alpha,$$

con α numero reale positivo, di solito minore o uguale a 0.1.

Nel nostro lavoro sono state considerate 11 diverse misure di discrepanza, di cui 6 associate a statistiche riguardanti i numeri pseudocasuali e 5 a statistiche collegate alla distribuzione delle cifre di tali numeri. Prevalentemente, abbiamo assunto come discrepanze o i valori di distribuzioni di tipo chi-quadro oppure le distanze, in valore assoluto, dei valori di distribuzioni simmetriche dalla media.

La notevole massa dei dati raccolti ci ha imposto di cercare un "test riassuntivo" per i risultati ottenuti, per ogni test, nelle varie replicazioni. Ciò allo scopo di confrontare, rispetto a ciascun test, la bontà dei generatori considerati. Abbiamo ritenuto opportuno, per tali motivi, modificare il procedimento di accettazione o rifiuto dell'ipotesi nulla, nella convinzione che valori eccessivamente bassi della discrepanza siano in disaccordo con l'ipotesi di indipendenza ed equidistribuzione così come quelli molto alti.

Abbiamo allora considerato due numeri σ e τ non negativi tali che:

$$\text{prob}(D \geq \sigma) = 0.1; \text{prob}(D \geq \tau) = 0.01$$

e diviso l'insieme dei valori possibili di D nei tre intervalli

$$I = [0, \sigma), J = [\sigma, \tau), L = [\tau, +\infty).$$

Dati m campioni casuali di ampiezza r della variabile casuale X , abbiamo

ritenuto di poterci attendere, in condizioni ideali, in media, che circa 0.9 m volte D assuma un valore d appartenente ad I, 0.09 m volte un valore in J e 0.01 m volte un valore in L.

Abbiamo allora pensato di valutare la discrepanza dei generatori dall'ipotesi H, rispetto ad un dato test ripetuto m volte, con un "test riassuntivo" di tipo chi-quadro applicato alla distribuzione dei valori assunti da D sulla partizione {I, J, L} di $[0, +\infty)$.

Abbiamo quindi ottenuto, per ciascuno dei nove generatori considerati, 11 tests riassuntivi per $r=200$ ed 11 per $r=1000$.

L'analisi campionaria ha confermato, in linea di massima, i risultati ottenuti con l'analisi globale. Alcuni esiti degni di rilievo sono i seguenti:

- i generatori con modulo 10^{10} risultati sufficienti nell'analisi globale risultano sufficienti, ma non eccezionali, nell'analisi campionaria. In particolare, nei test sulla frequenza delle coppie delle cifre e, in misura inferiore, nei test sulla frequenza delle cifre i valori della discrepanza assumono con una frequenza troppo elevata valori bassi.
- i generatori con modulo 2^{35} apparsi accettabili nell'analisi globale, risultano buoni dal punto di vista delle analisi campionarie. L'unica notazione degna di rilievo appare la percentuale relativamente elevata, del 3%, di esiti negativi, al livello di significatività dell'1%, del generatore 5 nel test delle varianze.
- per i generatori con modulo primo l'analisi campionaria dà dei buoni risultati, con qualche eccezione per i gap tests. Per qualche cifra, infatti, la percentuale di esiti negativi, al livello di significatività dell'1%, supera il 3%. Si arriva ad un massimo di esiti negativi del 4.5% per la cifra 7 nel generatore 9.

BIBLIOGRAFIA

1. J.H.Ahrens-U.Dieter-A.Grube, *Pseudo-random numbers. A new proposal for the choice of multipliers*, Computing 6°. 1970
2. G.Ascoli, *Lezioni di Algebra*, Ed. Tirrenia. Torino. 1965
3. R.R.Coveyou-R.Macpherson, *Fourier Analysis of uniform random generators*, Journal of the Association for Computing Machinery 14°. 1967
4. M.Greenberger, *An a priori determination of serial correlation in Computer generated random numbers*, Mathematics of Computation 15. 1961
5. D.E.Knuth, *The art of computer programming*, Vol:2. Seminumerical Algorithms. Addison-Wesley. London. 1969
6. A.Maturo, *Numeri pseudocasuali*, Montefeltro Ed. Urbino. 1982
7. A.Maturo-N.Cera, *Generazione di numeri pseudocasuali per mezzo di relazioni di ricorrenza su campi di Galois*, Facoltà di Architettura. Pescara. 1983
8. A.Maturo-N.Cera, *Confronto fra alcuni generatori di numeri pseudocasuali*,

Facoltà di Architettura, Pescara. 1983

9. A. Maturo, *Analisi di Fourier di successioni di numeri pseudocasuali: considerazioni e proposte*, Facoltà di Architettura. 1983
10. A. Rizzi, *Generazione di distribuzioni statistiche mediante un elaboratore elettronico*, Istituto di Statistica e ricerca sociale G. Gini. Roma. 1977
11. C. S. Smith, *Multiplicative pseudo-random numbers generators with prime modulus*, Journal of the Association for Computing Machinery. 18°. 1971
12. R. C. Tausworthe, *Random numbers generated by linear recurrence modulo two*, Mathematics of Computation 19°. 1965

CONFRONTO FRA INFERENZA BAYESIANA E TEST CLASSICI PER L'ANALISI DI SUCCESSIONI DI NUMERI PSEUDOCASUALI

Giuseppe Di Biase e Antonio Maturo (*)

Sommario - L'impostazione classica della verifica della casualità di successioni di numeri pseudocasuali si propone essenzialmente di scoprire se, ammettendo che l'ipotesi nulla H_0 sia vera, il verificarsi di una successione sia un evento "poco probabile" o "molto probabile". Nel primo caso si è "propensi a rinunciare" all'idea iniziale che H_0 sia vera, nel secondo caso, invece, l'idea viene confermata.

L'impostazione bayesiana, invece, fondata soprattutto su una visione soggettiva del Calcolo delle Probabilità si propone di confrontare H_0 con altre ipotesi possibili e di valutare le probabilità di tali ipotesi sulla base della conoscenza della successione.

Tuttavia, se ci si limita ad effettuare l'analisi bayesiana usuale non si ottengono risultati soddisfacenti. A nostro parere ciò dipende dal fatto che essa può essere considerata sostitutiva del solo test classico delle frequenze e, quindi, come quest'ultimo fornisce delle condizioni necessarie, ma del tutto insufficienti per la verifica di casualità delle successioni. Il lavoro si propone, quindi, di mettere a punto le idee di base per analizzare le successioni pseudocasuali da un punto di vista bayesiano.

Abstract - Classical approach for the check of the "randomness" of a sequence of pseudorandom numbers essentially proposes to evaluate if, in the hypothesis that H_0 (null hypothesis) is true, the occurrence of a sequence is a "not very probable" or a "very probable" result. In the former case it is "inclined to give up" the initial hypothesis that H_0 is true, whereas in the latter case the hypothesis is confirmed. Bayesian approach, on the contrary, is above all based on a subjective view of the Probability Theory, and it proposes to compare H_0 with other possible hypotheses on the basis of the knowledge of the sequence. However, if we limit ourselves to perform the usual Bayesian analysis we shall not obtain satisfying results. In our opinion this is due to the fact that these tests are substitutes of the only frequency test and, then, like this, they give necessary conditions but not sufficient for the check of "randomness" of a sequence. So in this paper we will consider to restate the basic ideas of Bayesian analysis of pseudo-random sequences.

(*) Dipartimento di Scienze, Storia dell'Architettura e Restauro
Viale Pindaro, 42 - 65127 Pescara

1. Successioni di numeri pseudocasuali e loro proprietà matematiche.

Sia N l'insieme dei numeri naturali maggiori di zero e sia $\{X_i\}_{i \in N}$ una successione di variabili casuali somiglianti ed indipendenti con distribuzione uniforme continua, ossia con densità di probabilità

$$f(x) = \begin{cases} 1 & \text{per } x \in (0,1) \\ 0 & \text{per } x \notin [0,1] \end{cases} \quad (1.1)$$

Per ogni $i \in N$, se x_i è una determinazione di X_i , la successione $\{x_i\}_{i \in N}$ si dice successione di numeri casuali.

Sia m un elemento di N «abbastanza grande», di solito dell'ordine di grandezza di almeno 10^8 .

Per esigenze di calcolo, dovute al fatto che gli elaboratori usano solo numeri razionali con un dato numero di cifre decimali, generalmente al posto delle X_i , si considerano delle variabili casuali $X_i^{(m)}$, $i \in N$ somiglianti ed indipendenti con distribuzione uniforme discreta normalizzata di parametro m , ossia tali che assumono i valori $0, 1/m, 2/m, \dots, (m-1)/m$, ciascuno con probabilità $1/m$.

La possibilità di sostituire le $X_i^{(m)}$ alle X_i , pur essendo variabili casuali di tipo diverso, le une discrete, le altre continue, si basa sul fatto che per ogni $i \in N$ la successione di variabili casuali $\{X_i^{(m)}\}_{m \in N}$ converge in distribuzione verso la variabile casuale X_i .

Infatti, basta osservare che le $X_i^{(m)}$ hanno la funzione caratteristica

$$H_m(u) = (1/m) (e^{iu} - 1) / (e^{iu/m} - 1)$$

e che

$$\lim_{m \rightarrow +\infty} H_m(u) = (e^{iu} - 1) / iu,$$

funzione caratteristica della variabile casuale uniforme continua.

In pratica una volta fissato m , la successione dei valori assunti dalle $X_i^{(m)}$, $i \in N$, viene sostituita da una successione $\{x_i\}_{i \in N}$ di valori ottenuti, per mezzo dell'elaboratore, con il seguente procedimento.

A partire da una formula ricorrente del tipo

$$y_{i+1} = f(\alpha_1, \alpha_2, \dots, \alpha_n; y_i), \quad i \in N, \quad (1.2)$$

con f funzione della variabile y_i , definita nell'insieme $I_m = \{0, 1, 2, \dots, m-1\}$, a valori in I_m e dipendente dai parametri $\alpha_1, \alpha_2, \dots, \alpha_n$, si ottiene una successione

$\{y_i\}_{i \in \mathbb{N}}$ di elementi dell'insieme I_m .

Successivamente, posto $x_i = y_i/m$ si ottiene una successione di numeri appartenenti al supporto $\{0, 1/m, \dots, (m-1)/m\}$ delle $X_i^{(m)}$.

Questo procedimento può apparire del tutto fuori di ogni logica. Infatti, invece di numeri a caso si considerano numeri, detti pseudocasuali, ottenuti, in maniera deterministica, da una formula.

Una giustificazione di tale procedimento segue dalla definizione soggettiva di numeri pseudocasuali (cfr. ad esempio Bruno Baldessari (1987), [1], Alfredo Rizzi (1989), [13]) secondo la quale la successione $\{x_i\}_{i \in \mathbb{N}}$ si dice pseudocasuale se

- (1) un operatore, al quale il generatore (1.2) sia noto, può univocamente generare la successione;
- (2) un individuo, al quale il generatore non sia noto, ritiene soggettivamente che la successione $\{x_i\}_{i \in \mathbb{N}}$ sia una realizzazione di una successione $\{X_i\}_{i \in \mathbb{N}}$ di variabili casuali uniformi discrete normalizzate di parametro m .

Infatti, se la funzione f e i suoi parametri sono assegnati in maniera opportuna il modo in cui si ottiene y_{i+1} a partire da y_i appare il prodotto di un «effetto di casualità» analogo a quello che si ottiene mescolando un mazzo di carte o ruotando la pallina di una roulette. Pensandoci bene, anche in questi casi i percorsi compiuti dalle carte o dalla pallina sono deterministici e la casualità consiste nel fatto che un osservatore non li conosce.

Nella formula (1.2), per ottenere successioni che soddisfino la definizione soggettiva di Baldessari, è necessario che l'operatore che genera la successione conosca bene le proprietà matematiche e statistiche di essa per fare in modo che appaia casuale ad un osservatore.

Ciò si ottiene assegnando f uguale ad una funzione il più possibile semplice. Il caso più utilizzato è quello in cui la (1.2) è del tipo

$$y_{i+1} = (a y_i + b) \bmod m, \quad i \in \mathbb{N}, \quad (1.3)$$

con a e b elementi di I_m .

Le proprietà matematiche e statistiche della successione $S = \{y_i\}_{i \in \mathbb{N}}$ che si ottiene dal generatore (1.3) variano notevolmente al variare dei parametri a , b e del valore iniziale (o seme) y_1 .

Per vari motivi (cfr. ad esempio Cera e Maturo (1983), [2] e [3], Maturo (1989), [8]), le scelte più opportune di m sono $m=10^s$, $m=2^s$ con $s \in \mathbb{N}$ e m numero primo.

Si dimostra che la successione S è periodica con periodo δ non superiore ad m .

Si dimostra, inoltre, (cfr. Cera e Maturo (1983) e (1990), [2] e [4], Maturo (1989), [8]) che, qualunque sia m , si ottiene il massimo periodo possibile $\delta=m$

se e solo se sono soddisfatte le seguenti condizioni

- (a) $\text{MCD}(a, m) = 1$,
- (b) p numero primo divisore di $m \Rightarrow a \bmod p = 1$,
- (c) 4 divisore di $m \Rightarrow a \bmod 4 = 1$,
- (d) $\text{MCD}(b, m) = 1$.

Queste condizioni non assicurano, però, la validità delle proprietà statistiche di indipendenza ed equidistribuzione. Ad esempio, per $a=1$ e $b=1$, esse sono soddisfatte qualunque sia m , ma la successione ottenuta appare tutt'altro che casuale.

Spesso è più conveniente, sia dal punto di vista statistico, sia perchè è particolarmente semplice verificare quali proprietà valgono, porre $b=0$. Si ottiene allora la seguente formula, detta generatore moltiplicativo,

$$y_{i+1} = (a y_i) \bmod m, \quad i \in \mathbb{N}. \quad (1.4)$$

Si può dimostrare che la successione S ottenuta a partire dal generatore moltiplicativo ha periodo δ minore o uguale ad un numero $\sigma(m)$ dipendente da m . Il valore $\sigma(m)$ è uguale ad $m-1$ per m numero primo, mentre è sensibilmente inferiore ad m negli altri casi. Ad esempio, $\sigma(m)=m/4$ per $m=2^s$, $s \geq 4$ e $\sigma(m)=m/20$ per $m=10^s$, $s \geq 5$. Si può inoltre dimostrare (cfr., ad es. Cera e Maturo (1983) e (1990), [2] e [4], Maturo (1989), [8]), che nei casi considerati, condizioni necessarie e sufficienti perchè sia $\delta=\sigma(m)$ sono le seguenti

- (I) caso m primo
- (IA) a e y_1 sono entrambi primi con m ,
- (IB) $a^i \bmod m \neq 1, \forall i \in \{1, 2, \dots, m-2\}$.

- (II) caso $m = 2^s$, con $s \geq 4$.
- (IIA) $a \bmod 8 = 3$ oppure $a \bmod 8 = 5$,
- (IIB) y_1 è dispari.

- (III) caso $m = 10^s$, con $s \geq 5$.
- (IIIA) y_1 è primo con 10,
- (IIIB) $a \bmod 200$ è uguale ad uno dei seguenti 32 numeri

3, 11, 13, 19, 21, 27, 29, 37, 53, 59, 61, 67, 69, 77, 83, 91, 109,
117, 123, 131, 133, 139, 141, 147, 163, 171, 173, 179, 181, 187, 189, 197.

2. Sull'analisi statistica dei generatori.

Consideriamo una successione $\{y_i\}_{i \in \mathbb{N}}$ di elementi di I_m ottenuti da una formula di tipo (1.2) e sia $x_i = y_i/m$.

Da un punto di vista classico la verifica della casualità della successione viene svolta nella maniera seguente.

Fissato un $n \in \mathbb{N}$ ci si pone il problema di poter accettare o meno l'ipotesi nulla H_0 che $\{x_1, x_2, \dots, x_n\}$ è un campione casuale della variabile casuale uniforme continua.

Si sottopone, a tale scopo, la successione finita $\{x_i\}_{i \in \{1, 2, \dots, n\}}$ di lunghezza n ad un insieme di test statistici.

Alcuni dei test più utilizzati, la cui descrizione dettagliata si può trovare nei lavori di Knuth (1969), [7], Maturo (1989), [11], Tausworthe (1965), [18], sono i seguenti

1. Test sulla media dei numeri;
2. Test sulla varianza dei numeri;
3. Test sulla frequenza dei numeri;
4. Test sulla frequenza delle cifre;
5. Test sulla frequenza delle coppie di numeri;
6. Test sulla frequenza delle coppie di cifre;
7. Test sulla correlazione fra i numeri;
8. Test sulla correlazione fra le cifre;
9. Run test;
10. Gap test;
11. Test sulle sequenze complete.

I metodi di controllo, non parametrici, considerati ad esempio in Cera e Maturo (1983) e (1990), [2] e [4], Maturo (1989), [8], si basano essenzialmente sul seguente tipo di procedimento:

(1) si fissa una statistica $D = D(X_1, X_2, \dots, X_n)$ funzione del campione casuale $\underline{X} = (X_1, X_2, \dots, X_n)$ di ampiezza n della variabile casuale X con distribuzione uniforme continua tale che $D \geq 0$. La D deve essere costruita in modo da assumere valori tanto più piccoli, quanto più è «accettabile» l'ipotesi H_0 e si dice **misura della discrepanza o discrepanza** fra la successione finita $\underline{x} = \{x_1, x_2, \dots, x_n\}$ e l'ipotesi H_0 ;

(2) si fissa un numero reale $\delta > 0$. Detto d il valore assunto da D per $\underline{X}$ che assume il valore $\underline{x}$, se $d \geq \delta$ si ritiene eccessiva la discrepanza e si giudica «non accettabile» H_0 ; se, invece, $d < \delta$ si ritiene sufficientemente piccola la discrepanza e si assume che H_0 sia «accettabile» rispetto alla statistica D .

Il numero δ è assegnato in modo da soddisfare una uguaglianza del tipo

$$\text{prob}(D \geq \delta) = \alpha,$$

con α numero reale positivo assegnato a seconda degli scopi da raggiungere. Ad esempio si pone $\alpha=0.1$ oppure $\alpha=0.01$ o $\alpha=0.001$.

Per fissare le idee, limitiamoci a considerare solo alcuni test e precisamente quelli relativi a

- (I) media dei numeri;
- (II) frequenza dei numeri;
- (III) frequenza delle coppie di numeri;
- (IV) runs.

(I) Si assume la seguente condizione sulla media: la media M delle x_i deve essere «vicina» a 0.5 , media della variabile uniforme continua. La situazione ideale è l'uguaglianza $M=0.5$.

Ricordiamo che, per un noto teorema, se $X_1, X_2, \dots, X_n$ sono variabili casuali indipendenti ed equidistribuite, con media m e varianza v , allora, per n «abbastanza grande», la variabile casuale

$$Z = [(X_1, X_2, \dots, X_n)/n - m] (n/v)^{1/2} \quad (2.1)$$

può considerarsi distribuita come una variabile casuale normale standard $N(0, 1)$.

Il test della media consiste nel fissare come discrepanza la statistica $D = |Z|$.

(II) Si divide l'intervallo $[0, 1]$ in h intervallini di uguale lunghezza con h molto minore dell'ampiezza n del campione e viene assunta la seguente condizione sulle frequenze: il numero di elementi appartenenti a ciascun intervallino deve essere vicino a n/h .

Fissata una successione finita $\{x_i\}$, con $i \in \{1, 2, \dots, n\}$, di elementi dell'intervallo $[0, 1]$, per un noto teorema, indicato con n_i , $i = 1, 2, \dots, h$, il numero di elementi della successione appartenenti all'intervallino I_i , per n abbastanza grande si può ritenere che la variabile casuale

$$D = (h/n) \sum_{i=1}^h (n_i/h - n_i)^2 = (h/n) \sum_{i=1}^h n_i^2 - h \quad (2.2)$$

abbia distribuzione chi-quadro ad $h-1$ gradi di libertà.

Nella situazione ideale, in cui ogni n_i è uguale a n/h , risulta $D=0$. Il test sulle frequenze dei numeri consiste nell'assumere come discrepanza la statistica data dalla (2.2).

(III) Si divide l'intervallo $[0, 1)$ in h intervallini di uguale lunghezza $I_1, I_2, \dots, I_h$. Si considera, poi, l'intervallo $[0, 1) \times [0, 1)$ come unione degli h^2 intervallini del piano, di uguale area, $I_{rs} = I_r \times I_s$, con $r, s = 1, 2, \dots, h$.

Se $\{x_1, x_2, \dots, x_n\}$ è il campione empirico, si esamina la successione delle coppie (x_i, x_{i+1}) , $i = 1, 2, \dots, n-1$. Si ammette la seguente condizione sulla frequenza delle coppie: poichè gli intervallini I_{rs} sono h^2 e le coppie sono $n-1$, la frequenza delle coppie in ciascun intervallino deve essere vicina a $(n-1)/h^2$.

Per n abbastanza grande, detto n_{rs} il numero di elementi della successione appartenenti all'intervallino I_{rs} , si può ritenere che la variabile casuale

$$D = [h^2/(n-1)] \sum_{r=1}^h \sum_{s=1}^h [(n-1)/h^2 - n_{rs}]^2 \quad (2.3)$$

abbia una distribuzione chi-quadro a h^2-1 gradi di libertà.

Il test sulle frequenze delle coppie di numeri consiste nell'assumere come discrepanza la statistica D data dalla (2.3).

(IV) Si assume la seguente condizione sui runs: il numero delle sottosuccessioni monotone massimali (dette runs) di una successione con n elementi deve essere «vicino» a $(2n-1)/3$.

La situazione ideale è quella in cui il numero di runs è esattamente uguale a $(2n-1)/3$. Ciò è basato sul fatto che, detta R la variabile casuale «numero di runs», si dimostra che, per n abbastanza grande, si può assumere che la variabile

$$Z = [R - (2n-1)/3] [90/(16n-29)]^{\frac{1}{2}} \quad (2.4)$$

sia normale standard.

Il test sui runs consiste nell'assumere come discrepanza la variabile casuale $D = |Z|$.

Le successioni di tipo (1.2) che si considerano sono, di solito, periodiche di periodo molto elevato, almeno dell'ordine di 10^{10} , mentre l'ampiezza dei campioni considerati nelle precedenti analisi statistiche è piccola rispetto a tali numeri, e va da $n=100$ fino a $n=100.000$ in alcuni casi.

Chiamiamo campionari i tests precedentemente esaminati.

Test di natura diversa, che chiamiamo di tipo «globale», sono utilizzati da Knuth (1969), [7], Maturo (1989), [9] e Rizzi (1989), [13]. Essi considerano il comportamento di successioni finite di lunghezza uguale al periodo della successione (1.2).

Anche tali test, tuttavia, sono fondati su un meccanismo di accettazione-rifiuto che si basa sulla distanza da valori ideali di certi valori ottenuti tramite opportune elaborazioni a partire dalla formula (1.2).

3. Le critiche ai test classici ed il problema della ricerca di analisi bayesiane “sostitutive”.

Generalizzando le considerazioni svolte nel paragrafo precedente possiamo quindi affermare che da un punto di vista classico l'analisi di casualità delle successioni di numeri pseudocasuali è basata sul seguente procedimento:

(a) Si considerano opportune statistiche $D_k = D_k(X_1, X_2, \dots, X_n)$, $k=1, 2, \dots, r$, che assumono valori reali positivi e che, in certe situazioni ritenute ideali, si annullano. Le D_k hanno, di solito, distribuzione asintotica o di tipo chi-quadro oppure uguale al valore assoluto di una distribuzione simmetrica rispetto allo zero e tabulata, come avviene ad esempio, per la normale standard o la legge di student.

(b) Per ogni k si fissa un numero $\alpha_k \in (0, 1)$ «piccolo», in genere $\alpha_k=0.01$ oppure $\alpha_k=0.001$ o un valore intermedio fra questi, detto livello di significatività, e si determina un numero reale β_k tale che $\text{prob}(D_k \geq \beta_k) = \alpha_k$.

La coppia $T_k = (D_k, \beta_k)$ si dice **test statistico** al livello di significatività α_k .

