
BOLLETTINO

UNIONE MATEMATICA ITALIANA

Sezione A – La Matematica nella Società e nella Cultura

LUCIANO PANDOLFI

Viaggi avventurosi e linee telefoniche: l'idea di controreazione

*Bollettino dell'Unione Matematica Italiana, Serie 8, Vol. 3-A—La
Matematica nella Società e nella Cultura (2000), n.1, p. 57–79.*

Unione Matematica Italiana

[<http://www.bdim.eu/item?id=BUMI_2000_8_3A_1_57_0>](http://www.bdim.eu/item?id=BUMI_2000_8_3A_1_57_0)

L'utilizzo e la stampa di questo documento digitale è consentito liberamente per motivi di ricerca e studio. Non è consentito l'utilizzo dello stesso per motivi commerciali. Tutte le copie di questo documento devono riportare questo avvertimento.

*Articolo digitalizzato nel quadro del programma
bdim (Biblioteca Digitale Italiana di Matematica)
SIMAI & UMI*

<http://www.bdim.eu/>

Viaggi avventurosi e linee telefoniche: l'idea di controreazione.

L. PANDOLFI

1. – Introduzione.

Il concetto di controreazione, o con termine inglese *feedback*, si impara a conoscere leggendo, da ragazzi, romanzi di mare e di avventure. Per esempio, il capitano di una nave, diciamo il Nautilus, per raggiungere un'isola lontana, vorrà seguire il cerchio massimo sulla superficie della terra che passa per il punto di partenza e quello di arrivo; si sa infatti che tra tutte le curve che congiungono due punti di una sfera, quella più corta è la metà più breve del cerchio massimo per i due punti.

Dunque il capitano, diciamo il capitano Nemo, parte dal punto iniziale seguendo una traiettoria che fa un certo angolo, letto sulla carta nautica, col meridiano... e non arriva da nessuna parte! Infatti, il piccolo errore di rotta inizialmente commesso, anche trascurando ogni errore successivo, su un percorso lungo farà passare la nave fuori dalla vista dell'isola. Si sa però come si rimedia a questo problema: ogni giorno viene «fatto il punto» ossia viene calcolata la posizione della nave, che non sarà quella teorica. E allora viene modificata la rotta, un po' più verso nord o verso sud, in modo da rimediare via via agli errori commessi. In questo modo la nave passerà in vista dell'isola e potrà raggiungerla.

Si noti che questo procedimento ha un costo: la traiettoria teorica è un arco di cerchio massimo, la traiettoria reale è un percorso zigzagante, che può essere sensibilmente più lungo.

Applicazioni, più o meno implicite, di quest'idea della correzione di errori «in corso d'opera» si trovano in meccanismi costruiti fin dall'antichità, per esempio negli orologi meccanici (si veda [7] per una storia di tali applicazioni). Si tratta però di applicazioni empiri-

che, non basate su alcuna teoria. In molti casi, la controeazione era fornita semplicemente dagli attriti insiti nel meccanismo, che impedivano alle sue parti di acquistare velocità troppo grandi, dato che *gli attriti stabilizzano il sistema*. È questo il caso del *regolatore di Watt*, noto dai corsi di Fisica: se si riduce troppo l'attrito delle sue parti il meccanismo inizia ad oscillare. A causa di questo fenomeno, e dell'importanza del regolatore di Watt nelle macchine termiche, J. C. Maxwell ne analizzò la teoria in un articolo pubblicato nel 1868⁽¹⁾, si veda [8]⁽²⁾.

Si considera comunemente il 1868 come inizio della *teoria* dei sistemi «reazionati». Di fatto però, come osservato da Bode in [4], i sistemi di regolazione meccanici, anche se descritti da equazioni differenziali non lineari piuttosto complesse, hanno pochi componenti e quindi l'origine reale della *teoria* dei sistemi reazionati va ricercata nel *progetto* dei primi amplificatori (reazionati) per le linee telefoniche.

In questo articolo vogliamo brevemente illustrare tre momenti cruciali dello sviluppo di questi amplificatori, due dei quali di tipo matematico, che ebbero luogo come anche il progetto stesso di amplificatore, nei *Bell Telephone Laboratories*.

L'importanza economica della diffusione e del possesso delle informazioni è stata uno degli stimoli forti per le innovazioni da quando, con l'introduzione dei giornali, queste hanno cominciato a sfuggire al controllo dei principi. Dopo l'introduzione dei giornali, il passo successivo è stato l'introduzione del telegrafo, che ha