(c) Sia d_k il valore assunto da D_k per $X_i = x_i$, con $i=1, 2, \dots, n$. Se $d_k \geq \beta_k$ si respinge l'ipotesi H_0 rispetto al test T_k , in caso contrario si accetta l'ipotesi H_0 rispetto al test T_k .

(d) Si considerano «accettabili», rispetto all'insieme dei test considerati, le successioni per le quali l'ipotesi H_0 è accettata rispetto ad ognuno dei test T_k .

La logica rispetto alla quale viene rifiutata l'ipotesi H_0 se $d_k \geq \beta_k$ si basa su due considerazioni.

(1) **Valutazione probabilistica:** l'evento $E_k = (D_k \geq \beta_k)/H_0$ ha probabilità α_k molto piccola di verificarsi. Poichè si è verificato, ciò induce a pensare che l'ipotesi H_0 sia falsa;

(2) **valutazione psicologica:** se $d_k \geq \beta_k$ allora il valore assunto da D_k è molto lontano da zero, «valore ideale».

Alcuni errori presenti in questo tipo di logica, (cfr. ad esempio Scozzafava (1989), [14], e Maturo (1989), [10]) sono:

(a) la confusione fra E_k e l'evento $A_k = H_0/(D_k \geq \beta_k)$. Nella considerazione (1), dal fatto che la probabilità di E_k è piccola si deduce che è piccola anche quella di A_k .

Conclusioni di questo tipo portano a conseguenze paradossali.

Ad esempio, si può osservare che, qualunque sia la successione finita $\underline{x}$ di lunghezza n , se gli x_i sono numeri con q cifre decimali, allora $\text{prob}(\underline{x}/H_0)=10^{-q}$ e quindi risulta $\text{prob}(\underline{x}/H_0)=10^{-mq}$.

Poichè, per ogni $\alpha_k > 0$, esiste un intero n tale che $10^{-mq} < \alpha_k$, fissato α_k è possibile trovare un intero m tale che, per ogni $n \geq m$, l'evento $\underline{x}/H_0$ ha probabilità minore di α_k .

Allora, ragionando come in (1), siamo indotti a respingere H_0 qualunque sia $\underline{x}$, purchè sia $n \geq m$.

(b) L'elemento di disturbo introdotto dalla (2). Il punto di vista psicologico porta a «scartare la coda» della distribuzione di D_k , perchè in essa vi sono valori molto lontani dal valore ideale, 0, di D_k . Ragionando da un punto di vista rigorosamente probabilistico, invece, se a è un qualsiasi numero reale non negativo minore di β_k e b è un qualsiasi numero reale tale che l'evento $E_{ak}=(a \leq D_k \leq b)/H_0$ abbia probabilità α_k , l'evento E_{ak} potrebbe giocare, nel meccanismo di accettazione e rifiuto del test, lo stesso ruolo dell'evento E_k . Ad esempio, ponendo $a=0$, il test porterebbe a «scartare» la «parte iniziale» della distribuzione.

(c) La mancata considerazione di ipotesi alternative ad H_0 . Ammesso che realmente l'evento $A_{ik}=H_i/(D_k \geq \beta_k)$ abbia probabilità molto piccola, ciò non è un motivo sufficiente per respingere H_0 . Infatti può capitare che vi siano numerose ipotesi alternative H_i , $i=0, 1, 2, \dots, m$ e che tutti i vari eventi $A_{ik}=H_i/(D_k \geq \beta_k)$ abbiano probabilità molto piccole. Allora, se si verifica che $d_k \geq \beta_k$ tutte le ipotesi devono essere respinte?

In questo lavoro ci proponiamo di mettere a punto le idee di base per un'analisi bayesiana delle successioni pseudocasuali.

Una analisi bayesiana presenta i seguenti vantaggi:

(a) oltre all'ipotesi H_0 considera un insieme di ipotesi alternative da confrontare con H_0 ;

(b) fornisce risultati esclusivamente probabilistici senza introdurre, nei ragionamenti, meccanismi estranei al Calcolo delle Probabilità, come quelli di accettazione e rifiuto, ed è quindi esente dai difetti logici del metodo classico.

A nostro parere, il rifiuto da parte di alcuni ricercatori di effettuare la verifica di casualità di una successione pseudocasuale di lunghezza n da un punto di vista bayesiano nasce dall'equivoco di considerare una tale successione semplicemente come un campione di ampiezza n della variabile casuale uniforme continua, senza tener conto dell'ordinamento. Infatti da un tale punto di vista appaiono «migliori» delle altre delle successioni banali e non certo

casuali quali ad esempio quella di termine generale $x_i = (i \bmod m)/m$, con $i \in \mathbb{N}$.

In altre parole, un insieme di n numeri reali distinti $x_1, x_2, \dots, x_n$, appartenenti all'intervallo $[0,1)$ si può considerare un campione della variabile casuale uniforme continua, ma esistono $n!$ successioni a cui corrisponde tale campione e solo alcune di esse soddisfano al requisito di apparire casuali ad un osservatore.

In definitiva, l'analisi bayesiana che considera la successione come un campione di ampiezza n che chiamiamo, in seguito, per semplicità, analisi bayesiana delle frequenze è sostitutiva unicamente del test classico delle frequenze, ampiamente verificato per praticamente tutte le successioni pseudocasuali che si considerano.

Come avviene per tale test, il fatto che l'ipotesi H_0 sia favorita da queste analisi bayesiane è una condizione necessaria, ma del tutto insufficiente affinché sia accettabile il fatto che la successione S soddisfi la definizione soggettiva di Baldessari.

Una generalizzazione dell'analisi bayesiana delle frequenze, considerata in Di Biase e Maturo (1990), [5] e [6], è la seguente

(1) si considera l'ipotesi H_0 come un elemento di una famiglia di ipotesi $(H_\alpha)_{\alpha \in I}$, con I insieme di indici, tale che per un certo valore μ di I sia $H_\mu = H_0$ e si assegna, a priori ad ogni ipotesi H_α , in maniera soggettiva, con opportuni criteri, il valore $f(H_\alpha)$ di una pseudodensità, detta **pseudodensità a priori**, ossia di una funzione reale non negativa, definita nell'insieme degli eventi elementari e che tiene conto di una assegnata relazione di preordine fra tali eventi; essa può essere, in particolare, una probabilità o una densità di probabilità (cfr. Scozzafava (1983) e (1989), [15] e [17], e Maturo (1989), [11]);

(2) a partire da una data successione $\underline{x} = \{x_1, x_2, \dots, x_n\}$ di numeri pseudocasuali nell'intervallo $[0, 1)$ si calcolano le verosimiglianze $p(\underline{x}/H_\alpha)$, $\alpha \in I$;

(3) con la formula del Teorema di Bayes

$$g(H_\alpha/\underline{x}) = k f(H_\alpha) p(\underline{x}/H_\alpha), \quad \text{con } k \text{ funzione solo di } \underline{x} \quad (3.1)$$

si calcolano i valori di una nuova pseudodensità $g(H_\alpha/\underline{x})$, detta **pseudodensità a posteriori**;

(4) si vede, con criteri convenzionali opportuni, se viene favorita l'ipotesi $\alpha = \mu$, ad esempio controllando se il punto di massimo di g è «sufficientemente vicino» a μ oppure se, fissata una opportuna partizione di I «è favorito» l'insieme della partizione a cui appartiene μ .

Osserviamo che, anche in questo caso generale, dall'ipotesi H_0 di indi-

pendenza ed equidistribuzione delle X_i , segue che ogni successione S' ottenuta come permutazione di $\underline{x}$ dà luogo alle stesse verosimiglianze e, quindi, agli stessi valori della g . Di conseguenza una tale analisi bayesiana fornisce informazioni solo sulle frequenze dei valori $\underline{x}_i$ assunti dalle X_i , e, pertanto, può sostituire solo il test classico delle frequenze.

Per poter eseguire la verifica di casualità delle successioni pseudocasuali da un punto di vista bayesiano è necessario, quindi, a nostro parere, effettuare una analisi bayesiana sostitutiva per ciascuno dei test classici. Ad esempio una verifica di casualità di tipo campionario può essere svolta considerando una analisi bayesiana sostitutiva per ciascuno degli 11 test del paragrafo precedente ed una verifica di casualità di tipo globale considerando una analisi bayesiana sostitutiva per ciascuno dei test di Knuth (1969), [7].

Proponiamo, quindi, il seguente procedimento generale.

Dalla successione finita $\underline{x} = \{x_1, x_2, \dots, x_n\}$ di numeri pseudocasuali ottenuta come illustrato nel paragrafo 1 se ne deduce un'altra del tipo:

$$T = \{z_1, z_2, \dots, z_{n+r}\}$$

con r intero positivo opportuno, $r \ll n$, dove le z_i sono vettori di uno spazio R^k , con $k \geq 1$, funzioni del vettore $(x_1, x_2, \dots, x_n) \in R^n$ e determinazioni di variabili casuali Z_i equidistribuite ed indipendenti.

Le Z_i dipendono da determinati parametri e, per opportuni valori di essi, hanno distribuzione uguale a quelle ideali richieste dal test classico che viene sostituito.

In definitiva si ammette che:

- (1) i vettori z_i siano una determinazione di variabili casuali $Z_i(\alpha_1, \alpha_2, \dots, \alpha_n)$ equidistribuite, indipendenti e dipendenti dai parametri $\alpha_1, \alpha_2, \dots, \alpha_n$;
- (2) in condizioni ideali di casualità i parametri $\alpha_1, \alpha_2, \dots, \alpha_n$ assumono i valori $\mu_1, \mu_2, \dots, \mu_n$;
- (3) si assume una distribuzione a priori di $(\alpha_1, \alpha_2, \dots, \alpha_n)$, in genere tale che privilegia il valore $(\mu_1, \mu_2, \dots, \mu_n)$, ad esempio che assume il massimo in corrispondenza a tale valore;
- (4) indicando con $f(\underline{\alpha})$ la pseudodensità a priori di $\underline{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_n)$ e con $p(\underline{z}/\underline{\alpha})$ la probabilità di $\underline{z} = (z_1, z_2, \dots, z_n)$ dato $\underline{\alpha}$, si calcola la pseudodensità a posteriori

$$g(\underline{\alpha}/\underline{z}) = k f(\underline{\alpha}) p(\underline{z}/\underline{\alpha}) \quad (3.2)$$

dove k è una funzione di $\underline{z}$. Se la $g(\underline{\alpha}/\underline{z})$ è una densità di probabilità si ha

$$k = \frac{1}{\int_{R^h} g(\underline{\alpha}, \underline{z}) d\alpha_1 d\alpha_2 \dots d\alpha_h}$$

(5) si guarda se la pseudodensità (3.2) «tende a favorire» il valore $(\mu_1, \mu_2, \dots, \mu_h)$, dando eventualmente dei criteri operativi, che hanno solo carattere convenzionale, di giudizio sul grado di soddisfacimento delle condizioni ideali.

Alcuni casi particolari sono ad esempio:

(1) analisi bayesiana sulle frequenze; si pone $z_i = x_i$, per ogni $i \in \{1, 2, \dots, n\}$;

(2) analisi bayesiana sulle coppie; si pone $z_i = [x_i, x_{i+1}]'$, per ogni $i \in \{1, 2, \dots, n-1\}$.
Risulta $r=1$;

(3) analisi bayesiana sulle k -ple; è una generalizzazione dei casi precedenti. Si pone $z_i = [x_i, x_{i+1}, \dots, x_{i+k-1}]'$, per ogni $i \in \{1, 2, \dots, n-k+1\}$. Risulta $r=k-1$;

(4) analisi bayesiana sui runs; supponiamo che per ogni i sia $x_i \neq x_{i+1}$, ciò che praticamente è sempre verificato; si pone:

$$z_i = \begin{cases} 0 & \text{per } i > 1 \text{ e } x_i < x_{i+1} < x_{i+2} \text{ oppure } x_i > x_{i+1} > x_{i+2} \\ 1 & \text{in caso contrario} \end{cases}$$

per ogni $i \in \{1, 2, \dots, n-2\}$.

La distribuzione degli z_i non soddisfa, però, alle condizioni di indipendenza. Allora il ragionamento si modifica considerando m successioni $\underline{z}^{(j)} = \{z_{ji}, \dots, z_{jn}\}$ casuali di ampiezza n e, per ognuna di esse, il numero

$$v_j = \sum_{i=1}^n z_{ji}$$

La distribuzione ideale dei v_j (cfr. Rizzi (1977), [12]) è quella normale con media $(2n-1)/3$ e varianza $(16n-29)/90$. I parametri su cui eseguire l'inferenza sono la media θ e la varianza σ^2 .

I v_j si possono considerare indipendenti e la (3.2) diventa

$$g(\underline{\alpha}/\underline{v}) = k f(\underline{a}) p(\underline{v}/\underline{a}). \quad (3.3)$$

BIBLIOGRAFIA

1. B. Baldessari (1987), *Aspetti probabilistici della crittografia*, Atti del 1° Simposio Nazionale su "Stato e Prospettive della Ricerca Crittografica in Italia", pp. 9-21.
2. N. Cera e A. Maturo (1983), *Confronto fra alcuni generatori di numeri pseudocasuali*, Facoltà di Architettura, Pescara.
3. N. Cera e A. Maturo (1990), *Generazione di numeri pseudocasuali per mezzo di relazioni di ricorrenza in campi di Galois*, Periodico di Matematica, n. 2, pp. 33-56.
4. N. Cera e A. Maturo (1991), *Analisi della bontà di alcuni generatori di numeri pseudocasuali per la cifratura dei messaggi e la simulazione*, Ratio Mathematica, 2, in corso di stampa.
5. G. Di Biase e A. Maturo (1990), *Analisi bayesiana di successioni pseudorandom*, Atti del XIV Conv. Ann. AMASES, Pescara 13-15 settembre 1990, pp. 293-313.
6. G. Di Biase e A. Maturo (1990), *Su un'analisi bayesiana per la verifica di casualità di successioni pseudorandom: software applicativo ed interpretazione dei risultati*, Atti del Conv. "Classificazione e analisi dei dati, Metodi, Software, applicazioni", Pescara 11-12 ottobre 1990, in corso di stampa.
7. D.E. Knuth (1969), *The art of Computer programming vol. 2*, Addison Wesley, London.
8. A. Maturo (1989), *Numeri pseudocasuali*, Libreria dell'Università, Pescara.
9. A. Maturo (1989), *Analisi di Fourier di successioni di numeri pseudocasuali*, Atti del 2° Simposio Nazionale su "Stato e Prospettive della Ricerca Crittografica in Italia", pp. 188-198.
10. A. Maturo (1989), *Probabilità e statistica con il calcolatore: problematiche di carattere logico ed operativo*, Convegno "Insegnamento integrato di probabilità soggettiva e statistica bayesiana", Teramo.
11. A. Maturo (1989), *Probabilità finitamente additiva a valori nel campo delle serie bilatere*, Rendiconti di Matematica, VII, 9, pp. 67-85.
12. A. Rizzi (1977), *Generazione di distribuzioni statistiche mediante un elaboratore elettronico*, Istituto di Statistica e Ricerca sociale C.Gini, Roma.
13. A. Rizzi (1989), *Verifiche di Pseudo-Casualità in crittografia*, Atti del 2° Simposio Nazionale su "Stato e Prospettive della Ricerca Crittografica in Italia", pp. 3-21.
14. R. Scozzafava (1989), *La probabilità soggettiva e le sue applicazioni*, Editoriale Veschi, Masson, Milano.
15. R. Scozzafava (1983), *Finitely Additive Probabilities and Proper 'Improper' Priors in Bayesian Statistics*, Institute of Statistics Mimeo Series, 1541, Chapel Hill (U.S.A.).

16. R. Scozzafava (1984), *A Survey of Some Common Misunderstandings Concerning the Role and Meaning of Finitely Additive Probabilities in Statistical Inference*, *Statistica*, 44, pp. 21-45.
17. R. Scozzafava (1982), *Probabilità s-additive e non*, *Bollettino U.M.I.*, 6, 1-A.
18. R.C. Tausworthe (1965), *Random numbers generated by linear recurrence modulo two*, *Mathematics of Computation*, 19.

UN MODELLO NEOKEYNESIANO DI EQUILIBRIO

Mauro Gallegati - Laura Gardini (*)

Sommario - Si presenta un modello postkeynesiano, ottenuto integrando l'impostazione marshall-minskiana con la determinazione kaleckiana dei profitti e con la teoria monetaria di Keynes. Si discute il modello non lineare alle differenze finite proposto e se ne esegue l'analisi parametrica, che permette di individuare un comportamento dinamico di carattere classico. Si determinano le condizioni di esistenza di un unico punto di equilibrio economicamente significativo, condizioni sufficienti di stabilità asintotica, e si commentano i risultati da un punto di vista economico.

1. INTRODUZIONE

La teoria dell'instabilità finanziaria di Minsky (1975, 1982) riprende una tradizione nella modellistica economica che può essere fatta risalire a Marshall. Secondo tale approccio, la stabilità dell'economia reale può essere modificata dalla parte monetaria, mediante creazioni di credito che permettono di espandere gli investimenti al di là del limite imposto dall'autofinanziamento. La scuola marshalliana e lo stesso Keynes enfatizzano più volte come un disordine creditizio si ripercuota sul funzionamento "tranquillo" dell'economia reale. In "The Economics of Industry" (1979, p. 119), Marshall analizza un tipico ciclo economico in cui il credito facile non è la causa della crescita, dovuta ad un evento esogeno, ma il fattore che trasforma la fase ascendente in "boom" e finisce poi per provocare la successiva recessione. La causa essenziale del ciclo è spesso oggetto di controversia tra i vari autori della scuola di

(*) Istituto di Scienze Economiche, Facoltà di Economia e Commercio, Università degli studi di Urbino.

Lavoro presentato al "Convegno di Matematica Applicata all'Economia e all'Ingegneria", Pescara, gennaio 1989; ed in parte realizzato con i contributi del M.P.I.

Cambridge, mentre un sostanziale accordo esiste sul fatto che la struttura finanziaria influenzi il comportamento ciclico dell'economia, secondo le ipotesi avanzate da Marshall.

A differenza dei contributi di Hayek (1931) e delle scuole monetariste (Sargent (1979), Friedman (1968), Lucas (1987)), la stabilità dei modelli "marshall-minskiani" non possiede quelle caratteristiche di ottimalità che dovrebbero discendere dall'applicazione dei teoremi Arrow-Debreu (Arrow-Hahn (1971), Debreu (1959)). I valori di equilibrio non sono "naturali", nel senso che interventi di politica economica sono in grado di migliorare in maniera permanente il loro livello.

Questo risultato si può ottenere ribaltando l'impostazione dell'approccio tradizionale che cerca di microfondare la struttura macroeconomica. In questo lavoro si è seguita la metodologia keynesiana. Essa consiste nel distinguere due fasi nello studio dei problemi economici: quello principale deve risolvere la determinazione dei vincoli di bilancio, quello secondario si occupa della distribuzione dei redditi individuali. In particolare, nel modello viene integrata l'impostazione "marshall-minskiana" con la determinazione kaleckiana dei profitti e con la teoria monetaria di Keynes. Questo permette di individuare un comportamento dinamico di carattere classico, espresso dalla convergenza ad uno stato stazionario. A differenza dello stato stazionario neo-classico, questa soluzione di lungo periodo non garantisce la piena occupazione del fattore lavoro (se non "by design or accident"). Il punto di equilibrio è infatti determinato dai particolari lavori dei parametri del modello, mentre

l'equilibrio di piena occupazione richiede che la domanda aggregata (nonché le previsioni imprenditoriali della medesima) intersechi l'offerta aggregata esattamente in corrispondenza del massimo livello di occupazione che i lavoratori accettano al salario atteso dagli imprenditori, che è il salario in base al quale viene definita la curva di offerta aggregata appropriata (Chick (1982), pag. 134).

Il presente lavoro è costituito da due parti: la descrizione ed il significato economico del modello proposto, e l'analisi parametrica del sistema.

Nel prossimo paragrafo ricaviamo la curva IS come luogo dei punti di equilibrio tra risparmi e investimenti nel piano profitti-prezzo del capitale, ed una curva LM che, nel medesimo piano, rappresenta l'insieme dei punti in cui la domanda uguaglia l'offerta. Si viene così ad avere un sistema non lineare alle differenze finite, la cui analisi qualitativa è sviluppata nel paragrafo 3. Si forniscono condizioni di esistenza e di stabilità di un unico punto di equilibrio economicamente significativo. Si discutono nelle conclusioni i risultati ottenuti.

2. DESCRIZIONE DEL MODELLO

2.1. L'EQUAZIONE IS

Dalle equazioni di contabilità nazionale abbiamo che in un'economia chiusa agli scambi con l'estero il reddito, Y , è pari a

$$Y_t = C_t + I_t + \bar{G} \quad (1)$$

dove $\bar{G} = G_c + G_i$, esogeno, rappresenta la spesa pubblica.

I consumi privati, C , sono determinati alla Kaldor (1955) dalla

$$C_t = (1-\tau) (cwN + c' (1-\theta) \pi_{t-1}) \quad (2)$$

dove c e c' (con $0 < c < 1$, e $0 < c' < 1$) sono, rispettivamente, le propensioni al consumo dei salariati e dei percettori di profitto; θ il tasso di ritenzione dei profitti, π ; w rappresenta il saggio unitario di salario; N , l'occupazione complessiva, e τ ($0 < \tau < 1$), l'aliquota media di imposizione fiscale diretta. In particolare:

$$N = N_c + N_i \quad (3)$$

$$N_c = (C + G_c) / p\lambda_c, \quad N_i = (I + G_i) / p\lambda_i \quad (4)$$

$$w/p = \lambda' \quad (5)$$

dove N_i e N_c rappresentano, rispettivamente, l'occupazione complessiva nei due settori beni di investimento e di consumo; λ_i e λ_c le produttività medie del lavoro nei due settori, che per semplicità assumiamo identiche, $\lambda_i = \lambda_c = \lambda$; λ' è la produttività marginale del lavoro stesso e p il prezzo della produzione corrente. Data la sommatoria delle singole curve di produttività marginale (derivabile dalla (5)), sarà la domanda aggregata a determinare il livello di occupazione corrispondente al punto di domanda effettiva.

I modelli post-Keynesiani si caratterizzano poi per avere due distinti livelli dei prezzi; abbiamo pertanto l'equazione del prezzo della produzione corrente

$$p = w (1 + \mu) / \lambda \quad (6)$$

dove μ è il "mark-up" unitario, per semplicità assumiamo che questo sia identico nei due settori, $\mu_i = \mu_c = \mu$.

Il mercato del lavoro è modellato in maniera tale che, pur rispettando il principio di massimizzazione del profitto, per cui il salario reale è pari

alla produttività marginale del lavoro, l'ammontare complessivo di occupazione, N^* , viene ad essere determinato dal punto di domanda effettiva, ED, come intersezione tra le curve di domanda, AD, e di offerta aggregata, AS, determinata dal prodotto $p\lambda N^d$, come risulta dalla Fig. 1 (si veda Chick (1982), pag. 253, ed il "Symposium" nel "Journal of Post Keynesian Economics", (1987)). Poiché in tale punto la curva di domanda ha una pendenza, determinata dalla propensione al consumo, minore di quella dell'offerta, l'equilibrio è stabile pur in presenza di disoccupazione involontaria. La produzione si "adequa", almeno nel breve periodo, al livello della domanda aggregata attesa che però, anche in presenza di aspettative razionali, non può essere esattamente prevista. Questo è imputabile al fatto che il tempo di produzione dei beni di investimento richiede un orizzonte temporale più lungo rispetto a quello delle merci disponibili, e che quindi ogni singolo operatore deve formare le proprie aspettative sulle aspettative degli altri, inficiando in tal modo il processo di convergenza all'"equilibrio di aspettative razionali" (si veda Frydman-Phelps (1983)).

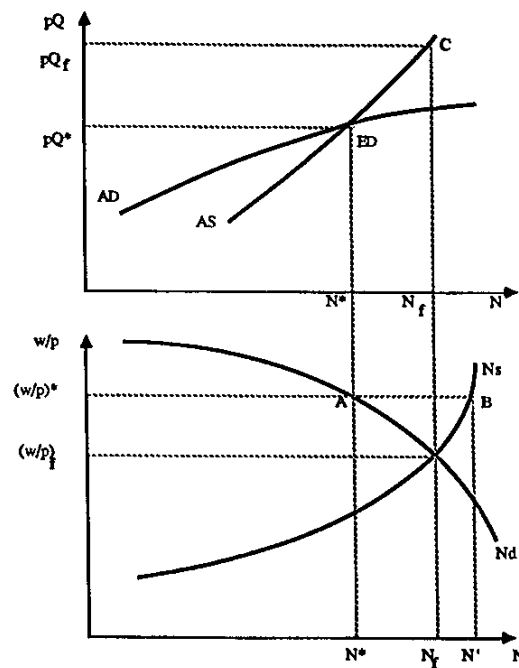


Fig. 1

L'offerta aggregata determina il punto (di domanda effettiva) sulla funzione di domanda aggregata, che mette in relazione la domanda di lavoro col salario reale che massimizza i profitti attesi, al quale le imprese decidono di operare. Se si vuole individuare una causa prima nella disoccupazione essa va dunque ricercata nell'insufficiente domanda aggregata piuttosto che in un salario reale "troppo elevato".

In quanto segue, assumiamo che tanto i salari quanto il capitale, siano esogeni nel breve periodo, definito come periodo coperto dall'accordo salariale contrattuale. Con queste assunzioni le decisioni degli imprenditori si limitano a:

- i) stabilire il livello degli investimenti, e quindi del reddito complessivo;
- ii) fissare i prezzi, tenuto conto delle condizioni di mercato, determinando in tal modo la distribuzione del reddito.

La funzione degli investimenti è definita dalla

$$I_t = a (V_t - p_t) + b I_t^f \quad (7)$$

dove $(V_t - p_t)$ svolge una funzione analoga alla "q" di Tobin (1968, 1969), con V prezzo dello stock di capitale, e p_t prezzo dell'investimento corrente, che assumiamo uguale a p : $p_t = p$; I_t^f rappresenta l'autofinanziamento, ed i due coefficienti sono positivi: $a > 0$ e $0 < b < 1$. Il prezzo dei beni di investimento e di consumo viene ad essere determinato secondo la regola kaleckiana del "mark up", dato il livello del salario nominale corrente, il grado di monopolio, la fase ciclica ed il livello della produttività. La Fig. 2 mostra graficamente la formulazione minskiana della teoria dell'investimento riferita ad una singola unità produttiva. Rispetto alla versione standard della IS-LM, che assume una relazione negativa tra investimenti e tasso di interesse, Minsky enfatizza la visione keynesiana di un carattere essenzialmente di tipo finanziario e speculativo che la detenzione di attività e l'investimento privato hanno in una moderna economia capitalistica, Minsky (1975).

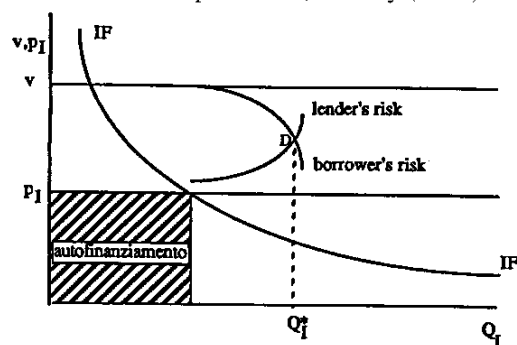


Fig. 2

L'equazione (7) rappresenta in forma lineare la formalizzazione della teoria finanziaria dell'investimento di Minsky. Il primo argomento nella funzione di investimento rappresenta la distanza verticale tra il prezzo dei "capital assets" e quello dell'investimento corrente. Nell'interpretazione di Minsky, la condizione $V > p$ rappresenta un "requisito di vitalità del sistema". Il rapporto tra EF ed il secondo argomento, IF, rappresenta il grado di "leverage", L, che comprende sia il "borrower's" che il "lender's risk" (sulla definizione di rischio del debitore e del creditore si veda Keynes (1936), pag. 112 ss). Durante una fase espansiva ci si può attendere tanto che lo stock di "leverage" diminuisca grazie all'aumento del corso delle azioni (salvo poi registrare un, possibilmente brusco, aumento dello stesso a seguito di una caduta dei prezzi delle emissioni azionarie), quanto che esso aumenti, come ipotizza Minsky (1982).

Si noti che, mentre un aumento del prezzo dei "capital assets" o dei profitti provoca un aumento dell'investimento, l'aumento del "leverage" ne causa una diminuzione rendendo più rigide le curve di rischio del debitore e del creditore. A sua volta il grado di indebitamento è funzione di V e π . Ma, mentre il "leverage" diminuisce senz'altro al crescere del prezzo dei "capital assets", più controverso è il caso in cui a crescere siano i profitti. Questi, da un lato consentono un maggior peso dell'autofinanziamento sul totale degli investimenti, dall'altro, se la propensione marginale ad investire è maggiore della propensione media, comporta un aumento del "leverage" al crescere dei profitti.

La teoria della fragilità sistemica si basa su quest'ultimo assunto. Durante la fase iniziale di un periodo di espansione, le imprese si indebitano in misura crescente, basandosi sulla prosperità che attualmente sperimentano. Le banche condividono lo stesso ottimismo e ottenere credito diviene più facile. Una sorta di circolo "virtuoso" si instaura: la remissione dei debiti conferma la bontà delle decisioni assunte in precedenza e interagisce sulle nuove richieste di finanziamento; è un processo di crescita cumulativo che ricorda molto da vicino quello descritto da Marshall e che finisce con lo sfociare in una crisi o in una recessione, a seconda dell'andamento delle aspettative degli operatori. In pratica, senza un organismo di coordinamento, sia esso un "Ministro della produzione" che pianifichi l'attività di investimento, o una Banca Centrale in grado di valutare l'andamento del grado di indebitamento, ogni singolo operatore tenderebbe ad indebitarsi in maniera crescente all'incremento dei profitti.

Il finanziamento interno IF_t è dato da

$$IF_t = \theta (1 - \tau) \pi_{t-1} \quad (8)$$

ed il finanziamento esterno, EF_t (l'impresa ha accesso a fonti esterne

soltanto domandando "finance"), da

$$EF_t = I_t - IF_t \quad (9)$$

Deriviamo ora la curva IS come luogo dei punti in cui i risparmi uguagliano gli investimenti. Usando (8), la (7) diviene:

$$I_t = a (V_t - p) + b \theta (1 - \tau) \pi_{t-1} \quad (10)$$

ed usando (3, 4, 6), la (2) diventa:

$$C_t = (1 - \tau) (\bar{G} + I_t + (1 + \mu) c' (1 - \theta) \pi_{t-1}) / \beta_o \quad (11)$$

dove si è definito

$$\beta_o = 1 + \mu - c (1 - \tau)$$

Da (1), utilizzando la relazione

$$Y_t = W_t + \pi_t = wN + \pi_t \quad (12)$$

otteniamo

$$\pi_t = C_t + I_t + \bar{G} - wN$$

ed usando (3, 4, 6)

$$\pi_t = (C_t + I_t + \bar{G}) \mu / (1 + \mu)$$

sostituendo in quest'ultima C_t ed I_t date rispettivamente in (11) e (10), otteniamo l'equazione IS:

$$\pi_t = \phi_o + \phi_1 V_t + \phi_2 \pi_{t-1}; \quad (13)$$

$$\begin{cases} \phi_o = (\bar{G} - ap) \mu / \beta_o & (>0) \\ \phi_1 = a \mu / \beta_o & (>0) \\ \phi_2 = (1 - \tau) (b \theta + c' (1 - \theta)) \mu / \beta_o & (>0) \end{cases} \quad (14)$$

2.2 L'EQUAZIONE LM

Integrando la teoria monetaria del "Treaties on Money" e della "General Theory", alla luce del successivo dibattito sulla "scuola svedese" (Ohlin (1973), Keynes (1973b, 1938)), si possono distinguere tre tipi di domanda di moneta:

- (i) la domanda transattiva, TR, ("active balances") a cui, per semplificare, assimiliamo la domanda per scopi precauzionali;
- (ii) la domanda speculativa, SP, ("hoards" o "inactive balances");
- (iii) la domanda di "finance" (Keynes (1937a), Asimakopoulou (1985)).

L'offerta di moneta viene ad essere determinata dalla base monetaria, H, emessa dalla banca centrale e dal moltiplicatore, $m = (1 + cu)/(re + cu)$, dove "cu" è pari al rapporto tra circolante e depositi, e "re" a quello tra riserve e depositi. Poiché entrambi riflettono le preferenze degli operatori e delle banche durante il ciclo, la quantità di moneta può essere determinata endogenamente (Minsky (1975)).

Vanno segnalate alcune differenze rispetto alla teoria tradizionale. In primo luogo la direzione di causalità moneta-reddito va qualificata. In base alla teoria keynesiana del "finance" è l'ammontare di produzione stabilito dalle imprese che "domanda" moneta, anche se, per via dei tempi tecnici di produzione, lo stock di moneta cresce prima del reddito. Non si ha quindi quell'effetto propulsivo proveniente dal settore monetario, teorizzato dalla scuola di Chicago. E' evidente come una politica monetaria restrittiva, tagliando le disponibilità di credito all'economia tramite un razionamento dello stesso, finisca con l'influenzare il sentiero e la velocità di accumulazione fino ad aggravare, ed in determinate circostanze provocare, una situazione di recessione.

In secondo luogo, una volta stabilito che l'aggiustamento fra investimenti e risparmi avviene attraverso variazioni dell'occupazione, anziché tramite cambiamenti nel tasso di interesse, si deve riconoscere che

il tasso di interesse non può essere determinato dalla domanda e dall'offerta di capitale, perché l'investimento provoca un uguale ammontare di risparmio. Quindi l'investimento "finanzia se stesso", qualunque sia il livello del tasso di interesse (Kalecki (1954), pag. 73).

Le determinanti del saggio di interesse vanno dunque ricercate altrove: segnatamente nel carattere sequenziale di un processo economico che si svolge in condizioni di incertezza.

La scuola monetarista propende a considerare esogena la quantità di moneta, o comunque determinata dalle decisioni delle autorità monetarie, mentre il filone post-keynesiano statunitense (cfr. Moore (1988)) ritiene completamente indeterminata la stessa per via delle erogazioni creditizie infraimprese. In questo scritto assumiamo esogena la quantità

complessiva della stessa. In quanto determinata dall'incontro tra le domande degli operatori e l'offerta di base monetaria aumentata per l'operare del moltiplicatore, la si può considerare di ammontare finito, mentre endogeno è il tasso di interesse.

Secondo la tradizione di Cambridge la domanda di moneta transattiva è una funzione lineare del reddito nominale disponibile; la seguente definizione

$$TR_t = k Y_t - \beta r_t \quad (15)$$

esprime la domanda transattiva come funzione decrescente del tasso di interesse per dato livello di reddito. Poiché k , $0 < k < 1$, dipende dalle abitudini e dallo stato generale di fiducia, la (15) è una funzione potenzialmente molto instabile (Roe (1973)).

La domanda speculativa è la somma di tutte le domande individuali. Seguendo il "Treatise", nel quale Keynes delinea una teoria della "propensity to hoard" o del "degree of bearishness" analiticamente compatibile con la nozione di preferenza per la liquidità sviluppata nella "General Theory", definiamo ciascuna funzione di domanda individuale come (si veda Delli Gatti (1987))

$$SP_j = \alpha_j (V - V^c); \alpha_j > 0 \text{ per } j = 1, \dots, n$$

dove V_j è il prezzo atteso del "capital asset" dall'unità j -esima.

Dalla sommatoria delle singole domande individuali si ricava la domanda complessiva

$$SP_t = \sum_{j=1}^n SP_j = h (V_t - V^c); h = \sum_{j=1}^n \alpha_j \quad (16)$$

La domanda di "finance" come funzione del "financing gap" delle imprese è espressa dalla (9). La domanda di moneta è così pari alla somma delle tre quantità:

$$M_t^d = EF_t + SP_t + TR_t \quad (17)$$

mentre l'offerta è data dalla

$$M_t^s = \bar{M} = H m \quad (18)$$

L'equilibrio nel mercato della moneta è dato dalla

$$M_t^s = M_t^d. \quad (19)$$

In un modello con due “assets” moneta ed azioni abbiamo in equilibrio la seguente equazione di arbitraggio

$$r_t = ((1 - \theta) \pi_{t-1} + \pi^c) / V_t \quad (20)$$

Ricaviamo ora l'equazione che chiamiamo LM. Dalla (19), usando (18), (17), e le definizioni dei termini che in essa compaiono date in (9, 15, 16), otteniamo

$$\bar{M} = I_t - \theta (1 - \tau) \pi_{t-1} + h(V_t - V^c) + kY_t - \beta r_t \quad (21)$$

Dalla (12), usando (3, 4, 6), otteniamo

$$kY_t = k\pi_t + (\bar{G} + I_t + (1 - \tau) c' (1 - \theta) \pi_{t-1}) k / \beta_0 \quad (22)$$

Sostituendo I_t e kY_t , dati rispettivamente in (7) e (22), nella (21) si ottiene, esplicitando π_t , la seguente equazione LM:

$$\pi_t = \gamma_0 + \gamma_1 V_t + \gamma_2 \pi_{t-1} + \gamma_3 r_t; \quad (23)$$

$$\begin{cases} \gamma_0 = (M + hV^c + ap) / k - \phi_0 / \mu & (>0) \\ \gamma_1 = - (a + h) / k - \phi_1 / \mu & (>0) \\ \gamma_2 = \theta (1 - \tau) (1 - b) / k - \phi_2 / \mu & (>< 0) \\ \gamma_3 = \beta / k & (>0) \end{cases} \quad (24)$$

3. ANALISI DEL MODELLO

3.1 RIDUZIONE AD UNA MAPPA

Le equazioni che costituiscono il modello che vogliamo studiare sono dunque la (13) e la (23) con la (20), che riscriviamo per comodità

$$\begin{cases} \pi_t = \phi_0 + \phi_1 V_t + \phi_2 \pi_{t-1} \\ \pi_t = \gamma_0 + \gamma_1 V_t + \gamma_2 \pi_{t-1} + \gamma_3 r_t \\ r_t = ((1 - \theta) \pi_{t-1} + \pi^c) / V_t \end{cases} \quad (25)$$

con i seguenti vincoli sui coefficienti:

$$\begin{cases} \theta \in [0, 1]; \pi^c, \phi_0, \phi_1, \phi_2 > 0 \\ \gamma_0 > 0, \gamma_1 > 0, \gamma_2 >< 0, \gamma_3 \geq 0 \end{cases} \quad (26)$$

Si osserva immediatamente che se $\gamma_3 = 0$ (oppure r è costante), il modello si riduce ad una mappa lineare:

$$\pi_t = \lambda \pi_{t-1} + \eta \quad ; \quad (27)$$

$$\lambda = \frac{\varphi_1 \gamma_2 - \varphi_2 \gamma_1}{\varphi_1 - \gamma_1} \quad , \quad \eta = \frac{\varphi_1 \gamma_0 - \varphi_0 \gamma_1}{\varphi_1 - \gamma_1} + \left(\frac{\gamma_3 \varphi_1 r}{\varphi_1 - \gamma_1} \right) \quad (28)$$

che discutiamo brevemente. Il punto di equilibrio della (27), per $\lambda \neq 1$, è dato da $\pi^* = \eta / (1 - \lambda)$. Dalle condizioni (26) segue $\eta > 0$, così che il punto di equilibrio è economicamente significativo, cioè $\pi^* > 0$, solo per $\lambda < 1$, ed è stabile per $-1 < \lambda < 1$. Tuttavia, il segno di λ è definito, in quanto nella (28) il denominatore è positivo, ed il segno della quantità al numeratore risulta dalla scrittura esplicita dei coefficienti in funzione dei parametri economici tramite le (14) e (24), e si ha:

$$(\varphi_1 \gamma_2 - \varphi_2 \gamma_1) > 0 \quad (29)$$

Quindi, se teniamo conto del vincolo (29), cioè $\lambda > 0$, possiamo concludere che la mappa lineare (27) ha un punto di equilibrio economicamente significativo nel dominio $0 < \lambda < 1$, che è (globalmente e asintoticamente) stabile con transitori di tipo monotono (sono escluse le oscillazioni improprie, cioè gli andamenti alternanti).

Consideriamo in seguito il sistema (26) con $\gamma_3 > 0$. Per ridurre il modello ad una mappa scritta in forma standard, cioè a una equazione non lineare alle differenze finite in forma esplicita,

$$\pi_t = f(\pi_{t-1}) \quad (30)$$

sono necessari alcuni passaggi algebrici. Dall'uguaglianza delle prime due equazioni in (25), ed usando la terza, si ottiene la seguente equazione del secondo ordine per V_t :

$$(\varphi_1 - \gamma_1) V_t^2 - [(\gamma_0 - \varphi_0) - (\varphi_2 - \gamma_2) \pi_{t-1}] V_t - \gamma_3 ((1 - \theta) \pi_{t-1} + \pi^e) = 0 \quad (31)$$

il cui discriminante, dato da

$$\Delta(\pi) = [(\gamma_0 - \varphi_0) - (\varphi_2 - \gamma_2) \pi]^2 + 4\gamma_3 (\varphi_1 - \gamma_1) ((1 - \theta) \pi + \pi^e) \quad (32)$$

è positivo se assumiamo $\pi_t > 0$ per ogni t (vedremo che si presenta effettivamente questo caso), per cui l'equazione (31) ha due radici reali di segno opposto (come segue dalla regola di Cardano). Denotiamo con V_1 la radice positiva,

$$V_t = [(\gamma_0 - \varphi_0) - (\varphi_2 - \gamma_2) \pi_{t-1}] + \sqrt{\Delta(\pi_{t-1})} / 2 (\varphi_1 - \gamma_1) \quad (33)$$

che è la sola radice che interessa in questo modello economico.

Sostituendo la (33) nella prima di (25) otteniamo una mappa di tipo (30), con

$$f(\pi) = \frac{1}{2} \left[\varphi_0 + \frac{\varphi_1 \gamma_0 - \varphi_0 \gamma_1}{\varphi_1 - \gamma_1} \right] + \frac{1}{2} \left[\varphi_2 + \frac{\varphi_1 \gamma_2 - \varphi_2 \gamma_1}{\varphi_1 - \gamma_1} \right] \pi + \frac{\varphi_1}{2(\varphi_1 - \gamma_1)} \sqrt{\Delta(\pi)} \quad (34)$$

che rappresenta il modello dinamico di interesse. Si noti che se $\pi_0 > 0$, allora V_t e π_t sono positivi per ogni $t > 0$.

Interessa ora determinare, se esistono, i punti di equilibrio della mappa (34) che sono economicamente significativi, cioè $\pi^* > 0$ e tali che i corrispondenti valori di equilibrio V^* siano positivi; chiamiamo ammissibili tali punti. Vedremo sotto quali condizioni esiste un unico punto di equilibrio ammissibile, alcune condizioni di stabilità ed alcuni comportamenti qualitativi della mappa (34).

3.2 PUNTI DI EQUILIBRIO

I punti di equilibrio della mappa (30) sono date dalle soluzioni dell'equazione $f(\pi) = \pi$. Con ovvi passaggi algebrici otteniamo l'equazione del secondo ordine

$$(1 - \varphi_2) A \pi^2 - B \pi + C = 0; \quad (35)$$

$$A = (\varphi_1 - \gamma_1) - (\varphi_1 \gamma_2 - \varphi_2 \gamma_1)$$

$$B = \varphi_3 (1 - \theta) \varphi_1^2 + \varphi_0 A + (1 - \varphi_2) (\varphi_1 \gamma_0 - \varphi_0 \gamma_1)$$

$$C = \varphi_3 \pi^c \varphi_1^2 + \varphi_0 (\varphi_1 \gamma_0 - \varphi_0 \gamma_1)$$

Le radici dell'equazione (35) sono punti stazionari del sistema (25), gli altri valori all'equilibrio sono:

$$V = ((1 - \varphi_2) \pi - \varphi_0) / \varphi_1, \quad r = ((1 - Q) \pi + \pi^c) / V \quad (36)$$

ed i punti ammissibili sono punti di equilibrio della mappa (34).

Proposizione 1. Esiste un unico punto di equilibrio ammissibile solo se

$$\varphi_2 < 1 \text{ e } A > 0 \quad (37)$$

dato da

$$\pi^* = \frac{B + \sqrt{B^2 - 4(1 - \varphi_2)AC}}{2(1 - \varphi_2)A} \quad (38)$$

Dimostrazione. Dall'equazione (35) segue immediatamente che $\phi_2 < 1$ è una condizione necessaria. Infatti, se $\phi_2 = 1$ tale equazione ha una sola radice, positiva, ma il corrispondente valore di V all'equilibrio è negativo, mentre se $\phi_2 > 1$, dalla (36) segue che, all'equilibrio, o V o π è negativo, in ogni caso non ammissibile.

Assumiamo quindi $\phi_2 < 1$, interessa ora vedere se le radici dalla (35) sono reali e positive. A questo scopo, riscriviamo le equazioni in (25) nella forma seguente

$$(IS) \quad \begin{cases} \pi_t = \phi_0 + \phi_1 V_t + \phi_2 \pi_{t-1} \\ \gamma_3 r_t = (\phi_0 - \gamma_0) + (\phi_1 - \gamma_1) V_t + (\phi_2 - \gamma_2) \pi_{t-1} \\ r_t V_t = (1-\theta) \pi_{t-1} + \pi^e \end{cases} \quad (39)$$

che ci consente una interpretazione geometrica, utile sia per considerazioni dinamiche che all'equilibrio. Infatti, all'equilibrio, dalla prima equazione si ottiene che l'espressione per V_t data in (36) diventa, sostituendovi V_t ,

$$r = \frac{A}{\gamma_3 \phi_1} \pi - \frac{(\phi_1 \gamma_0 - \phi_0 \gamma_1)}{\gamma_3 \phi_1} \quad (40)$$

$$r\pi = \frac{\gamma_0}{(1-\phi_2)} r + \frac{(1-\theta)\phi_1}{(1-\phi_2)} \pi + \frac{\pi^e \phi_1}{(1-\phi_2)} \quad (41)$$

I valori di equilibrio sono dati geometricamente, nel piano (π, r) , dall'intersezione della retta di equazione (40) con l'iperbole di equazione (41). Siccome tutti i coefficienti dell'equazione (41) sono positivi, il grafico dell'iperbole, i cui asintoti sono le rette di equazione $\pi = \phi_0/(1-\phi_2)$ ed $r = (1-\theta)\phi_1/(1-\phi_2)$, è qualitativamente simile a quello riportato in Fig. 3. Il termine noto nell'equazione della retta (40) è negativo, e se $A \leq 0$ i valori di equilibrio di π sono minori di $\phi_0/(1-\phi_2)$ così che, come prima, o V o π è negativo, ossia inammissibili. Segue che la condizione $A > 0$ è necessaria per avere un punto di equilibrio ammissibile. La sufficienza della (37) segue immediatamente delle proprietà geometriche delle curve (40) e (41) infatti, vi sono due intersezioni reali e solo quella con $\pi > \phi_0/(1-\phi_2)$ è ammissibile. Possiamo quindi concludere che se vale la (37), l'equazione (35) ha due radici reali, la maggiore delle quali, data in (38), fornisce il punto di equilibrio, e la dimostrazione è completa.

In seguito assumeremo che le condizioni in (37) siano soddisfatte.

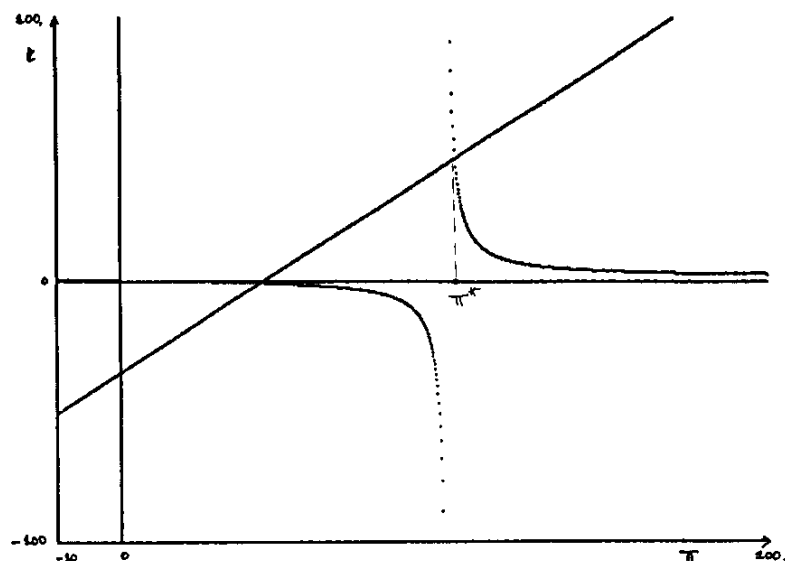


Fig. 3

Osservazione 1. Considerando la famiglia di mappe $f(\pi)$ in funzione del parametro γ_2 , si nota che nell'interpretazione geometrica (40) - (41) del punto di equilibrio ammissibile, il solo parametro che varia è la pendenza della retta. Diminuendo γ_2 , la pendenza aumenta ed il punto di equilibrio π^* (γ_2) diminuisce, al limite si ha

$$\lim_{\gamma_2 \rightarrow -\infty} \pi^*(\gamma_2) = \phi_0 / (1 - \phi_2) \quad (42)$$

3.2 ALCUNE OSSERVAZIONI QUALITATIVE

Il comportamento locale della mappa $f(\pi)$, cioè per condizioni iniziali prossime al valore di equilibrio π^* , può dipendere dalle proprietà di stabilità locali di tale punto. Per esempio, è ben noto che se π^* è stabile (i. e. $|f'(\pi^*)| < 1$), allora gli iterati π_i convergono a π^* monotonicamente o con oscillazioni improprie, a seconda del segno di $f'(\pi^*)$.

Nel caso in esame si ha, da (34) e (33):

$$f'(\pi) = \phi_2 + \phi_1 V'(\pi) \quad (43)$$

dove

$$V'(\pi) = \frac{1}{2} \left\{ \frac{\frac{\varphi_2 - \gamma_1}{\varphi_1 - \gamma_1} [(\varphi_0 - \gamma_0) + (\varphi_2 - \gamma_2) \pi] + 2\gamma_3(1-\theta)}{\sqrt{\Delta(\pi)}} - \frac{\varphi_2 - \gamma_2}{\varphi_1 - \gamma_1} \right\} \quad (44)$$

che deve essere valutata in π^* , dato in (38). Nel seguito useremo anche le seguenti espressioni

$$V'(\pi^*) = \frac{\gamma_3(1-\theta) - V^*(\varphi_2 - \gamma_2)}{\sqrt{\Delta(\pi^*)}} \quad (45)$$

$$V'(\pi^*) = \left\{ \gamma_3(1-\theta) - (\varphi_2 - \gamma_2) [(1 - \varphi_2) \pi^* - \varphi_0] / \varphi_1 \right\} \frac{1}{\sqrt{\Delta(\pi^*)}} \quad (46)$$

che si ottengono dalla (44) usando rispettivamente (33) e (36).

E' chiaro che il segno di $f'(\pi^*)$ non è immediatamente leggibile, tuttavia con ragionamenti qualitativi di tipo geometrico, si può vedere che i due tipi di transitorio (monotono e alternante) possono presentarsi entrambi. Assegnato π_{t-1} , il sistema (39) può leggersi nel piano (V, r) come l'intersezione di una retta con una iperbole, la cui soluzione ammissibile fornisce il valore V_t da usare nell'equazione IS per calcolare π_t , e così via. Se risulta $\pi_t > \pi_{t-1}$ allora nel successivo sistema da risolvere sia l'iperbole che la retta sono (parlando rozzamente), "spostate verso l'alto", ma il nuovo valore di V può essere maggiore o minore del precedente, a seconda delle "deviazioni relative", e quindi π_t può essere maggiore o minore di π_{t-1} . In modo analogo si ragiona se $\pi_t < \pi_{t-1}$.

A scopo illustrativo, consideriamo la famiglia di mappe a un parametro $f(\pi, \gamma_2)$, per fissati valori degli altri parametri, siano

$$\begin{cases} \theta = .5, \pi^c = 1, \varphi_0 = 10, \varphi_1 = .5, \varphi_2 = .8, \\ \gamma_0 = 800, \gamma_1 = -.4, \gamma_3 = .8 \end{cases} \quad (47)$$

In Fig. 4a si può vedere la mappa $f(\pi, -.5)$, il cui punto di equilibrio π^* è stabile con $0 < f'(\pi^*) < 1$, ed in Fig. 4b la mappa $f(\pi, -2)$, in cui π^* è stabile con $-1 < f'(\pi^*) < 0$. L'andamento qualitativo del grafico della mappa $f(\pi)$ in quest'ultimo caso (cioè l'essere una funzione unimodale), suggerisce che il comportamento dinamico della mappa per i valori inferiori di γ_2 può essere più complesso di quello descritto precedentemente. Dalla teoria delle mappe unimodali (Collet-Eckmann (1980), Devaney (1986), Kelsey (1988)), sappiamo che se un punto di equilibrio π^* in funzione di un parametro (nel nostro caso γ_2) biforca con $f'(\pi^*) =$

-1, e sono soddisfatte opportune condizioni di trasversalità (Guckenheimer-Holmes, teorema 3.5.1), allora per valori del parametro vicini al valore di biforcazione il punto di equilibrio è instabile, ed appare un'orbita di periodo due, diciamo C_2 , stabile.

Ci chiediamo se è possibile, nel nostro caso, una biforcazione di questo tipo (flip o raddoppiamento di periodo). Mostriamo che o non esistono, o ne esistono due. Infatti, osservando che per $\gamma_2 \rightarrow -\infty$ si ha $\forall^* > 0$, da (42) e (36), e $\sqrt{\Delta(\pi^*)} \rightarrow \infty$ da (32), si ottiene infine, da (43),

$$\lim_{\gamma_2 \rightarrow -\infty} f'(\pi^*) = \varphi_2$$

Questo limite assicura che per valori di $|\gamma_2|$ sufficientemente grandi, π^* è stabile con $0 < f'(\pi^*) < 1$. Da ciò segue che se esiste un valore di biforcazione flip primario, sia γ_2^F (Forward bifurcation), tale che per $\gamma_2 > \gamma_2^F$ ($\gamma_2 < \gamma_2^F$) si ha $f'(\pi^*) > -1$ ($f'(\pi^*) < -1$), allora deve esistere un altro valore di biforcazione flip primario, sia γ_2^R (Reverse bifurcation), tale che per $\gamma_2 > \gamma_2^R$ ($\gamma_2 < \gamma_2^R$) si ha $f'(\pi^*) < -1$ ($f'(\pi^*) > -1$).

In Appendice A si dimostra che se esiste un'orbita di periodo 2 (stabile o instabile), i due valori dell'orbita, siano $\pi^{(1)}$ e $\pi^{(2)}$, sono le radici della seguente equazione

$$\pi^2 - S\pi + P = 0 \quad (48)$$

dove S e P sono dati in funzione dei parametri del modello.

Nell'esempio dato sopra, la mappa $f(\pi, -3)$ ha un punto di equilibrio instabile con $f'(\pi^*) < -1$, si veda la Fig. 5a, ed in Fig. 5b si mostra la mappa $f^2(\pi, -3)$ con un'orbita C_2 stabile.

Per quanto riguarda i valori di biforcazione si può osservare che ai valori di biforcazione flip, l'equazione (48) deve avere due radici coincidenti, cioè deve essere $S^2 - 4P = 0$, e scrivendo esplicitamente questa relazione di funzione di γ_2 si ottiene un'equazione del tipo

$$\eta_0 \gamma_2^2 + \eta_1 \gamma_2 + \eta_2 = 0 \quad (49)$$

Se il discriminante della (49) è positivo, allora i due valori di biforcazione flip sono le radici di tale equazione.

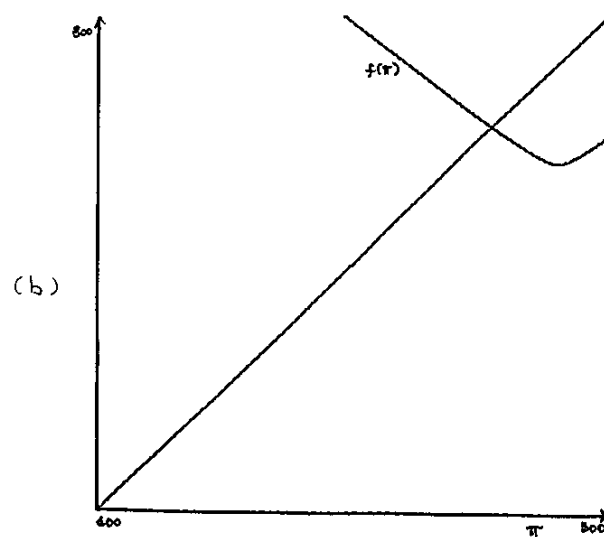
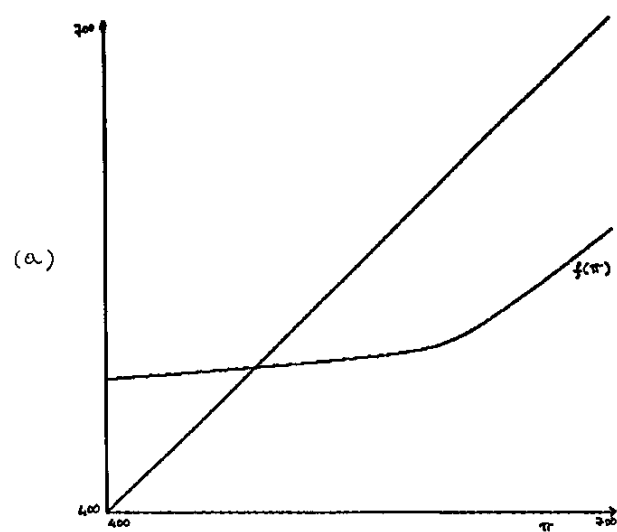


Fig. 4

Continuando con lo stesso esempio, le radici della (49) sono $\gamma_2^F = -2.5$, $\gamma_2^R = -13.16$. Diminuendo γ_2 in questo intervallo si sono trovate le successioni di biforcazioni flip "Forward" dell'orbita C_2 (in Fig. 6a si mostra un'orbita di periodo elevato a $\gamma_2 = -4.25$ ed in Fig. 6b un'orbita della mappa caotica a $\gamma_2 = -5$), di un'orbita C_3 , e di altre ancora, quindi le successioni di biforcazioni flip "Reverse" di diverse orbite fino ad una C_2 per γ_2 prossimo a γ_2^R , ed infine un punto di equilibrio stabile per $\gamma_2 < \gamma_2^R$.

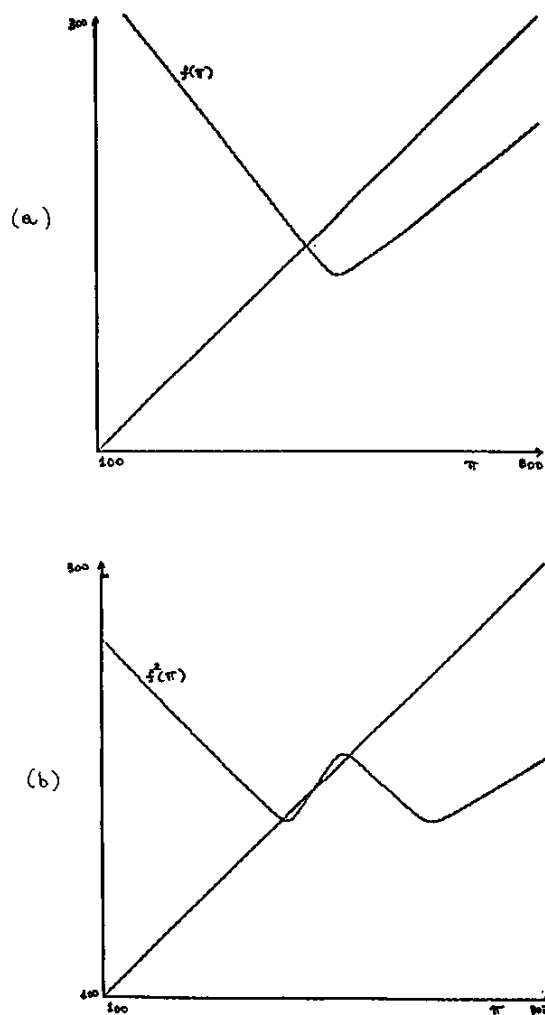


Fig. 5

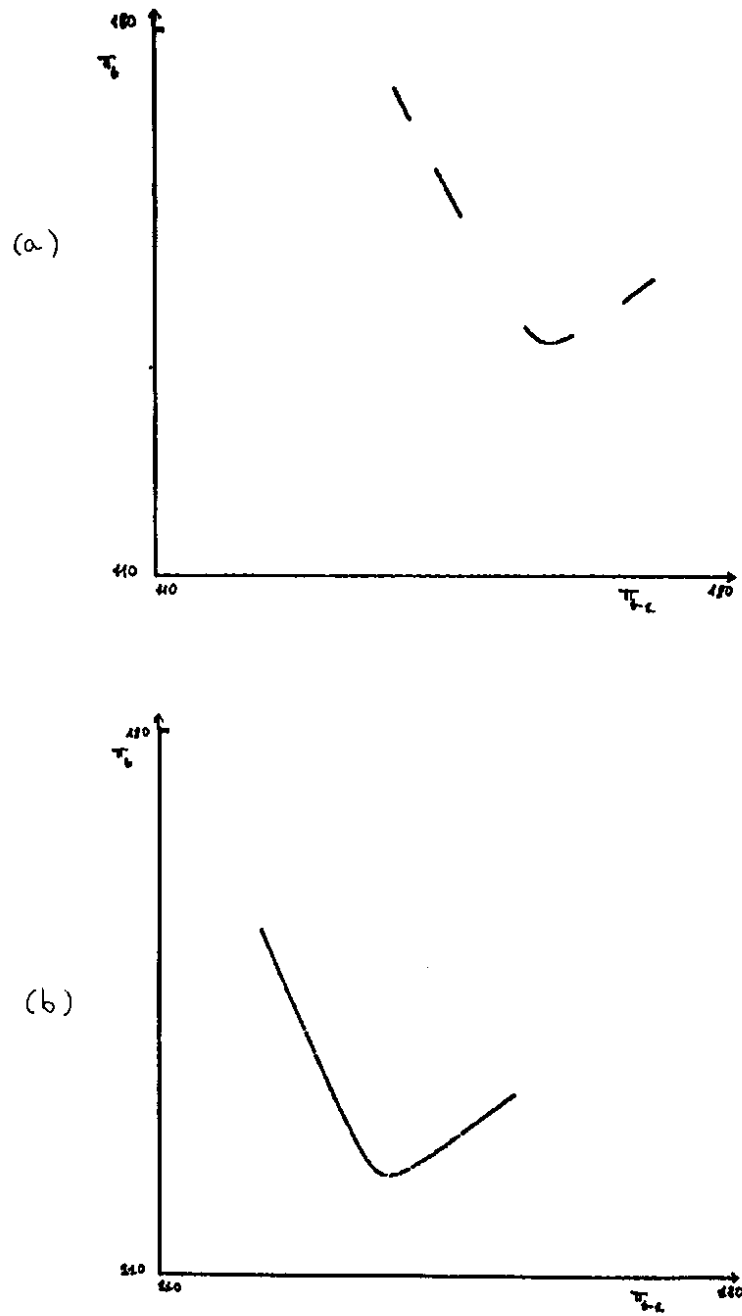


Fig. 6

Ci chiediamo se questo comportamento qualitativo è “generale” per la mappa $f(\pi)$ data in (34), e se dinamiche non standard di questo tipo si abbiano in corrispondenza di valori dei parametri che siano economicamente significativi. Risponderemo rigorosamente alla seconda domanda nella prossima sezione, e qualitativamente alla prima nell’ultima sezione.

3.4 CONDIZIONI DI STABILITÀ

E’ interessante fare un confronto fra il modello lineare discusso in 3.1 e quello non lineare. Per quest’ultimo, le condizioni di esistenza di un punto di equilibrio ammissibile sono:

$$\lambda < 1 \quad \text{e} \quad \varphi_2 < 1 \quad (50)$$

in quanto la condizione $A > 0$ è equivalente alla $\lambda < 1$, che è l’analoga condizione del modello lineare. Il “dominio di esistenza è quindi diminuito, tuttavia la restrizione non è rilevante nel modello in esame, in quanto da un punto di vista economico la condizione $\varphi_2 < 1$ può ritenersi quasi sempre soddisfatta.

Per quanto riguarda la stabilità, supposto che valgano le (50), si è visto che il modello lineare è stabile per $\lambda > -1$, ma economicamente significativo solo per $\lambda > 0$, per via del vincolo (29). Vedremo che, come il caso lineare, anche il modello non lineare ha un punto di equilibrio stabile con transitori di tipo monotono quando i parametri sono economicamente rilevanti. Allo scopo, premettiamo la seguente

Proposizione 2. Assumendo soddisfatte le (26) e (50):

$$\gamma_2 \leq \varphi_2 \Rightarrow \lambda < f'(\pi^*) \quad (51)$$

Infatti, per $\gamma_2 = \varphi_2$ si ha dalla (43), per sostituzione diretta,

$$\varphi_2 = \lambda < f'(\pi^*) = \varphi_2 + \varphi_1 \gamma_3 (1 - \theta) / \sqrt{\Delta(\pi^*)}$$

e per $\gamma_2 < \varphi_2$ si ha, usando (43) e (44),

$$\begin{aligned} \lambda < f'(\pi^*) &\Leftrightarrow V'(\pi^*) > -(\varphi_2 - \gamma_2) / (\varphi_1 - \gamma_1) \\ &\Leftrightarrow \sqrt{\Delta(\pi^*)} + [(\varphi_0 - \gamma_0) + (\varphi_2 - \gamma_2) \pi^*] + 2 \gamma_3 (1 - \theta) \frac{\varphi_1 - \gamma_1}{\varphi_2 - \gamma_2} > 0 \end{aligned}$$

e l’ultima disuguaglianza è soddisfatta.

Dimostriamo ora la seconda condizione sufficiente per la stabilità

asintotica locale di π^* .

Proposizione 3. Assumendo soddisfatte le (26) e (50):

$$-1 + \gamma_1 (1 + \varphi_2) / \varphi_1 < \gamma_2 < \varphi_2 - \varphi_1 \gamma_3 (1 - \theta) \Rightarrow -1 < \lambda < f'(\pi^*) < 1 \quad (52)$$

Infatti, dalla (43) e (45) si ha:

$$\begin{aligned} f'(\pi^*) < \varphi_2 &\Leftrightarrow V'(\pi^*) < 0 \\ &\Leftrightarrow \gamma_3 (1 - \theta) < (\varphi_2 - \gamma_2) V^* \\ &\Leftrightarrow \gamma_3 \varphi_1 (1 - \theta) < (\varphi_2 - \gamma_2) [(1 - \varphi_2) \pi^* - \varphi_0] \end{aligned} \quad (53)$$

ed essendo $\pi^* > \varphi_0 / (1 - \varphi_2)$, abbiamo che

$$\gamma_2 < \varphi_2 - \varphi_1 \gamma_3 (1 - \theta) \Rightarrow f'(\pi^*) < \varphi_2 (< 1)$$

La dimostrazione è completa osservando che

$$\lambda > -1 \Leftrightarrow \gamma_2 > -1 + \gamma_1 (1 + \varphi_2) / \varphi_1$$

Le seguenti osservazioni sono utili per l'interpretazione economica del modello:

- (I) La condizione $\gamma_2 < \varphi_1$ può ritenersi quasi sempre soddisfatta.
 (II) Se nella Proposizione 3 aggiungiamo il vincolo (29), la (52) diventa:

$$\gamma_1 \varphi_2 / \varphi_1 < \gamma_2 < \varphi_2 - \varphi_1 \gamma_3 (1 - \theta) \Rightarrow 0 < \lambda < f'(\pi^*) < 1$$

- (III) Se supponiamo che il vincolo $\gamma_2 < 0$ sia più realistico, allora la disuguaglianza in (53) è soddisfatta, ed una condizione sufficiente di stabilità è che sia $\lambda > -1$ (oppure $\lambda > 0$):

$$\begin{aligned} -1 + \gamma_1 (1 + \varphi_2) / \varphi_1 < \gamma_2 < 0 &\Rightarrow -1 < \lambda < f'(\pi^*) < 1 \\ \gamma_1 \varphi_2 / \varphi_1 < \gamma_2 < 0 &\Rightarrow 0 < \lambda < f'(\pi^*) < 1 \end{aligned}$$

- (IV) Una condizione necessaria per l'esistenza di un'orbita di periodo due (e quindi di dinamiche locali non standard) è che sia $\lambda < -1$.

- (V) Le condizioni sufficienti di stabilità date sopra non dipendono da φ_0 e γ_0 , tuttavia questi parametri influenzano i coefficienti dell'equazione (49), si può quindi concludere che φ_0 e γ_0 influenzano la stabilità di π^* solo nel dominio $\lambda < -1$. Per esempio, se il discriminante di tale equazione è negativo allora ci aspettiamo che non vi siano biforcazioni di tipo flip e quindi $f'(\pi^*) > -1$ per ogni γ_2 ammissibile (il limite superiore di γ_2 è determinato dal vincolo $\lambda < 1$).

3.5 RISULTATI NUMERICI

Riassumiamo brevemente i risultati dell'analisi svolta numericamente sulla famiglia $f(\pi)$ con i vincoli (26), (50) e $\varphi_1 < \varphi_2$. Si è investigato, con esperimenti numerici, il comportamento qualitativo delle famiglie $f(\pi, \gamma_2)$ per diversi valori di $(\gamma_0 - \varphi_0)$, diminuendo γ_2 dall'intervallo di stabilità dato nella Proposizione 3.

Si è trovato che per $(\gamma_0 - \varphi_0)$ inferiore ad un opportuno "valore di soglia", il punto di equilibrio è stabile per ogni $\gamma_2 < 0$. Un esempio con $\gamma_0 = 200$ (gli altri parametri sono dati in (47)) è riportato in Fig. 7a, dove si rappresenta l'intera famiglia $f(\pi, \gamma_2)$ con il grafico della funzione $f'(\pi^*(\gamma_2))$. (Si noti che a questi casi si applicano i ragionamenti fatti nell'osservazione (V) di cui sopra).

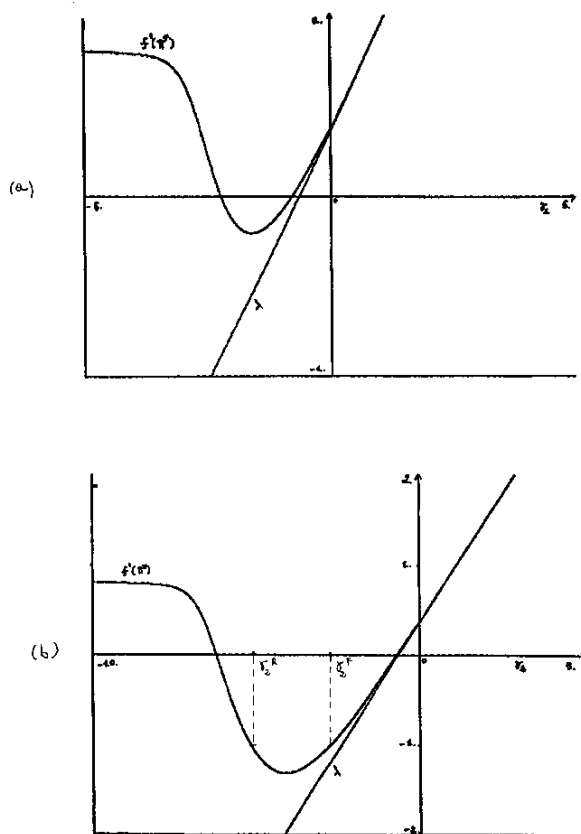


Fig. 7

Aumentando $(\gamma_0 - \phi_0)$ appare ed aumenta l'intervallo di instabilità di π^* con $f'(\pi^*) < -1$ (i.e. $\gamma_2 \in (\gamma_2^R, \gamma_2^F)$), in cui si hanno biforcazioni di tipo flip Forward e Reverse. Il "grado di complessità" raggiunto cresce con le dimensioni dell'intervallo, per esempio, la dinamica associata alla famiglia rappresentata in Fig. 7b per $\gamma_0 = 400$ (gli altri parametri sono dati in (47)) è il seguente:

$$\pi^*, C_2, C_4, C_2, \pi^*$$

cioè, diminuendo γ_2 il punto di equilibrio stabile diventa instabile a γ_2^F ed appare un'orbita di periodo due stabile, che biforca a sua volta, cioè diviene instabile ed appare un'orbita di periodo quattro stabile, seguita da una biforcazione flip Reverse a cui la C_4 diviene instabile ed appare una C_2 stabile e, dopo l'ultima biforcazione a γ_2^R , il punto di equilibrio π^* è di nuovo stabile, ed è stabile per ogni $\gamma_2 < \gamma_2^R$.

In Fig. 8 si riporta la funzione $f'(\pi^*(\gamma_2))$ della famiglia che è stata un esempio nelle precedenti sezioni, in cui è evidente un ampio intervallo (γ_2^R, γ_2^F) , con molte "finestre" di raddoppiamento di periodo e caos in esso.

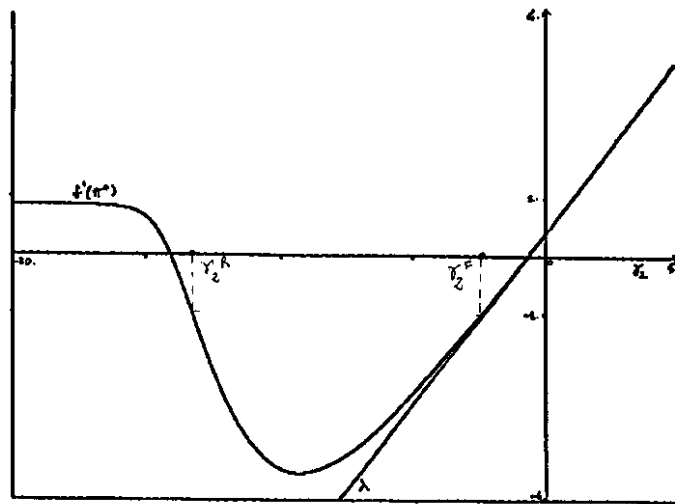


Fig. 8

CONCLUSIONI

Nella tradizione Marshall-Minsky-Kindleberger, l'economia reale è soggetta a vari tipi di shock che possono essere amplificati dal comportamento "non neutrale" del sistema creditizio. A differenza però dei contributi di Hayek e Friedman, il sistema individuato in questo lavoro non presenta quelle situazioni di ottimismo che derivano dalle premesse dei modelli neowalrasiani. Questo perché la moneta viene ad essere analizzata nella duplice veste di "intermediario degli scambi" e di "riserva di valore". Interrompendo la continuità logica e temporale caratteristica dei sistemi alla Arrow-Debreu, i modelli minskiani sono in grado di tener conto di fenomeni di ciclo e, possibilmente, di crisi.

La loro struttura resta, in ogni caso, reale. Shock monetari faranno passare il sistema da zone di stabilità ad altre di instabilità ma se, come sopra ipotizzato, il tasso di interesse viene endogenizzato, la soluzione finale di equilibrio sarà, se pur non ottimale, di tipo classico. Una volta che le autorità bancarie si preoccupino dell'inflazione e della solidità finanziaria del sistema, situazioni di tipo instabile possono emergere. Da questo non segue logicamente la validità dell'assunto di "neutralità delle politiche economiche" nella nuova macroeconomia classica, o la potenziale "dannosità" delle stesse, come sostenuto dalla scuola monetarista. In primo luogo, per entrambi i filoni teorici, essendo la quantità di moneta fissata esogenamente dalla Banca Centrale, aumenti della domanda si ripercuoteranno sul tasso di interesse, determinandone un rialzo: in termini del nostro modello, questo provocherà il passaggio a regioni instabili. In secondo luogo, la politica economica si rivela efficace nel ridurre il "vuoto occupazionale" che si determina in un'economia decentralizzata ogni volta che la domanda aggregata intersechi, come di solito avviene, la curva di offerta ad un livello inferiore a quello di pieno impiego.

In conclusione, pare legittimo ipotizzare una sostanziale stabilità della struttura "reale" dell'economia, se pur non di tipo "steady state". L'instabilità può sorgere una volta che influenze finanziarie, ad esempio relazioni tra cash-flow attesi che devono validare la struttura del debito accumulato, interagiscono con componenti reali. In tal senso non si può sostenere una presunta maggiore stabilità dell'economia reale su quella finanziaria, perché le moderne economie capitalistiche sono del tipo "monetario" dove una separazione tra le due componenti è logicamente scorretta, oltre che storicamente irrealistica.

RINGRAZIAMENTI

Gli autori ringraziano il Prof. H.P. Minsky per gli utili commenti e suggerimenti ad una precedente versione del modello.

APPENDICE A

Una condizione necessaria affinché esista un'orbita di periodo due con valori $\pi^{(1)} = x$ e $\pi^{(2)} = y$, $x \neq y$, è che questi lavori soddisfino il seguente sistema:

$$(A. 1) \quad \begin{cases} y = \varphi_0 + \varphi_1 V_x + \varphi_2 x \\ x = \varphi_0 + \varphi_1 V_y + \varphi_2 y \\ (\varphi_1 - \gamma_1) V_x^2 - [(\gamma_0 - \varphi_0) - (\varphi_2 - \gamma_2) x] V_x - \gamma_3((1 - \theta) x + \pi^e) = 0 \\ (\varphi_1 - \gamma_1) V_y^2 - [(\gamma_0 - \varphi_0) - (\varphi_2 - \gamma_2) y] V_y - \gamma_3((1 - \theta) x + \pi^e) = 0 \end{cases}$$

Eliminando V_x e V_y si ottiene

$$(A. 2) \quad \begin{cases} ay + bx + cxy + dy + ex + g = 0 \\ ax + by + cxy + dx + ey + g = 0 \end{cases}$$

con

$$\begin{aligned} a &= (\varphi_1 - \gamma_1) \\ b &= \varphi_2(\varphi_1\gamma_2 - \varphi_2\gamma_1) \\ c &= -\varphi_2(\varphi_1 - \gamma_1) - (\varphi_1\gamma_2 - \varphi_2\gamma_1) \\ d &= -\varphi_0(\varphi_1 - \gamma_1) - (\varphi_1\gamma_0 - \varphi_0\gamma_1) \\ e &= \varphi_0(\varphi_1\gamma_2 - \varphi_2\gamma_1) + \varphi_2(\varphi_1\gamma_0 - \varphi_0\gamma_1) - \gamma_3(1 - \theta)\varphi_1^2 \\ g &= \varphi_0(\varphi_1\gamma_0 - \varphi_0\gamma_1) - \gamma_3\pi^e\varphi_1^2 \end{aligned}$$

Geometricamente il sistema (A.2) rappresenta l'intersezione di due iperboli, essendo $(ab - c^2/4) = -\varphi_1^2(\varphi_2 - \gamma_2)^2/4 < 0$.

Sottraendo le due equazioni in (A.2) ed eliminando la soluzione ovvia $x=y$ si ha $(x+y) = (e-d)/(a-b)$. Esplicitando x oppure y e sostituendo in una delle (A.2) si ottiene la seguente equazione per $\pi=y$ e $\pi=x$

$$(A.3) \quad \pi^2 - S\pi + P = 0 \quad ; \quad S = \frac{e-d}{a-b}, \quad P = \frac{aS^2 + ds + g}{a + b - c}$$

BIBLIOGRAFIA

1. Arrow K. J. e F. H. Hahn, *General Competitive Analysis*, (Holden Day, San Francisco, 1971).
2. Asimakopoulou A. , "*Finance, Saving and Investment in Keynes' Economics*", Cambridge Journal of Economics, (1985).
3. Chick V., *La Macroeconomia dopo Keynes*, (Il Mulino, Bologna, 1982).
4. Collet P. e J.P. Eckman, *Iterated Maps on the Interval as Dynamical System*, (Birkhauser, Boston, 1980).
5. Debreu G., *Theory of Value*, (John Wiley, 1959)
6. Delli Gatti D., "*Il pensiero economico di Keynes (1923-1936) e l'interpretazione post-keynesiana soggettivista*", tesi di dottorato (non pubblicata), (1987).
7. Devaney R.L., *An Introduction to Chaotic Dynamical System*, (Benjamin, Cummings Publishing, Menlo Park, CA, 1986).
8. Friedman M., "*The Role of Monetary Policy*", American Economic Review, Vol. LVIII, N. 1 (1968).
9. Frydman R. e E. Phelps, *Individual Forecasting and Aggregate Outcomes*, (Cup, Cambridge, 1983).
10. Guckenheimer J. e P. Holmes, *Nonlinear Oscillations Dynamical System and Bifurcations of Vector Fields*, (Spriger-Verlag, N. Y., 1986).
11. Hayek F., *Prices and Production*, (Routledge, Londra, 1931).
12. Kaldor N., *Alternative Theories of Distribution (1955)*. Ristampato in *Essays on Value and Distribution* (Duckworth, Londra, 1960).
13. Kalecki M., *Theory of Economic Dynamics*, (Cup. Cambridge, 1971).
14. Kelsey D., "*The Economics of Chaos or the Chaos of Economics*", Oxford Economic Papers 40 (1988), 1-31.
15. Keynes J.M., *A Treatise on Money*, (Macmillian, Londra, 1930).
16. Keynes J.M., "*Alternative Theories of the Rate of Interest*", Economic Journal (1937 a).
17. Keynes J.M., "*The Ex-Ante Theory of the Rate of Interest*", Economic Journal (1937 b).
18. Keynes J.M., "*Mr. Keynes' Finance*", Economic Journal (1938).
19. Lucas R.E., *Studi sulla Teoria del Ciclo Economico*, (Giuffré, Milano, 1987).
20. Marshall A., *L'Economia della Produzione*, (1879, Tr. it. Isedi, Milano, 1975).
21. Minsky H.P., *John Maynard Keynes*, (Columbia Univ. Press N.Y., 1975).
22. Minsky H.P., *Can "It" Happen Again? Essays on Instability and Finance*, (M.E. Sharpe Armonk, N. Y., 1982).

23. Moore B., *Horizontalism vs Verticalism*, (in corso di stampa 1989).
24. Ohlin B., "Some Notes on the Stockholm Theory of Saving and Investment", *Economic Journal* (1937).
25. Roe F., "The Case for Flow of Funds and National Balance Sheet Accounts", *Economic Journal* (1973).
26. Sargent T., *Macroeconomic Theory* (Academic Press, 1979).
27. Symposium on the Marginal Productivity of Labor, *Journal of Post Keynesian Economics* (1987).
28. Tobin J. e W. Brainard, "Pitfalls in Financial Model Building", *American Economic Review, Papers and Proceedings* (1968).
29. Tobin J., "A General Equilibrium Approach to Monetary Theory", *Journal of Money, Credit and Banking* (1969).

SULLE COALIZIONI IN TEORIA DEI GIOCHI

Stefano Innamorati

Riassunto. In questo lavoro si presentano alcune correlazioni tra le Geometrie Finite e le applicazioni economiche della Teoria dei Giochi.

0. INTRODUZIONE.

Ogni distribuzione di risorse tra i membri di una organizzazione sociale provoca la formazione di coalizioni nel tentativo di risolvere i conflitti dettati da ragioni di interessi particolari o collettivi, nasce così un “gioco” tra gli individui nell’esprimere la propria decisione mirando, in una ampia visione cooperativa, all’obiettivo comune di migliorare la condizione globale di un determinato gruppo per trarne un vantaggio effettivo. Lo studio delle soluzioni di un gioco è di importanza fondamentale per le applicazioni all’analisi dei conflitti di interessi nel campo economico, militare e sociale. Una soluzione di un gioco può essere interpretata, da un punto di vista economico, come la descrizione dei possibili pagamenti che possono sorgere a seconda delle strategie e delle condizioni del mercato, oppure da un punto di vista socio-politico come l’espressione mediante votazione della volontà di una assemblea, in questo contesto, la soluzione equitabile rappresenta l’eguaglianza del voto. L’approccio combinatorico tramite l’impiego di opportune tecniche e di originali metodi di verifica ha assunto particolare rilevanza nell’indagine delle manifestazioni dei fenomeni collettivi. Scopo del presente lavoro è illustrare le mutue correlazioni tra le strutture geometrico-combinatorie e le coalizioni che esse stabiliscono, in particolare si esaminano giochi che derivano dalle geometrie a blocchi e dalla teoria delle matroidi, verificandone la risolubilità.

1. GIOCHI SEMISEMPlici.

Sia $P = \{1, \dots, n\}$ un n -insieme di elementi detti *punti* (o *giocatori*). Indichiamo con 2^P l’insieme delle parti di P ; tale insieme è detto *l’insieme delle*

coalizioni (cf. [1], [13], [15], [22]). Sia $\mathcal{W} \subseteq 2^P$. La coppia $(P, \mathcal{W})$ si dice *Gioco Semisemplice* [14], [20], se valgono i seguenti assiomi dovuti a Shapley (1962) [16]:

$$A1. - \mathcal{W} \neq \emptyset, \mathcal{W} \neq 2^P.$$

$$A2. - \forall W \in \mathcal{W}, \forall A \in 2^P \mid W \subset A \Rightarrow A \in \mathcal{W}.$$

$$A3. - \forall W \in \mathcal{W} \Rightarrow W^c \notin \mathcal{W}.$$

Gli elementi di $\mathcal{W}$ si dicono *coalizioni vincenti*. L'intersezione di tutte le coalizioni vincenti, eventualmente vuota, si dice vertice del gioco e si indica con V .

Gli elementi di $\mathcal{W}^c$ si dicono *coalizioni non vincenti*. La coppia $(P, \mathcal{W}^c)$ è detta il *complementare di un gioco semisemplice*. L'insieme $\mathcal{B}$ delle coalizioni B tali che

$$B, B^c \in \mathcal{W}^c \quad (1.1)$$

è detto *insieme delle coalizioni bloccanti*.

L'insieme $\mathcal{L}$ delle coalizioni L tali che

$$L \in \mathcal{W}^c, L^c \in \mathcal{W} \quad (1.2)$$

viene chiamato *insieme delle coalizioni perdenti*. Risulta dunque:

$$\begin{cases} 2^P = \mathcal{W} \cup \mathcal{B} \cup \mathcal{L}; \\ \mathcal{W} \cap \mathcal{B} = \mathcal{W} \cap \mathcal{L} = \mathcal{B} \cap \mathcal{L} = \emptyset. \end{cases} \quad (1.3)$$

Un gioco si dice *semplice*, nel senso di Von Neumann e Morgenstern, se è privo di coalizioni bloccanti. Ovvero se:

$$L \in \mathcal{W}^c \Leftrightarrow L^c \in \mathcal{W} \quad (1.4)$$

Ovviamente si ha:

$$B \in \mathcal{B} \Leftrightarrow \forall W \in \mathcal{W}, B \cap W \neq \emptyset, B^c \cap W \neq \emptyset. \quad (1.5)$$

In letteratura insiemi che verificano la (1.5) si dicono *blocking sets* rispetto alla famiglia $\mathcal{W}$ (cf. [8], [9], [10], [11], [12]).

La famiglia delle coalizioni bloccanti può essere individuata mediante una sottofamiglia di $\mathcal{W}$, la sottofamiglia delle coalizioni vincenti minimali. Una coalizione vincente W si dice *minimale* se non contiene propriamente alcuna

coalizione vincente. Analogamente una coalizione perdente L si dice massimale se non è contenuta propriamente in alcuna coalizione perdente. Denotiamo con μ l'insieme delle coalizioni vincenti minimali e con η l'insieme delle coalizioni perdenti massimali.

Si prova facilmente il seguente:

LEMMA 1.1 .- *Un insieme $B \in 2^P$ è un blocking set rispetto a $\mathcal{W}$ se e solo se lo è rispetto a μ . In altre parole:*

$$B \in \mathcal{B} \Leftrightarrow \forall W \in \mu \quad B \cap W \neq \emptyset \quad \text{e} \quad B^c \cap W \neq \emptyset. \quad (1.6)$$

Il lemma 1.1 è un teorema di riduzione, esso permette di individuare le coalizioni bloccanti esaminando il comportamento di una coalizione solo rispetto alle minimali.

E' di notevole interesse poter stabilire se una famiglia di sottoinsiemi è l'insieme delle coalizioni perdenti o bloccanti di un qualche gioco, in altri termini se un gioco può essere definito mediante alcune proprietà delle coalizioni perdenti o bloccanti. Al riguardo si prova che:

LEMMA 1.2 .- *Sia P un insieme finito ed $\mathcal{L} \subset 2^P$. Se valgono le*

$$\mathcal{L} \neq \emptyset, \quad \mathcal{L} \neq 2^P; \quad (1.7)$$

$$\forall L \in \mathcal{L}, \quad A \in 2^P : A \subset L \Rightarrow A \in \mathcal{L}; \quad (1.8)$$

$$\forall L \in \mathcal{L}, \quad L^c \notin \mathcal{L} \quad (1.9)$$

allora $\mathcal{L}$ è l'insieme delle coalizioni perdenti di un gioco semisemplice. Viceversa se $\mathcal{L}$ è l'insieme delle coalizioni perdenti di un gioco semisemplice allora $\mathcal{L}$ gode delle proprietà suindicate.

Dimostrazione.- Se $\mathcal{L} \subset 2^P$ verifica le (1.7), (1.8) ed (1.9) allora i complementari degli elementi di $\mathcal{L}$ verificando gli assiomi A1, A2 ed A3 sono coalizioni vincenti di un gioco semisemplice. Viceversa sia $\mathcal{L}$ l'insieme delle coalizioni perdenti di un gioco semisemplice allora la (1.7) e la (1.9) seguono per passaggio al complementare dalle A1 e A3 rispettivamente. Dimostriamo la (1.8). Sia $L \in \mathcal{L}$, $A \in 2^P$ tale che $A \subset L$ e supponiamo per assurdo che $A \notin \mathcal{L}$ allora $A \in \mathcal{W}$ oppure $A \in \mathcal{B}$ in virtù della (1.3). Se $A \in \mathcal{W}$ segue per la A2 $L \in \mathcal{W}$ contro l'ipotesi. Se $A \in \mathcal{B}$ si ha che $W \cap A \neq \emptyset$ per ogni $W \in \mathcal{W}$ ne segue che $W \cap L \neq \emptyset$ per ogni $W \in \mathcal{W}$. Allora due sono le possibilità: o esiste un $W' \in \mathcal{W}$ tale che $W' \subset L$ che implica $L \in \mathcal{W}$, oppure $W \cap L^c \neq \emptyset$ per ogni $W \in \mathcal{W}$ che implica $L \in \mathcal{B}$. In entrambe si contraddice l'ipotesi.

Nell'ambito delle coalizioni vincenti si possono distinguere quelle che contengono propriamente almeno una coalizione bloccante, che chiameremo *fortemente vincenti*, e quelle che non ne contengono alcuna, dette *semplicemente*

vincenti. In modo analogo le coalizioni perdenti contenute in almeno una coalizione bloccante si diranno *coalizioni fortemente perdenti*, le altre *semplicemente perdenti*.

Si dimostra il seguente

LEMMA 1.3. - *Sia P un insieme finito ed $\mathcal{B} \subset 2^P$. Se valgono le*

$$\mathcal{B} \neq \emptyset, \mathcal{B} \neq 2^P; \quad (1.10)$$

$$\forall B, B' \in \mathcal{B}, A \in 2^P: B \subset A \subset B' \Rightarrow A \in \mathcal{B}; \quad (1.11)$$

$$\forall B \in \mathcal{B}, B^c \in \mathcal{B}; \quad (1.12)$$

allora esistono coalizioni fortemente vincenti di un gioco semisemplice per il quale $\mathcal{B}$ è un insieme di coalizioni bloccanti di un gioco semisemplice. Inoltre l'insieme delle coalizioni bloccanti di un gioco semisemplice verifica le (1.10) (1.11) e (1.12).

Dimostrazione. Se $\mathcal{B} \subset 2^P$ verifica le (1.10), (1.11) e (1.12) allora denotiamo con $\mathcal{W}$ gli elementi W di 2^P tali che

$$\begin{cases} \forall B \in \mathcal{B}, W \not\subset B \\ \exists B' \in \mathcal{B} : B' \subset W. \end{cases}$$

L'insieme $\mathcal{W}$ verifica gli assiomi A1, A2 ed A3. Infatti la A1 segue dal fatto che $P \in \mathcal{W}$ e la (1.10) implica che $\mathcal{W} \neq 2^P$. La A2 è ovvia. Dimostriamo la A3. Sia $W \in \mathcal{W}$ e supponiamo per assurdo che $W^c \in \mathcal{W}$. Esiste allora $B' \in \mathcal{B}$ tale che $B' \subset W^c$, che implica $W \subset B'^c$ ma in virtù della (1.12) $B'^c \in \mathcal{B}$ e dunque si contraddice l'ipotesi. Gli elementi di $\mathcal{W}$ sono allora coalizioni fortemente vincenti di un gioco semisemplice. La seconda parte del lemma è ovvia.

Si possono caratterizzare le coalizioni semplicemente perdenti o semplicemente vincenti

LEMMA 1.4. - *Le coalizioni semplicemente vincenti o perdenti sono tutti e soli i blocking sets rispetto alle coalizioni bloccanti.*

Dimostrazione. Sia $\mathcal{B}$ la famiglia delle coalizioni bloccanti di un gioco semisemplice. Denotato con $\mathcal{W}$ l'insieme delle coalizioni vincenti, indichiamo con $\mathcal{W}^s \subseteq \mathcal{W}$ le coalizioni semplicemente vincenti, ovvero

$$\forall B \in \mathcal{B}, W \in \mathcal{W}^s \quad B \not\subset W.$$

Essendo $B \in \mathcal{B}$ si ha

$$\forall B \in \mathcal{B}, W \in \mathcal{W}^s \quad W \not\subset B,$$

onde

$$\forall B \in \mathcal{B}, W \in \mathcal{W}^s \quad B \not\subset W^c.$$

Ne segue

$$\forall B \in \mathcal{B} \quad \mathcal{B} \cap W \neq \emptyset, \quad B \cap W^c \neq \emptyset$$

Le coalizioni semplicemente vincenti sono blocking sets rispetto agli elementi di $\mathcal{H}$. In modo analogo si procede per le coalizioni semplicemente perdenti. Sia ora X un blocking sets rispetto alle coalizioni bloccanti cioè

$$\forall B \in \mathcal{B}, \quad X \cap B \neq \emptyset, \quad X^c \cap B \neq \emptyset.$$

Si ha ovviamente $X \notin \mathcal{B}$. Se $X \in \mathcal{W}$ allora X non potendo contenere coalizioni bloccanti è una coalizione semplicemente vincente. In modo analogo si ha che se $X \in \mathcal{L}$ allora X è una coalizione semplicemente perdente.

Ci poniamo il seguente problema: esistono strategie per bloccare la vincita a due gruppi disgiunti?

LEMMA 1.5 .- *Sia P un insieme finito. Ogni sottoinsieme proprio $\mathcal{C}$ di 2^P chiuso rispetto all'operazione di complementazione è un sottoinsieme di coalizioni bloccanti di un gioco semisemplice se e solo se $\emptyset, P \notin \mathcal{C}$.*

Dimostrazione. Sia $\mathcal{C} \subset 2^P$, $\mathcal{C} \neq \emptyset$ e $\mathcal{C} \neq 2^P$ tale che:

$$\begin{aligned} \forall H \in \mathcal{C} \quad &\Rightarrow \quad H^c \in \mathcal{C}. \\ \emptyset, P &\notin \mathcal{C}. \end{aligned}$$

Denotiamo con $\mathcal{W}$ gli elementi W di 2^P tali che:

$$\begin{cases} \forall H \in \mathcal{C} & W \not\subset H \\ \exists K \in \mathcal{C} & : \quad K \subset W. \end{cases}$$

L'insieme $\mathcal{W}$ verifica gli assiomi A1, A2 ed A3. Dimostriamo la A1. $P \notin \mathcal{C}$ implica $P \in \mathcal{W}$ allora $\mathcal{W} \neq \emptyset$. Essendo $\mathcal{C} \neq \emptyset$ si ha $\mathcal{W} \neq 2^P$. La A2 è ovvia. Dimostriamo la A3. Sia $W \in \mathcal{W}$ supponiamo per assurdo che $W^c \in \mathcal{W}$. Esiste allora $K \in \mathcal{C}$ tale che $K \subset W^c$, che implica $W \subset K^c$. Essendo $K^c \in \mathcal{C}$ si contraddice l'ipotesi. L'insieme $\mathcal{W}$ è dunque un insieme di coalizioni fortemente vincenti di un gioco semisemplice. Viceversa se $\emptyset, P$ appartenessero a $\mathcal{C}$ il gioco risulterebbe privo di coalizioni vincenti onde l'asserto.

La seconda parte del presente paragrafo è dedicata all'analisi di alcuni esempi concreti alla luce della risolubilità del gioco.

Una soluzione § semplice principale normalizzata di un gioco $(P, \mathcal{W})$ è una n -pla ordinata di numeri reali non negativi tali che:

$$\S = (x_1, \dots, x_n) \quad : \quad \begin{cases} \sum_{i \in W} x_i = 1 & \forall W \in \mu \\ \sum_{i \in H} x_i > 1 & \forall H \in \mathcal{W} \cup \mathcal{B} - \mu. \end{cases}$$

Una soluzione si dice equitabile se $x_i = x_j$ con $i \neq j$, ([17], [20]).

E' noto, cf. [3], [11], [14], che:

RISULTATO 1.6 .- *Un gioco semisemplice ha soluzioni semplici principali solo se:*

- (1) *non esistono coalizioni fortemente vincenti minimali;*
- (2) *devono esserci almeno due coalizioni vincenti che si intersecano in esattamente un punto;*
- (3) *il vertice del gioco deve avere al più un punto.*

ESEMPIO 1.7 .- Sia S un n -insieme finito e V un sottoinsieme proprio e non vuoto di S . Indichiamo con $\mathcal{W}$ il sottoinsieme dell'insieme delle parti di S i cui elementi sono tutti e soli i sottoinsiemi di S che contengono V . L'insieme $\mathcal{W}$ verifica banalmente gli assiomi A1, A2, ed A3 e la coppia $(S, \mathcal{W})$ è un gioco semisemplice. Il vertice del gioco è ovviamente V . La coppia $(S, \mathcal{W})$ è un gioco semplice se e solo se $|V|=1$. Infatti supponiamo che la coppia $(S, \mathcal{W})$ non sia un gioco semplice allora $\mathcal{B} \neq \emptyset$. Esiste dunque almeno un sottoinsieme B di S tale che $B \cap V \neq \emptyset$ e $B^c \cap V \neq \emptyset$. Ne segue che V ha almeno due elementi distinti. Viceversa supponiamo che V abbia almeno due elementi distinti e fissiamo un $v \in V$ l'insieme $\{v\}$ è banalmente una coalizione bloccante e pertanto la coppia $(S, \mathcal{W})$ non è un gioco semplice. Il gioco ammette soluzioni se e solo se è un gioco semplice. Infatti in base alla (3) del risultato 1.6 ammette soluzione solo se $|V|=1$. Viceversa se è un gioco semplice si ha $\mu=V=\{v\}$ e la n -pla $(1, 1, \dots, 1)$ è una soluzione del gioco.

ESEMPIO 1.8 .- Sia X un insieme finito. Sia $P=2^X$ l'insieme dei giocatori. Poiché l'insieme $\mathcal{L}$

$$\mathcal{L} = \{A \in 2^P : \forall H, K \in A \quad H \cap K \neq \emptyset\},$$

soddisfa alle (1.7), (1.8) ed (1.9) definiamo $\mathcal{L}$ come l'insieme delle coalizioni perdenti di un gioco semisemplice. Le coalizioni bloccanti sono le famiglie B

di parti di X tali che:

$$\exists H, H' \in B, K, K' \in B^c : H \cap H' = \emptyset, K \cap K' = \emptyset.$$

Poiché se $A \in \mathcal{L}$ allora $\emptyset \notin A$ risulta che $\emptyset$ è l'elemento di ogni coalizione vincente. Inoltre se E è un elemento del vertice del gioco si ha

$$\forall A \in \mathcal{L} E \notin A \Rightarrow \{E\} \notin \mathcal{L} \Rightarrow \exists H, K \in \{E\} H \cap K = \emptyset \Rightarrow E \cap E = E = \emptyset.$$

Pertanto il gioco ha per vertice $\{\emptyset\}$ ed è un caso particolare dell'esempio 1.7.

ESEMPIO 1.9. - Sia X un n -insieme finito. Sia $P=2^X$ l'insieme dei giocatori. Poiché l'insieme $\mathcal{L}$ delle famiglie concatenate di parti di X :

$$\mathcal{L} = \{A \in 2^P : \forall H, K \in A \quad H \subseteq K\}$$

soddisfa alle (1.7), (1.8) ed (1.9) definiamo $\mathcal{L}$ come l'insieme delle coalizioni perdenti di un gioco semisemplice. Le coalizioni bloccanti sono le famiglie B di parti di X tali che:

$$\exists H, H' \in B, K, K' \in B^c : H \neq H \cap H', K \neq K \cap K'.$$

Tale gioco non ha soluzioni. Infatti sia $C \in P$ con $\emptyset \neq C \neq X$. Ovviamente $\{C, C^c\}$ è una coalizione bloccante. Siano $c_0 \in C, c_1 \in C^c$. La famiglia di parti concatenate di X costruita nel seguente modo:

$$\emptyset \subset \{c_0\} \subset \{c_0, c_1\} \subset \{c_0, c_1, c_2\} \subset \dots \subset \{c_0, c_1, c_2, \dots, c_{n-1}\} = X$$

dove $c_i \neq c_j$ per $i \neq j$; è un elemento massimale di $\mathcal{L}$. La famiglia complementare è una coalizione vincente minimale che contiene la coalizione bloccante $\{C, C^c\}$. Pertanto in virtù della (I) del risultato 1.6 il gioco non ha soluzioni.

ESEMPIO 1.10. - Sia $P=2^X$ l'insieme dei giocatori come nell'esempio 1.9. La cardinalità della differenza simmetrica definisce una distanza in P :

$$\forall A, B \in P \quad d(A, B) = |A \cup B| - |A \cap B|.$$

Il diametro di un sottoinsieme di P è il massimo delle distanze fra i suoi elementi. Poiché l'insieme $\mathcal{L}$ delle famiglie di parti di X aventi diametro minore di $|X|/2$:

$$\mathcal{L} = \{S \in 2^P : \text{diam}(S) < |X|/2\},$$

soddisfa alle (1.7), (1.8) ed (1.9) definiamo $\mathcal{L}$ come l'insieme delle coalizioni perdenti di un gioco semisemplice. Le famiglie di parti con diametro $|X|/2$ sono le coalizioni bloccanti. Il gioco ammette soluzioni solo se $|X|$ è dispari.

2. GIOCHI DI BLOCCHI.

Sia P un insieme finito e β una famiglia di parti di P , la coppia (P, β) si dice *spazio di blocchi*. Gli elementi di S e di β si dicono rispettivamente *punti* e *blocchi*, [2], [7], [4], [18].

In virtù delle A2 e A3 le coalizioni vincenti formano quella che si chiama una famiglia di intersezione cioè una famiglia di parti a due a due intersecantisi:

$$\forall W, W' \in \mathcal{W} \Rightarrow W \cap W' \neq \emptyset. \quad (2.1)$$

Sia (P, β) uno spazio di blocchi. Se la famiglia β verifica la (2.1) si può definire un gioco semisemplice $(P, \beta+)$, dove $\beta+$ è la famiglia dei soprainsiemi di β (gli elementi di 2^P che contengono almeno un elemento di β), che si dice *gioco di blocchi*.

Uno spazio proiettivo $PG(r, q)$ di dimensione r ed ordine q rispetto all'insieme β dei sottospazi di dimensione h dove $r/2 \leq h \leq r$ è uno spazio di blocchi (P, β) verificante la (2.1). La coppia $(P, \beta+)$ è un gioco di blocchi. Esiste un'ampia bibliografia in [2] sullo studio delle coalizioni bloccanti. Sulla risolubilità relativamente al caso $h=r-1$ si può provare generalizzando un risultato contenuto in [3], che

LEMMA 2.1. - *Il gioco di blocchi generato dalla famiglia degli iperpiani di uno spazio proiettivo finito ammette un'unica soluzione e precisamente la soluzione equitabile.*

Dimostrazione. Sia $(P, \mathcal{W})$ il gioco di blocchi costituito dai punti di uno spazio proiettivo $PG(r, q)$ di dimensione r ed ordine q e dai soprainsiemi degli iperpiani. Esso ammette l'unica soluzione equitabile in virtù del teorema 6.1 di [3]. Inoltre le componenti di una soluzione devono soddisfare il sistema di θ_r equazioni in θ_r incognite

$$\left\{ \sum_{i \in W} x_i = 1 \quad W \in \mu; \right.$$

la cui matrice dei coefficienti è la matrice di incidenza dello spazio che ha determinante diverso da zero, cf. [4].

ESEMPIO 2.2 .- Denotiamo con P l'insieme dei punti di uno spazio affine $AG(r, q)$ di dimensione r ed ordine q . Poiché in esso nessuna famiglia di sottospazi gode della proprietà (2.1) consideriamo una famiglia di intersezione. Ad esempio possiamo prendere in esame la famiglia β delle unioni di due sottospazi non paralleli di dimensione h e k dove $h+k=r$. Supponiamo $h \leq k$. La coppia $(P, \beta+)$ è un gioco di blocchi. Tra le coalizioni bloccanti vi sono gli h -blocking sets, i k -blocking sets (cf. [5], [6], [19]) e le unioni di un h -blocking set con un sottospazio di dimensione k . Poiché un elemento di β privo di un punto è una coalizione bloccante, ogni coalizione in β è fortemente vincente. Pertanto il gioco non è risolubile in base al risultato 1.6.

ESEMPIO 2.3 .- Sia (P, β) uno spazio di blocchi, $x \in P$, denotiamo con (x) l'insieme dei blocchi passanti per x . Si definisce una distanza δ tra i punti:

$$\forall x, y \in P \quad \delta(x, y) = |(x) \cup (y) - (x) \cap (y)|.$$

Indichiamo con $\text{diam}(A)$ il diametro di un sottoinsieme A di punti ovvero il massimo delle distanze dei suoi elementi. Sia $\text{diam}(P)=p$. L'insieme $\mathcal{L}$ di parti di P aventi diametro minore di un numero intero $d < p/2$ soddisfa alle (1.7), (1.8) ed (1.9). Si può assumere $\mathcal{L}$ come l'insieme delle coalizioni perdenti di un gioco semisemplice. Il gioco così definito è risolubile solo se p è dispari.

ESEMPIO 2.4 .- Sia (P, β) uno spazio di blocchi, $S \subset P$ si dice di classe $[m_1, \dots, m_b]$ se $\forall B \in \beta \mid B \cap S \mid \in \{m_1, \dots, m_b\}$. Denotiamo con b la cardinalità massima di un blocco. Sia k un intero tale che $1 < k < b/2$, gli insiemi di classe $[1, \dots, k]$ in virtù del lemma 1.2 definiscono un gioco semisemplice. Le coalizioni vincenti sono gli insiemi tali che ogni blocco interseca in almeno $b-k$ punti. Le coalizioni bloccanti sono quelle tali che esiste almeno un blocco che interseca in s punti con $s \leq [b/2]$ ed esiste almeno un blocco che interseca in t punti con $t \geq [b/2]$. Ogni coalizione vincente è fortemente vincente. Infatti sia W una coalizione vincente. Si può definire il minimo numero m di punti di intersezione tra W ed i blocchi di (P, β) . Si ha $m \geq [b/2]$. Sia $B \in \beta$ un blocco tale che $|B \cap W| = m$. Indichiamo con $p_1, p_2, \dots, p_m$ i punti di $B \cap W$. L'insieme $W - \{p_1, \dots, p_u\}$ dove $u = m - [b/2]$ è una coalizione bloccante del gioco $(P, \beta+)$ poiché il blocco B interseca $W - \{p_1, \dots, p_u\}$ in esattamente $[b/2]$ punti. In virtù del risultato 1.6 il gioco non è risolubile.

3. GIOCHI SEMIMATROIDALI.

Sia P un insieme finito ed $\mathcal{F}$ una famiglia di parti di P . La coppia $(P, \mathcal{F})$ si dice matroide se valgono i seguenti assiomi dovuti a Whitney (cf. [21]):

I1. - $\mathcal{F} \neq \emptyset$

I2. - $\forall I \in \mathcal{I}, \forall J \in 2^P : J \subset I \Rightarrow J \in \mathcal{I}$

I3. - $\forall I, J \in \mathcal{I}, X \in 2^P : I, J \subseteq X, \text{ massimali in } X \Rightarrow |I| = |J|$.

Gli elementi di $\mathcal{I}$ si dicono indipendenti della matroide. Gli elementi di $2^P - \mathcal{I}$ si dicono dipendenti.

In [2] si definisce gioco matroidale un gioco semisemplice il cui gioco complementare è una matroide. Se i complementari degli indipendenti non sono indipendenti una matroide definisce un gioco semisemplice, che chiameremo gioco semimatroidale, assumendo gli elementi di $\mathcal{I}$ come coalizioni perdenti. Ovviamente si ha:

LEMMA 3.1 .- *Un gioco semimatroidale è matroidale se e solo se è privo di blocking sets.*

ESEMPIO 3.2 .- Sia V uno spazio vettoriale su un campo K . Sia $\mathcal{A}$ l'insieme delle parti di V costituite da vettori linearmente indipendenti. La coppia $(V, \mathcal{A})$ è una matroide. Indicato con $\mathcal{W}$ l'insieme delle parti di V costituite da vettori dipendenti complementari di vettori indipendenti, la coppia $(V, \mathcal{W})$ è un gioco semimatroidale. Le coalizioni bloccanti sono i sottoinsiemi di V tali che essi ed i propri complementari sono formati da vettori dipendenti. Ogni sottospazio è una coalizione bloccante. E' un esempio di gioco mai semplice essendo il vettore nullo una coalizione bloccante. Sia W una coalizione perdente. Sia $\overline{W}$ il sottospazio generato da W . Si ha $\overline{W}^c \subset W^c$. Ne segue che ogni coalizione vincente è fortemente vincente. Pertanto per la (1) del risultato 1.6 il gioco non ha soluzioni.

ESEMPIO 3.3 .- Sia V un insieme di n vertici. Supponiamo inoltre che $n \geq 5$. Sia $\mathcal{G}$ l'insieme dei grafi su V . Sia $\mathcal{A} \subset \mathcal{G}$ l'insieme dei grafi privi di cicli su V . Denotato con E l'insieme di tutti i lati su V risulta che $|E| = n(n-1)/2$. Essendo $n \geq 5$ i complementari degli elementi di $\mathcal{A}$ non appartengono ad $\mathcal{A}$. Infatti un grafo su n vertici senza cicli ha al più $n-1$ lati. Pertanto il complementare ha almeno $n(n-1)/2 - (n-1) = (n-1)(n-2)/2$ lati. Tale numero è maggiore di $n-1$ per $n \geq 5$. Ne segue che indicato con $\mathcal{W}$ l'insieme dei grafi che contengono almeno un ciclo la coppia $(E, \mathcal{W})$ è un gioco semimatroidale. Le coalizioni bloccanti sono i grafi con cicli i cui complementari hanno cicli. Un grafo G privo di cicli su V avendo al più $n-1$ lati ha un grafo G' con n lati che lo contiene. Poiché G' è una coalizione bloccante e $G'^c \subset G^c$, si ha che ogni coalizione vincente del gioco è fortemente vincente. Il gioco pertanto non ha soluzioni in virtù della (1) del risultato 1.6.

BIBLIOGRAFIA

1. K.J.ARROW, *Social Choise and Individual Values*, J. Wiley and Sons, New York, 1951, 2nd edition, 1963.
2. L.BERARDI-F.EUGENI, *Blocking sets e Teoria dei Giochi*, Atti Sem. Mat. Fis. Univ. Modena 34 (1988), 165-196.
3. L.BERARDI-F.EUGENI, *Blocking sets and game theory II*, Proc. "Blocking sets", Giessen 1989, Springer verlag to appear.
4. T.BETH-D.JUNGnickel-M.LENZ, *Design Theory*, Wissenschaftsverlag, Mannheim-Wien-Zurich, 1985.
5. A.BEUTELSPACHER, *Blocking sets and partial spreads in finite projective space*, Math.Z. 145 (1975), 211- 229.
6. A.BRUEN, *Blocking sets in finite projective planes*, SIAM J. Appl. Math. 23 (1971), 380-392.
7. M.CERASOLI-F.EUGENI-M.PROTASI, *Introduzione alla Matematica Discreta*, Zanichelli Bologna, 1988.
8. J.DIPAOLA, *On a restricted class of block design games*, Canad. J. Math. 18 (1966), 225-236.
9. J.DI PAOLA, *On minimum blocking coalitions in small projective plane games*, SIAM J. Appl. Math. 17 (1969), 378-392.
10. G.EBERT, *Blocking sets in projective spaces*, Canad. J. Math. 4 (1978), 856-862.
11. A.J.HOFFMAN-M.RICHARDSON, *Block design games*, Canad. J. Math. 13 (1961), 110-128.
12. F.MAZZOCCA-G.TALLINI, *On the non existence of blocking sets in PG (n, q) and AG (n, q) for all large enough n*, Simon Stelvin I (1985), 43-50.
13. G.PICKERT, *Abstimmungssysteme und taktische konfigurationen*, Atti Convegni Lincei: "Teorie Combinatorie", Roma 1973, vol. II, 17 (1976), 235-248.
14. M.RICHARDSON, *On finite projective games*, Proc. Amer. Math. Soc., 7 (1956), 458-465.
15. J.ROSENMULLER, *Some topics of cooperative Game Theory*, in Modern Applied Math. - Opti. & Op. Res.- North Holland 1982.
16. L.S.SHAPLEY, *Simple Games - An outline of the Theory*, Rand Corporation P-series Report, 1962.
17. G.P.SZEGO-G.GAMBARELLI, *Discontinuous solutions in n-persons Games*, in "New quantitative techniques for Economic Analysis", Accademic Press 1982.
18. G.TALLINI, *Geometrie d'incidenza e matroidi*, IAC Roma, Quad. Serie III - N. 127 (1981), 1-34.
19. G.TALLINI, *On blocking sets in finite projective and affine spaces*, Annals of Discrete Math. 37 (1988), 433-450.

20. J.VONNEUMANN-O.MORGENSTERN, *Theory of Games and Economic Behavior*, Princeton Univ. Press, Princeton, 2^d ed., 1947.
21. H.WHITNEY, *On the abstract properties of linear dependence*, Amer. J. Math. 57, 507-533 (1935).
22. H.ZEITLER, *Koalitionbildung bei Abstimmungen*, Math. Phys. Semesterberichte 20 (1973), 86-113.

ON IRREDUCIBLE BLOCKING SETS IN PROJECTIVE PLANES (*)

Stefano Innamorati and Antonio Maturo

Summary. - In a paper of Bruen and Silverman [7], it is proved that in a Desarguesian projective plane of square order q , $q > 4$, in the interval of the admissible cardinalities of irreducible blocking sets there are integers k such that there is no irreducible blocking set with k points. In this paper we prove that in a finite projective plane there is a sub-interval in which for any integer k there is at least one irreducible blocking set with k points.

1. INTRODUCTION

Throughout this note, we denote by $\pi = \pi_q$ a finite projective plane of order q , where q is not necessarily a power of a prime.

A blocking set of π is a set K of points which contains no line but intersects every line. A blocking set is said to be irreducible if it contains no blocking set properly, otherwise it is said to be reducible. The index of a blocking set K is the minimum number of lines whose union contains K .

The following results are well-known (see [5], [6], [11]).

1.1 RESULT .- *Let K be an irreducible blocking set in π . Then*

$$(a) \quad q + \sqrt{q} + 1 \leq |K| \leq q\sqrt{q} + 1.$$

$$(b) \quad |K| = q + \sqrt{q} + 1 \text{ iff } q \text{ is a square and } K \text{ is a Baer subplane.}$$

(*) Communicated at Conference "Giornate di Geometrie Combinatorie"
Pescara, Marzo 1989.

(c) $|K| = q\sqrt{q} + 1$ iff q is a square and K is a Hermitian arc.

Examples and other results concerning the blocking sets in finite projective planes can be found in [1], [2], [4], [8], [9].

In particular, in [1] it is proved that:

1.2 RESULT .- *Let π be a Desarguesian projective plane of order q , where $q=p^h$, p a prime, $q>2$. Let $m(q)$ be the function defined as follows:*

$$m(q) = \begin{cases} \sqrt{q} & \text{if } q \text{ is a square;} \\ (q+1)/2 & \text{if } q \text{ is a prime;} \\ p^{h-d} & \text{otherwise, where } d \text{ denotes the} \\ & \text{greatest divisor of } h \text{ different from } h. \end{cases}$$

Then for any integer k , with $q+m(q)+1 \leq k \leq q^2-m(q)$, there exists a blocking set with k points.

An obvious question is, whether there exist irreducible blocking sets in π for each cardinality belonging to the interval $(q+\sqrt{q}+1, q\sqrt{q}+1)$. In [7] is proved the following:

1.3 RESULT .- *If π is a Desarguesian projective plane, q square, $q>4$, then in the interval $(q+\sqrt{q}+1, q+\sqrt{2q}+1-1/(2q))$ there is no irreducible blocking set.*

In order to give an answer to this question we shall prove the following assertions:

(I) *For $q>4$, a finite projective plane has at least one irreducible blocking set of index 4 of cardinality k for any integer k with $2q-1 \leq k \leq 3q-5$.*

(II) *For $q>4$, a Desarguesian projective plane has at least one irreducible blocking set of index 4 of cardinality k for any integer k with $2q-1 \leq k \leq 3q-3$.*

We note that a complete characterization of irreducible blocking sets in the case $q=3$ can be found in [11], in the case $q=4$ in [2], in the case $q=5$ in [3]. Moreover we recall the following two results, see [8], [10], on the lower bound of the interval of (II).

1.4 RESULT .- *In a Desarguesian projective plane of order q , for every proper divisor d of q or of $q-1$, there exists an irreducible blocking set of index 3 having $2q+1-d$ points.*

1.5 RESULT. - *If in a Desarguesian projective plane of order q there is an irreducible blocking set B of index 3 whose cardinality is less than or equal to $2q-1$, then B has $2q+1-d$ points exactly, where d is a proper divisor of q or of $q-1$.*

2. IRREDUCIBLE BLOCKING SETS

We begin with the following

2.1 THEOREM. - *In a finite projective plane of order q , $q \geq 4$, for any integer k with $2q-1 \leq k \leq 3q-5$ there exists at least one irreducible blocking set of index 4 having k points.*

Proof.- Let π and σ be two lines of a finite projective plane, let V be their intersection point, let $R_i, i=0, 1, \dots, q-1$, be the points on π different from V and let $S_j, j=0, 1, \dots, q-1$, be the points on σ different from V . Denote by ℓ the line $R_0 S_0$. The set

$$K = \pi \cup \sigma \cup \ell - \{V, R_0, S_0\}$$

is a well-known irreducible blocking set, with $3(q-1)$ points called a triangle without vertices. Let $T_r, r=1, 2, \dots, q-1$, be the points of ℓ distinct from R_0 and S_0 . Consider an arbitrary point S on σ , different from V and S_0 , and an arbitrary point T_1 on ℓ , different from R_0 and S_0 . Denote by R the intersection point of the line ST_1 with π . Put

$$U_1 = R_0 S \cap VT_1,$$

$$U_2 = R_0 S \cap RS_0.$$

There are two possible cases:

(a) $U_1 = U_2$;

(b) $U_1 \neq U_2$.

Let n be an integer such that $0 \leq n \leq q-3$ if case (a) holds and $0 \leq n \leq q-4$ if case (b) holds. Moreover let $U_{2+i}, i=1, \dots, n$ be n arbitrary points of $R_0 S$, distinct two by two, and different from R_0, S, U_1, U_2 . Denote by R_i and by T_i the intersection points of the line $S_0 U_i$ with π and ℓ respectively, for any $i=1, 2, \dots, n+2$. The set

$$K' = K - \{R_1, R_2, \dots, R_{n+2}, T_1, T_2, \dots, T_{n+2}\} \cup \{U_1, U_2, \dots, U_{n+2}\}$$

is an irreducible blocking set of index 4 with $3(q-1)-(n+1)$ points if case (a) holds and with $3(q-1)-(n+2)$ if case (b) holds. If case (a) holds, since $0 \leq n \leq q-3$, we obtain irreducible blocking sets of cardinality $3(q-1)-r$ for any integer r such that $1 \leq r \leq q-2$. If case (b) holds, since $0 \leq n \leq q-4$, we have irreducible blocking sets of cardinality $3(q-1)-r$ for any integer r such that $2 \leq r \leq q-2$.

Now we deal with Desarguesian case.

2.2 COROLLARY. - *In a Desarguesian projective plane of even order greater than 2 for any integer k with $2q-1 \leq k \leq 3q-3$ there exists at least one irreducible blocking set of index 4 having k points.*

Proof. - The assertion follows by the proof of the previous theorem and by observing that in a Desarguesian projective plane the order is even if and only if case (a) holds.

Finally we prove the following:

2.3 PROPOSITION - *In a Desarguesian projective plane of odd order greater than 3 for any integer k with $2q-1 \leq k \leq 3q-3$ there exists at least one irreducible blocking set of index 4 having k points.*

Proof. - By using the same notation as in the proof of the previous theorem, since in a Desarguesian projective plane the order is odd if and only if case (b) holds, we prove that if $U_1 \neq U_2$ it is possible to construct an irreducible blocking set of index 4 with $3(q-1)-1$ points, so the assertion follows by the proof of the theorem. Let us denote by $R_i, i=1, 2$, the intersection points of the line $S_0 U_i$ with π and by $T_i, i=1, 2$, the intersection points of the line $U_i V$ with π . Put $Q = S_0 U_2 \cap T_1 V$ and $S'' = R_0 Q \cap \pi$. The set

$$K'' = K - \{T_1, R_1, R_2\} \cup \{U_1, Q\}$$

is a blocking set of index 4 with $3(q-1)-1$ points. In order to prove that K'' is irreducible, since the line ST_1 contains R_2 , it is sufficient to prove that the line $S'' T_1$ passes through the point R_1 . Put

$$X = S'' T_1 \cap \pi.$$

Denote by $(ABCD)$ the cross-ratio of four collinear points A, B, C and D and by (P) the perspectivity of centre a point P . It results:

$$(VR_2 R_0 X) \stackrel{(T_1)}{=} (VSS_0 S'') = (S'' S_0 SV).$$

Moreover we have:

$$(VR_2R_0R_1)^{(U_1)} = (QR_2U_2S_0)^{(S)} = (QT_1U_1V)^{(R_0)} = (S''S_0SV)$$

It follows that

$$(VR_2R_0R_1) = (VR_2R_0X)$$

and then

$$R_1 = X.$$

REFERENCES

1. L.Berardi-F.Eugeni, *On the cardinality of blocking sets in PG(2,q)*, J. Geom. 22 (1984), 5-14.
2. L.Berardi-F.Eugeni, *Blocking sets in projective plane of order four*, Annals of Discrete Math. 37 (1988), 43-50.
3. L.Berardi-S.Innamorati, *Irreducible blocking sets in a projective plane of order five*, Atti Sem. Mat. Fis. Univ. Modena, (to appear).
4. A.Beutelspacher-F.Eugeni, *Sui blocking sets di dato indice con particolare riguardo all'indice tre*, Boll. U.M.I., 4 A (1985), 441-450.
5. A.A.Bruen, *Blocking sets in projective planes*, Siam. J. Appl. Math. 21 (1971), 380-392.
6. A.A.Bruen-J.A.Thas, *Blocking sets*, Geom. Dedicata 6 (1977), 193-203.
7. A.A.Bruen-R.Silverman, *Arcs and blocking sets II*, Europ. J. Combinatorics 8 (1987), 351 - 356.
8. P.J.Cameron, *Four lectures on finite geometry*, Finite Geometries (ed. C.A.Baker, L.M.Batten) Lecture Notes in Pure and Appl. Math. 103, M.Dekker, New York, 1985, 27-63.
9. J.W.P.Hirschfeld, *Projective geometries over finite fields*, Clarendon Press, Oxford, (1979).
10. T.Szonyi, *Combinatorial problems for abelian groups arising from geometry*, to appear in Periodica Polytechnica.
11. G.Tallini, *On blocking sets in finite projective and affine spaces*, Annals of Discrete Math. 37 (1988), 433-450.

SISTEMI DINAMICI DIGITALI E LORO POSSIBILI APPLICAZIONI

Giovanni Melzi

SUNTO - Alla definizione assiomatica di Sistema Dinamico Digitale (SDD) fornita attraverso opportuni legami funzionali tra insiemi di parole su un alfabeto dato, seguono alcune suggestive interpretazioni di tipo algebrico-combinatorio, analitico e topologico dell'idea di SDD.

ABSTRACT - In this paper some developments of the *Dynamic Digital System* (SDD) have been exposed. A SDD is defined as a dynamic system whose instantaneous state and input-output are made by codified information, particularly in words on a given alphabet. Some algebraic-combinatory, analytical and topological suggestions have been exposed afterwards.

0. In questa nota si espongono, e in qualche caso si anticipano, alcuni sviluppi relativi alla definizione di *Sistema Dinamico Digitale* (SDD). Questi sviluppi si riferiscono a lavori, in corso di pubblicazione di E. Allevi, G. Melzi, F. Mercanti ed altri.

Un sistema dinamico, nel senso usuale del termine, è caratterizzato dal fatto di scambiare con l'esterno *grandezze* (per lo più fisiche) e di possedere uno *stato istantaneo* descrivibile con uno o più *attributi misurabili*. Non ci si scosta molto dalla verosimiglianza pensando e dicendo che un sistema dinamico è interamente descritto da certe *variabili numeriche* funzioni le une delle altre, e tutte funzioni del tempo.

Un esempio tipico di sistema dinamico è quello di un circuito complesso caratteristico dell'attuale ingegneria elettronica. Altri esempi molto significativi sono offerti dalla fisica dei sistemi con un elevato numero di gradi di libertà. Nel caso di tali sistemi dinamici risulta naturale studiarne il movimento con gli usuali strumenti analitici: funzioni, equazioni differenziali, algoritmi (numerici) e così via.

Al contrario un SDD è definito come un sistema dinamico il cui stato istantaneo ed il cui input-output siano costituiti da *informazione codificata*, tipicamente informazione codificata in *parole su un dato alfabeto*. La motivazione della ricerca su oggetti astratti di questo tipo è molto agevole e suggestiva: un sistema percettivo, un ecosistema, un distretto fisiologico, un reattore chimico, una rete nervosa, un sistema economico e perfino un gruppo animale o umano si lasciano spontaneamente immaginare, salvo inevitabili specificazioni, come SDD.

1. Dal punto di vista della modellazione matematica di un SDD sembra corretto procedere ad una definizione assiomatica nel modo seguente, utilizzando linguaggio e simbologia introdotti in MELZI-MERCANTI [1] e MELZI [2].

Siano T un *alfabeto* costituito da un numero finito di caratteri e $\mathcal{G}$ una *grammatica* costituita dall'insieme (finito o infinito numerabile) delle *parole* su T . Sia $P(\mathcal{G})$ l'insieme delle parti di $\mathcal{G}$. Sia $B(\mathcal{G})$ l'algebra di Boole i cui elementi siano elementi di $P(\mathcal{G})$, ossia *insiemi di parole*. Tipicamente, per gli scopi attuali, si opera su insiemi finiti, ma ciò non interdice l'ovvia generalizzazione ad insiemi anche infiniti. Si definiscono inoltre alcune *operazioni fra parole* indicate con i simboli $\neg, \vdash, \perp, \mid, O$ ([1], §2), osservando solamente che si possono definire assiomaticamente molte altre operazioni tra parole e anche tra *insiemi* (finiti o infiniti) di *parole* ossia tra gli elementi di $P(\mathcal{G})$. L'estensione delle *operazioni tra parole* alle *operazioni tra insiemi* (finiti o infiniti) di *parole* si esegue mediante formule del tipo

$$X \neg Y = \{ Z \mid \exists x \exists y ((x \in X) \wedge (y \in Y) \wedge (Z = x+y)) \}$$

e analoghe.

Si può quindi definire un' *algebra* A , generalizzazione dell'algebra $B(\mathcal{G})$, ottenuta con l'aggiunzione delle operazioni $\neg, \vdash, \perp, \mid, O$ (o altre ancora, all'occorrenza). In modo del tutto analogo si introducono i concetti di *A-espressione*, ossia espressione in A , di *A-equazione*, di *A-funzione* e così via ([1], §3).

2. Sia ora

$$\underline{X} = [C_0 \mid C_1 \dots C_p \mid C_{p+1} \dots C_{p+q}] \quad (1)$$

un vettore ad $n=p+q+1$ componenti su $P(\mathcal{G})$. Si supponga che il tempo sia suddiviso in intervalli $t_0, t_1, \dots, t_h, \dots$ e che ciascuno degli insiemi $C_0, \dots, C_n$ ($n=p+q+1$) sia variabile in funzione di t_i ($i=0, 1, \dots, h, \dots$). Più precisamente si supponga che siano C_j^i ($j=0, 1, \dots, n$) i valori assunti dalle variabili C_j ($j=0, 1,$

..., n) al tempo t_i ($i=0, 1, \dots, h, \dots$) e che sussistano i seguenti legami funzionali

$$C_j^{i+1} = \emptyset_j (C_0^i \mid C_1^i \dots C_p^i \mid C_{p+1}^i \dots C_q^i), \quad (2)$$

essendo $\emptyset_j$ ($j=0, 1, \dots, n$) certe A-funzioni prefissate, per le quali il tempo t_i non è una variabile indipendente. In questo senso si può dire che le $\emptyset_j$ non dipendono dal tempo.

In tal modo si cerca di esprimere la circostanza che le componenti del vettore (1) siano le une funzioni delle altre, ossia che gli *insiemi di parole siano funzioni variabili con il tempo*, ma che le dipendenze funzionali siano costanti nel tempo, ossia siano delle caratteristiche fisiche del sistema.

3. Lo studio di un SDD definito da un sistema di A-funzioni del tipo (2) dà subito luogo a consistenti problemi algebrico-combinatori ed anche analitici. Per esempio si può pensare al movimento di un SDD come all'evoluzione temporale di "folle di parole" dotate di proprietà di eccitazione e di attenuazione reciproche. Tutto ciò è analogo a quanto avviene in un sistema dinamico in senso tradizionale, salvo ovviamente il fatto che attualmente esiste una notevole ricchezza di proprietà di tipo *digitale* (in contrapposizione con il carattere prettamente analogico dei sistemi dinamici).

4. I legami concettuali tra la teoria ancora embrionale dei SDD e l'analisi (reale o complessa) dei sistemi dinamici tradizionali sono molto ricchi ed impegnativi. Ecco alcune suggestioni di tipo algebrico-combinatorio, analitico e topologico.

Anzitutto si può rappresentare un insieme finito o infinito di parole su un alfabeto con coordinate reali, e in tal modo un *insieme* finito o infinito di parole appare come un *punto* in uno spazio dotato in generale di un numero finito o infinito numerabile di coordinate. Allora il *movimento* di un SDD può essere visto come una *traiettoria* in uno spazio in generale pre-hilbertiano (o, in casi assai significativi, hilbertiano). E' evidente che con tale immagine geometrica le proprietà di continuità o di discontinuità acquistano un notevole significato intuitivo.

Un'altra suggestiva interpretazione è fornita dalla possibilità di descrivere la *transizione* da uno stato all'altro di un SDD in *termini* algebrico-combinatori, vertenti su insiemi (finiti o infiniti) di parole (finite). Per tale via le proprietà di eccitazione e di attenuazione reciproche tra folle di parole possono essere descritte e misurate in termini topologici e metrici in spazi funzionali opportuni.

Significative applicazioni di tipo naturalistico ed economico sono in corso di studio, come appare dalla pur sommaria nota bibliografia (MELZI [3]).

MERCANTI [4], [5]).

BIBLIOGRAFIA

1. G. MELZI e F. MERCANTI, *An algebraic-combinatory Theory of real nervous System*, Rend. Sem. Matem. di Brescia, Italy, 9 (1988), 107-121.
2. G. MELZI, *Sulla definizione di semiautoma*, Rend. Sem. Matem. di Brescia, Italy, 10 (1988), 23-50.
3. G. MELZI, *Optical Illusions as an Example of Fuzzy - perception*, TÜV, Verlag, Rheinland (1986), 231-248.
4. F. MERCANTI, *Una suggestiva analogia tra fenomeni biologici e fenomeni economici a soglia*, Atti Conv. Matem. Applic. Econom. e Ingegn., Università Studi Chieti, (1990), in questo volume.
5. F. MERCANTI, *Interpretazione economica dei predicati di canale in un Sistema Digitale Multicanale*, Atti Conv. Matem. Applic. Econom. e Ingegn., Università Studi Chieti, (1990), in questo volume.
6. E. ALLEVI, *Una speciale classe di sistemi dinamici*, In corso di pubblicazione.

UNA SUGGESTIVA ANALOGIA FRA FENOMENI BIOLOGICI E FENOMENI ECONOMICI A SOGLIA

Fabio Mercanti

SUNTO - Viene fornita una descrizione in termini algebrico-combinatori che sembra particolarmente adatta per una modellazione matematica dei *fenomeni a soglia* tipici della biologia e della economia.

ABSTRACT - An algebraic-combinatory description which appears particularly suitable to mathematical models of *threshold phenomena* of the biological and economical processes has been presented. An analogy, probably not yet known, between energetical investment phenomena in biological and economical system has been described.

O. In questa nota si descrive un'analogia probabilmente non ancora nota fra fenomeni di investimento energetico in sistemi biologici ed in sistemi economici. Il lavoro si inquadra nel filone delle ricerche attualmente dirette da G. Melzi.

L'analogia di cui si parla consiste in quanto segue. Si ammette anzitutto che il fenomeno biologico della percezione consista nella *saturazione di un segnale* per mezzo di risorse energetiche limitate, nel senso descritto in MELZI [2], [3], MELZI-MERCANTI [1]. Se si osserva che ogni processo economico si lascia descrivere, a prescindere da tutti i possibili e anzi inevitabili disturbi osservazionali, per mezzo di strategie ottimali di investimento di risorse limitate, l'analogia annunciata diventa ovvia.

Sono in corso ricerche tendenti alla modellazione matematica in questo senso di alcuni suggestivi meccanismi economici. Ad esempio, il modello classico di Straffa di produzione di merci per mezzo di merci, il modello della gestione di un portafoglio di n titoli, e altre.

Questa nota tende a chiarire solo il primo termine dell'analogia, e cioè certi meccanismi di controllo nel sistema percettivo ottico descritti in MELZI [2] e qui sviluppati e predisposti per il successivo esame analogico.

Il presente lavoro riprende e sviluppa [2] quantificando con una certa precisione la *topologia metrica* degli stati di un tale sistema percettivo.

1. La simbologia e la terminologia occorrenti per la trattazione che segue si trovano in [1]. In [1] è definito un *semiautoma*

$$\Sigma = \langle H(F, X_0), \Omega \rangle$$

costituito da un'automa $H(F, X_0)$ e da un osservatore Ω (detto in [3] *osservatore interno*) in grado di interagire con H stesso. Gli stati di H (tra i quali lo stato iniziale X_0) sono insiemi di *parole finite* costruite su un dato *alfabeto finito* T . Tali insiemi di parole sono gli elementi di un'algebra $\mathcal{A}$, estensione dell'algebra di Boole $\mathcal{B}$, le cui operazioni sono, oltre a quelle dell'algebra $\mathcal{B}$, quelle di *selezione*, *selezione destra*, *selezione sinistra*, *taglio* e *concatenazione*, dette brevemente *operazioni*. Le *operazioni* sono verosimili modelli delle operazioni di *sintesi polipeptidica*.

Se ad un dato tempo (discreto) l'automa H è nello stato X e riceve lo stimolo Y dall'osservatore Ω , allora lo stato che H assume al tempo successivo in virtù dello stimolo Y è

$$F(X, Y),$$

dove F è una *funzione omogenea* ([1], § 5) definita su $\mathcal{A}$.

2. In [2] si considerano un semiautoma $\Sigma = \langle H(F, X_0), \Omega \rangle$ e un altro osservatore ω . Nel seguito Ω e ω saranno anche chiamati, come in [3], rispettivamente, *osservatore interno* ed *osservatore esterno* (ad H). Quando l'automa $H(F, X_0)$ assume lo stato X , l'osservatore esterno "propone" all'osservatore interno Ω di raggiungere un altro stato a partire da X , e così via. Quando ω propone lo stato X a Ω si dice che H è *non-equilibrato*. Quando lo stato X è assunto *dopo* la scelta operata da Ω si dice che H è *riequilibrato* ([2], § 8).

Al processo P in cui H assume gli stati X_i sotto l'azione degli stimoli Y_i ($i = 1, 2, \dots, k$) si associa in modo del tutto naturale il numero

$$\sum_i F(x_i, y_i), \quad (1)$$

con $x_i = |X_i|$ e $y_i = |Y_i|$, detto *traccia* di P .

Il segno Σ di sommatoria si riferisce alla ordinaria somma di anello booleano.

Il funzionamento complessivo del semiautoma Σ è il seguente: l'osservatore esterno ω squilibra casualmente in ogni tempo l'automa H e l'osservatore interno Ω riequilibra H in ogni tempo, per modo che, con l'investimento energetico (1), il processo P è complessivamente mantenuto equilibrato.

Tuttavia la strategia di riequilibrio in ogni tempo a disposizione di Ω non è univocamente determinata. Si possono prevedere le seguenti ipotesi fra loro alternative di comportamento di Ω .

2.1. Per ogni stimolo da riequilibrare Ω porta H allo stato X_0 e da questo allo stato che riequilibra il nuovo stimolo.

2.2. L'osservatore Ω sceglie lo stimolo riequilibrante in modo che la relativa traccia sia interpretabile come la *distanza* fra gli stati X_{i-1} , X_i in una topologia degli stati, opportunamente definita.

La proprietà triangolare della distanza $d(X, Y)$ fra stati di H , come è detto al § 5, assicura la vantaggiosità del comportamento (2.1) rispetto al comportamento (2.2), per ogni singolo passo X_{i-1} , X_i .

3. Il processo tra ω e Ω descritto nel § 2 può essere visto come uno schema matematico dell'equilibrio fra gli stimoli esterni e la risposta proteica in un'area percettiva del sistema nervoso centrale. In particolare la traccia (1) può essere interpretata come *energia dispersa* nel processo P . Il meccanismo di nonequilibrio-riequilibrio descritto nel § 2 può allora essere visto come un *modello concettuale della percezione sensoriale*. In particolare in [2] si tratta della percezione visiva. Ivi si considerano ω come un modello della indeterminatezza degli stimoli provenienti dalla realtà fisica e Ω come un modello della corteccia visiva che controlla l'ingresso percettivo.

I meccanismi di equilibrio fra ω e Ω possono essere descritti con l'algebra $\mathcal{A}$ di [1] (cfr. § 1), come è sommariamente detto in [2].

4. L'illusione ottica presa in considerazione in [2] è la seguente. Due sorgenti luminose di uguali caratteristiche ottiche emettono luce alternativamente per opportuni periodi di tempo nell'ordine del secondo. La percezione che se ne ha ad opportuna distanza dalle sorgenti è quella di un'unica immagine che oscilla con continuità dalla posizione di una sorgente alla posizione dell'altra, passando attraverso le posizioni intermedie $P_1, P_2, \dots, P_n$, dalle quali sembra provenire. Man mano che l'osservatore si avvicina alle due sorgenti l'impressione di continuità aumenta, ma scompare quasi di colpo quando la posizione O dell'osservatore, ossia il corrispondente angolo $\alpha = P_1 O P_n$, raggiunge un punto

critico. Vi è dunque una sorta di *soglia* che varia, di poco da individuo a individuo, o in particolari condizioni di osservazione.

5. L'interpretazione del fenomeno in termini combinatori degli stati dell'area visiva della corteccia come un semiautoma è la seguente. Siano $S_1, S_2, \dots, S_n$ gli stati del semiautoma Σ corrispondenti alla saturazione degli stimoli $S'_1, S'_2, \dots, S'_n$ che al semiautoma verrebbero forniti dalle immagini provenienti dai punti $P_1, P_2, \dots, P_n$ (§4). In base a quanto detto nei precedenti paragrafi ogni stato S_i ($i = 1, 2, \dots, n$) è un sottoinsieme dell'insieme $\mathcal{G}$ di tutte le parole sull'alfabeto T (§1). Si può attribuire a ciascun insieme S_i ($i = 1, 2, \dots, n$) un *peso* p ossia si può definire un'applicazione

$$\pi : P(\mathcal{G}) \rightarrow \mathcal{N}$$

dell'insieme $P(\mathcal{G})$ delle parti dell'insieme $\mathcal{G}$ di tutte le parole sull'alfabeto T (§1), definendo $\pi(p)$ come la cardinalità dell'insieme (discreto) $p \in P(\mathcal{G})$.

Nell'insieme $P(\mathcal{G})$ delle parti di $\mathcal{G}$ è inoltre facile definire una *distanza* $d(p, q)$ tra due parti p, q , dotata delle usuali proprietà $d(p, p) = 0$, $d(p, q) \geq 0$, $d(p, q) = d(q, p)$ e della proprietà triangolare. Si dimostra facilmente che una tale funzione può essere il peso della somma $p \oplus q$ di anello booleano di p e q , ossia

$$d(p, q) = \pi(p \oplus q).$$

Allora per la proprietà triangolare, detto S_0 lo stato iniziale del semiautoma Σ , per ogni $i = 1, 2, \dots, n-1$ si ha

$$d(P_i, P_{i+1}) < d(P_i, S_0) + d(S_0, P_{i+1}).$$

Invece l'essere

$$\sum_{i=1}^n d(P_i, P_{i+1}) \geq d(P_1, S_0) + d(S_0, P_n) \quad (2)$$

dipende ovviamente dal numero dei lati della poligonale (astratta) $Q_1, Q_2, \dots, Q_n$, corrispondente alla successione $P_1, P_2, \dots, P_n$ delle immagini virtuali (§4), valendo certamente nella (2) il segno inferiore per $n = 2$.

6. Nella (2) del §5 si scorge una suggestiva interpretazione dell'illusione ottica da cui si è partiti (§4). Assimilando il peso p di un insieme P di parole, ossia la cardinalità di tale insieme, all'energia occorrente per sintetizzare tali parole, diventa evidente che la lunghezza di ogni singolo lato della poligonale

$P_1, P_2, \dots, P_n$ (§ 5) è tanto più piccola quanto più due insiemi consecutivi, corrispondenti ad un lato di tale poligonale, hanno numerose parole in comune, a causa della ben nota definizione di somma di anello booleano. Quanto più piccolo è n (nella (2)) ossia quanto minore è la distanza angolare delle due sorgenti luminose, tanto più sarà soddisfatta la (2) con il segno inferiore. Più precisamente esiste un intero n_0 , da dirsi *soglia* per la (2), tale che per $n < n_0$ valga la (2) con il segno inferiore e per $n \geq n_0$ valga la (2) con il segno superiore.

7. In termini di percezione visiva si può ora dire che quanto più sono vicine le sorgenti luminose alternative tanto più piccolo è il dispendio energetico della percezione a stadi intermedi. Il semiautoma Σ risparmia energia di saturazione investendo piccoli incrementi di energia *per percepire le sorgenti luminose virtuali (inesistenti)* (§ 6) $P_2, P_3, \dots, P_{n-1}$ piuttosto che cancellando le parole corrispondenti alla percezione di P_1 e inversamente. La soglia n_0 corrisponde, da questo punto di vista, alla *soglia di convenienza energetica nel processo di saturazione percettiva degli stimoli*.

Ciò spiega in maniera assai suggestiva la scomparsa improvvisa dell'illusione ottica al crescere dell'angolo apparente α fra le due sorgenti alternative (§ 4):

la scomparsa avviene in corrispondenza al superamento della soglia n_0 , ossia dell'angolo α corrispondente.

Invece la casuale scomparsa del fenomeno per qualche breve intervallo di tempo ad angolo α costante si spiega assai bene con l'esistenza di soglie energetiche nel rifornimento energetico continuo di risorse metaboliche all'apparato visivo.

Si può difendere l'opinione che la descrizione in termini algebrico-combinatori che qui è stata esposta è assai più adatta alla modellazione matematica dei fenomeni a soglia tipici della biologia e della economia (macro e micro) del tradizionale apparato analitico fondato essenzialmente sui concetti di proporzionalità e continuità.

BIBLIOGRAFIA

1. G. MELZI e F. MERCANTI, *An algebraic-combinatory Theory of real nervous System*, Rend. Sem. Matem. di Brescia, Italy, 9 (1988), 107-121.
2. G. MELZI, *Optical illusion as an Example of Fuzzy-perception*, TÜV, Verlag, Rheinland (1986), 231-248.
3. G. MELZI, *Sulla definizione di semiautoma*, Rend. Sem. Matem. di Brescia, Italy, 10 (1988), 23-50.

UN PROBLEMA DI OTTIMIZZAZIONE NON-SMOOTH IN STATISTICA ROBUSTA

Luciano Stefanini (*)

Sommario - In questa nota viene preso in esame un problema di stima parametrica robusta noto come *Least Median of Squares Regression (LMS Regression)*. Esso consiste nel determinare i parametri da stimare mediante la minimizzazione della mediana dei quadrati dei residui, in analogia con il metodo classico dei minimi quadrati in cui se ne minimizza la media.

In due recenti lavori di D.L. Souvaine - J.M. Steele [11] e di J.M. Steele - W.L. Steiger [12], sono stati proposti due algoritmi per il caso lineare, basati sulle proprietà degli iperpiani di regressione. Tali algoritmi non sono però applicabili al caso non-lineare e perdono di efficienza, nel caso lineare, col crescere del numero dei parametri del modello.

L'approccio qui adottato consente di studiare alcune caratteristiche della regressione LMS come problema di ottimizzazione vincolata a due livelli, in cui si deve minimizzare una funzione obiettivo soggetta a vincoli che contengono, a loro volta, un (sotto)problema di ottimizzazione. Ciò fa intervenire funzioni non-differenziali (*nonsmooth*) della classe poliedrico-convessa, le quali rendono conto della natura intrinsecamente combinatoria della stima LMS. La struttura del problema consente di ricavare un algoritmo di tipo Branch & Bound, applicabile indifferentemente sia al caso lineare che al caso non-lineare.

INTRODUZIONE

Nella letteratura recente sulla stima parametrica, a partire dai primi lavori di P. Huber (1964) e di F.R. Hampel (1971), è stato introdotto in modo sistematico il concetto di robustezza delle stime, precisato tramite diverse classi di stimatori robusti. Recentemente, P.J. Rousseeuw (in [9] e poi in [8]), in analogia con il

(*) Facoltà di Economia e Commercio - Urbino
Lavoro svolto con il contributo del Ministero della Pubblica Istruzione

modo di procedere nel caso dei minimi quadrati (in cui si minimizza la Media dei quadrati dei residui), ha proposto un criterio di stima che consiste nel minimizzare la Mediana dei quadrati dei residui (Least Median of Squares Regression), il cui *breakdown point* è $1/2$.

Se con $r_1, r_2, \dots, r_m$ si indicano i quadrati dei residui, la loro mediana è definita come un valore M che minimizza la somma degli scarti assoluti:

$$M \in \arg \min_{\mu} \sum_{j=1}^m |r_j - \mu|$$

ovvero

$$\sum_{j=1}^m |r_j - M| = \min_{\mu} \sum_{j=1}^m |r_j - \mu|$$

In due recenti lavori, D.L. Souvaine - J.M. Steele [11] e J.M. Steele - W.L. Steiger [12] hanno proposto due algoritmi efficienti per il caso lineare, basati sulle proprietà degli interpiani di regressione. Tali algoritmi non sono però applicabili nel caso di modelli di regressione non-lineare. Inoltre, essi perdono rapidamente di efficienza col crescere del numero p dei parametri da stimare (la loro complessità computazionale è dell'ordine $O(m^p)$, ciò che li rende vantaggiosamente applicabili per bassi valori di p).

In questo lavoro si prende in esame la stima LMS nel caso di un modello di regressione generale, in cui i residui possono essere definiti a partire da una qualsiasi funzione (di classe $C^{(1)}$) dei parametri e dei dati.

Il problema viene dapprima formulato in termini di ottimizzazione a due livelli (*bilevel programming*) e, utilizzando la teoria classica della dualità, viene impostato come problema non-lineare che, descritto in forma compatta, contiene una funzione poliedrica convessa che ne caratterizza la natura *nonsmooth* e combinatoria.

Infine, utilizzando la particolare struttura del problema di ottimizzazione, viene proposto un metodo di tipo branch & bound per ottenere una procedura di minimizzazione globale applicabile sia al caso di un modello lineare che non-lineare.

FORMULAZIONE DEL PROBLEMA DI STIMA:

Si consideri il problema di determinare p parametri $\theta = (\theta_1, \theta_2, \dots, \theta_p)^T$ che definiscono un modello del tipo

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

dove:

$$x \in \mathbf{R}^q, \theta \in \mathbf{R}^p, f: \mathbf{R}^q \times \mathbf{R}^p \longrightarrow \mathbf{R},$$

sulla base di un numero m di osservazioni date dai vettori:

$$y = (y_1, y_2, \dots, y_m)^T,$$

$$x_i = (x_{i1}, x_{i2}, \dots, x_{iq})^T, \quad i = 1, 2, \dots, m.$$

Indichiamo i quadrati dei residui con

$$r_i(\theta) = [y_i - f(x_i, \theta)]^2, \quad i = 1, 2, \dots, m.$$

Fissato $\theta \in \mathbf{R}^p$, la mediana $M(\theta)$ dei quadrati dei residui, funzione di θ , è una soluzione del problema di minimo seguente:

$$M(\theta) \in \arg \min_{\mu} \sum_{j=1}^m |r_j(\theta) - \mu| \quad (1)$$

ed essendo $r_j(\theta) \geq 0 \quad \forall j$, si ha $M(\theta) \geq 0$. Se il minimo di tale problema, per θ fissato, non è unico, si può scegliere il più piccolo degli argomenti μ .

La stima LMS consiste nel determinare il vettore che minimizza $M(\theta)$:

$$\text{Min}_{\theta} M(\theta). \quad (2)$$

Definendo le variabili

$$d_j^+ = \max \{0, r_j(\theta) - \mu\} \quad \text{e} \quad d_j^- = \max \{0, \mu - r_j(\theta)\} \quad (3)$$

$$\text{ed avendosi } r_j(\theta) - \mu = d_j^+ - d_j^- \quad \text{e} \quad |r_j(\theta) - \mu| = d_j^+ + d_j^-,$$

il problema (2) può essere posto nella forma seguente:

$$\begin{aligned} &\text{Min}_{\theta, \mu} \quad \mu \\ &\text{s.t.} \quad \mu \geq 0 \end{aligned} \quad (4)$$

$$\begin{aligned} &\sum_{j=1}^m (d_j^+ + d_j^-) = \text{Minimo} \\ &\text{s.t.} \end{aligned} \quad (5)$$

$$\mu = r_j(\theta) - d_j^+ + d_j^- \quad \text{per} \quad j = 1, 2, \dots, m \quad (6)$$

$$d_j^+ \geq 0, \quad d_j^- \geq 0. \quad (7)$$

Il problema (4)-(7) contiene due livelli di ottimizzazione, il secondo comparando fra i vincoli del primo (*Bilevel Programming Problem*, cfr. J. Bard, [1]).

Consideriamo il sottoproblema (5)-(7) per θ fissato. Se $\lambda_1, \lambda_2, \dots, \lambda_m$ sono le variabili duali associate ai vincoli (6), il problema duale di (5)-(7), corrispondente al calcolo della mediana, è il seguente (θ fissato):

$$\text{Max } \sum_{j=1}^m \lambda_j r_j(\theta) \quad (8)$$

s.t.

$$\sum_{j=1}^m \lambda_j = 0 \quad (9)$$

$$-1 \leq \lambda_i \leq 1, \quad i = 1, 2, \dots, m. \quad (10)$$

Poiché, per θ fissato, le funzioni che intervengono sono lineari, le seguenti condizioni di Kuhn-Tucker sono necessarie e sufficienti affinché il vettore $(\mu, d_1^+, \dots, d_m^+, d_1^-, \dots, d_m^-, \lambda_1, \dots, \lambda_m)^T$ sia soluzione della coppia di problemi (5)-(7) e (8)-(10).

$$\mu + d_i^+ - d_i^- - r_i(\theta) = 0 \quad \text{per } i = 1, 2, \dots, m \quad (11)$$

$$\sum_{j=1}^m (d_j^+ + d_j^-) - \sum_{j=1}^m \lambda_j r_j(\theta) = 0 \quad (12)$$

$$d_j^+ \geq 0, d_j^- \geq 0, \mu \geq 0 \quad (13)$$

$$\sum_{j=1}^m \lambda_j = 0 \quad (14)$$

$$-1 \leq \lambda_j \leq 1 \quad (15)$$

Il problema di stima LMS può allora essere formulato nel modo seguente, nelle $3m+p+1$ variabili $\mu \in \mathbf{R}$, $\theta \in \mathbf{R}^p$, $d^+ \in \mathbf{R}$, $d^- \in \mathbf{R}^m$, $\lambda \in \mathbf{R}^m$:

$$\begin{aligned} & \text{(P) Min } \mu \\ & \text{s.t. (11) - (15).} \end{aligned}$$

OSSERVAZIONI SUL PROBLEMA DI OTTIMIZZAZIONE

Esaminiamo ora il problema (8)-(10), che può essere riscritto, per valori fissati di θ , nella forma:

$$\begin{aligned} \text{Max } & 2 \sum_{j=1}^m |v_j r_j(\theta) - r_j(\theta)| \\ \text{s.t. } & \\ & 2 \sum_{j=1}^m v_j = m \\ & 0 \leq v_j \leq 1 \end{aligned}$$

dove si è posto $v_j = (\lambda_j + 1) / 2$ per ogni $j = 1, 2, \dots, m$

Come noto, il problema (5) è un caso particolare di *knapsack problem* con vincoli sulle variabili, la cui soluzione (B.G. Dantzig [3]) richiede semplicemente di ordinare i termini $r_j(\theta)$ in senso non-decrescente:

$$r_1(\theta) \geq r_2(\theta) \geq \dots \geq r_m(\theta), \quad (16)$$

di scegliere $t = m/2$ se m è pari, ovvero $t = (m+1)/2$ se m è dispari e di porre

$$v_j = \begin{cases} 1 & \text{se } j < t \\ m/2 - t + 1 & \text{se } j = t \\ 0 & \text{se } j > t \end{cases}$$

e quindi

$$\lambda_j = \begin{cases} +1 & \text{se } j < t \\ m - 2t + 1 & \text{se } j = t \\ -1 & \text{se } j > t \end{cases}$$

Ne segue che i valori di λ_j sono tutti interi dell'insieme $\{-1, 0, 1\}$ e, indicato con $\mathbf{L}$ l'insieme dei vettori λ che, al variare di θ , sono candidati a risolvere il problema (8)-(10), si ha che

$$\mathbf{L} \subset \{-1, 0, 1\}^m \text{ è un insieme finito.}$$

Il problema (8)-(10) può allora essere scritto nella forma seguente:

$$\Phi(\theta) = \text{Max} \left\{ \sum_{j=1}^m r_j(\theta) \lambda_j; \lambda \in \mathbf{L} \right\} \quad (17)$$

Poiché una funzione del tipo

$$h(x) = \text{Max} \{ x^T \lambda; \lambda \in L \}, \text{ al variare di } x,$$

è poliedrica convessa (*polyhedral convex*), caratterizzata da un numero finito di iperpiani di supporto, e $\Phi(\theta) = h(r(\theta))$, la funzione (17) è composta da $r(\theta)$ che, nelle nostre ipotesi, è differenziabile e da $h(r)$ che è invece convessa, continua ma non ovunque differenziabile.

Il problema

$$\text{Min}_{\theta} \quad \Phi(\theta)$$

è stato studiato da R. Fletcher ([5] e [6]), il quale ha proposto un algoritmo globalmente convergente.

Si osservi che, essendo

$$\Phi(\theta) = \sum_{j=1}^m |r_j(\theta) - M(\theta)|, \quad (18)$$

$\Phi(\theta)$ rappresenta una misura di dispersione dei quadrati dei residui attorno alla loro mediana e la minimizzazione della (18) ha di per sé interesse. La specializzazione dell'algoritmo di Fletcher a (18) è descritta in un lavoro precedente [13].

Inoltre, la funzione $\Phi(\theta)$ esprime la natura non differenziabile del problema (P). Infatti, tramite la definizione (17), possiamo esprimere (P) nella forma seguente:

$$\begin{array}{ll} \text{Min} & \mu \\ \text{s.t.} & \\ \mu = r_i(\theta) - d_i^+ + d_i^- & \text{per } i = 1, 2, \dots, m \end{array}$$

$$\sum_{j=1}^m [d_j^+ + d_j^-] = \Phi(\theta)$$

$$d_j^+ \geq 0, d_j^- \geq 0, \mu \geq 0$$

In questa ultima formulazione, in cui la funzione obiettivo è lineare e le funzioni dei vincoli sono convesse, il vettore λ non compare esplicitamente bensì è contenuto nella funzione $\Phi(\theta)$ che compare in un vincolo.

UN ALGORITMO DI TIPO BRANCH AND BOUND

Come si è visto nella sezione precedente, i vettori λ che entrano come variabili nel problema (P) appartengono ad un insieme finito

$$L \subseteq \{\lambda \mid \lambda_i \in \{-1, 0, 1\}, \sum_{j=1}^m \lambda_j = 0\}$$

Supponiamo di conoscere il vettore λ soluzione di (P) e definiamo i tre insiemi disgiunti di indici:

$$\Lambda^+ = \{j \mid \lambda_j = +1\}, \quad \Lambda^- = \{j \mid \lambda_j = -1\} \text{ e} \\ \Lambda^0 = \{j \mid \lambda_j = 0\} \text{ (eventualmente } \Lambda^0 \text{ è vuoto).}$$

Dal problema (8)-(10) si ha che i primi due insiemi hanno cardinalità uguale a $t = \lfloor m/2 \rfloor$, il troncamento intero di $m/2$, e che l'unione dei tre insiemi è tutto $\{1, 2, \dots, m\}$.

Si osservi, per inciso, che, se $\lambda_i = 0$ allora $d_i^- + d_i^+ = 0$, da cui anche $d_i^- = d_i^+ = 0$ e quindi $\mu = r_i(\theta)$. In ogni caso, inoltre,

$$\begin{aligned} \lambda_j = -1 &\Rightarrow d_j^- = 0, \quad d_j^+ = r_j(\theta) - \mu \geq 0 \\ \lambda_j = +1 &\Rightarrow d_j^+ = 0, \quad d_j^- = \mu - r_j(\theta) \geq 0 \end{aligned}$$

Allora, fissato il vettore λ che soddisfi (9) e (10), il calcolo di θ è riconducibile al problema seguente:

$$(Q) \quad \begin{array}{ll} \text{Min}_{\theta, \mu} & \mu \\ \text{s.t.} & \end{array}$$

$$\begin{aligned} r_j(\theta) - \mu &\geq 0 && \text{per } j \in \Lambda^- \\ \mu - r_j(\theta) &\geq 0 && \text{per } j \in \Lambda^+ \\ \mu &= r_j(\theta) && \text{per } j \in \Lambda^0 \end{aligned}$$

Il problema precedente, contenente le variabili μ e θ , equivale a (P) nell'ipotesi che siano noti a priori gli insiemi Λ corrispondenti al vettore ottimo λ . Se invece tali insiemi di indici non si riferiscono all'ottimo λ , è immediato verificare che la soluzione (μ_Q, θ_Q) di (Q) è legata alla soluzione (μ_P, θ_P) di (P) della disuguaglianza

$$\mu_P \leq \mu_Q,$$

ossia (Q) fornisce un *upper bound* per (P).

Inoltre, se gli insiemi Λ in (Q) sono sostituiti da tre loro sottoinsiemi

$$V^+ \subseteq \Lambda^+, V^- \subseteq \Lambda^- \text{ e } V^0 \subseteq \Lambda^0$$

si ottiene un problema

$$\begin{aligned}
 & \text{(BQ)} \quad \text{Min}_{\theta, \mu} \mu \\
 & \quad \text{s.t.} \\
 & \quad r_j(\theta) - \mu \geq 0 \quad \text{per } j \in V^- \\
 & \quad \mu - r_j(\theta) \geq 0 \quad \text{per } j \in V^+ \\
 & \quad \mu = r_j(\theta) \quad \text{per } j \in V^0 \\
 & \quad \mu \geq 0
 \end{aligned}$$

la cui soluzione $(\mu_{\text{BQ}}, \theta_{\text{BQ}})$ è tale che $\mu_{\text{BQ}} \leq \mu_Q$, per cui (BQ) fornisce un *lower bound* a (Q).

Queste proprietà suggeriscono una procedura di ricerca di tipo *branch and bound* in cui ogni percorso (nodo finale) dell'albero di ricerca corrisponde ad una scelta degli insiemi V . Una strategia di branching è la seguente: si considerano m possibili livelli, corrispondenti alle m componenti del vettore λ ; ad ogni livello sono creati tre *branch*, corrispondenti alle tre possibili scelte seguenti:

$$\begin{array}{ccc}
 0 & & \text{livello } j-1 \\
 0 & 0 & 0 \\
 j \in V^+ & j \in V^0 & j \in V^- \quad \text{livello } j
 \end{array}$$

Ad ogni nodo del livello m si ha una terna di insiemi Λ per il problema (Q) mentre ad ogni livello intermedio si ha un problema (BQ).

L'algoritmo può allora essere schematizzato nel modo seguente:

STEP 0 (INIZIALIZZAZIONE)

Dato $\theta = \theta_0$ iniziale (qualsiasi), calcola $r_i(\theta)$ per $i = 1, 2, \dots, m$ e risolvi il problema (8)-(10) (richiede solo un ordinamento).

Definisci i tre insiemi Λ e risolvi il corrispondente problema (Q) ottenendo (μ_Q, θ_Q) . Definisci $t^+ = |\Lambda^+|$, $t^0 = |\Lambda^0|$ e $t^- = |\Lambda^-|$ uguali alle cardinalità dei tre insiemi Λ .

Per quanto visto, μ_Q è una soluzione iniziale di (P).

Poni $j = 0$ (indice iniziale di livello), $V^+ = V^- = V^0 =$ Insieme vuoto.

STEP 1. (BRANCHING AL LIVELLO SUCCESSIVO)

Incrementa l'indice di livello: $j \leftarrow j + 1$.

Se $|V^+| < t^+$ allora: $s_j \leftarrow +1$

$V^+ \leftarrow V^+ \cup \{j\}.$
 altrimenti, se $|V^0| = 0$ allora: $s_j \leftarrow 0$
 $V^0 \leftarrow \{j\}$
 altrimenti: $s_j \leftarrow -1$
 $V^- \leftarrow V^- \cup \{j\}.$

(Il vettore di stato s_j indica il tipo di branching effettuato al livello j).

STEP 2. (CALCOLO DEL BOUND)

Risolvi il problema (BQ). Se non esistono soluzioni ammissibili, allora vai allo step 4.

Se la soluzione esiste, e $j=m$, allora vai allo step 5; se invece $j < m$ e $\mu_{BQ} \geq \mu_Q$ vai allo step 3.

Altrimenti vai allo step 1.

STEP 3. (BRANCHING AL MEDESIMO LIVELLO)

Se $s_j = -1$ allora vai allo step 1.

Se $s_j = 0$ e $|V^-| < \tau^-$ allora: $s_j \leftarrow -1$

$V^- \leftarrow V^- \cup \{j\}$

$V^0 \leftarrow$ Insieme Vuoto

vai allo step 2.

Se $s_j = 0$ e $|V^-| = \tau^-$ allora: vai allo step 4.

Se $s_j = +1$ e $|V^0| = 0$ allora: $s_j \leftarrow 0$

$V^0 \leftarrow \{j\}$

$V^+ \leftarrow V^+ - \{j\}$

vai allo step 2.

Se $s_j = +1$ e $|V^0| = 1$ e $|V^-| = \tau^-$ allora: vai allo step 4.

Se $s_j = +1$ e $|V^0| = 1$ e $|V^-| < \tau^-$ allora: $s_j \leftarrow -1$

$V^- \leftarrow V^- \cup \{j\}$

$V^+ \leftarrow V^+ - \{j\}$

vai allo step 2.

STEP 4. (BACKTRACK AL LIVELLO PRECEDENTE)

Se $s_j = 0$ allora: $V^0 \leftarrow$ Insieme Vuoto

Se $s_j = +1$ allora: $V^+ \leftarrow V^+ - \{j\}$

Se $s_j = -1$ allora: $V^- \leftarrow V^- - \{j\}$

Decrementa $j \leftarrow j - 1.$

Se $j = 0$ allora: STOP.

altrimenti: vai allo step 3.

STEP. 5. (TROVATA UNA SOLUZIONE)

Se $\mu_{BQ} < \mu_Q$ allora:

$$\begin{aligned} \mu_Q &\leftarrow \mu_{BQ} \\ \theta_Q &\leftarrow \theta_{BQ} \end{aligned}$$

Vai allo step 4.

L'algoritmo appena illustrato, che, come si è visto, enumera tutti i nodi dell'albero di ricerca relativo ai possibili insiemi di indici che definiscono il problema (P), determina la soluzione ottima globale di (P). Esso è stato codificato in linguaggio FORTRAN 77 per un Personal Computer che utilizza il sistema operativo MS-DOS. L'esperienza computazionale svolta ha mostrato che scegliendo il vettore iniziale θ della stima ai minimi quadrati del modello ed ordinando le osservazioni sulla base del calcolo della mediana con tale valore iniziale di θ , il numero dei nodi esaminati effettivamente è ragionevolmente contenuto.

BIBLIOGRAFIA

1. Bard, J. F., *Convex two-level Optimization*, Math. Progr., 40, 1988.
2. Clarke, F.H., *Nonsmooth Analysis and Optimization*, J. Wiley & Sons, 1983.
3. Dantzig, G.B., *Discrete Variable Extremum Problems*, Oper. Res., 5, 1957.
4. Demyanov, V.F. - Pallaschke, D. (eds), *Nondifferentiable Optimization: Motivations and Applications*, IIASA - Springer Verlag, 1985.
5. Fletcher, R., *Practical Methods of Optimization*, Vol. 2., J. Wiley & Sons, 1981.
6. Fletcher, R., "A Model Algorithm for Composite Nondifferentiable Optimization Problems", Math. Progr. Study 17, 1982.
7. Huber, P., *Robust Statistics*, J. Wiley & Sons, 1981.
8. Hampel, F.R. - Ronchetti, E.M. - Rousseeuw, P.J. - Stahel, W.A., *Robust Statistics*, J. Wiley & Sons, 1986.
9. Rousseeuw, P.J., *Least Median of Squares Regression*, J.A.S.A., 79, 388, 1984.
10. Shor, N.Z., *Minimization Methods for Nondifferentiable Functions*, Springer Verlag, 1985.
11. Souvaine, D.L. - Steele, J.M., *Time - and Space - Efficient Algorithms for Least Median of Squares Regression*, J.A.S.A., 82, 399, 1987.
12. Steele, J.M. - Steiger, W.L., *Algorithms and Complexity for Least Median of Squares Regression*, Discr. Appl. Math., 13, 1986.

13. Stefanini, L., *Su un problema di stima parametrica robusta*, XII Congresso Annuale A.M.A.S.E.S., Palermo, 1988.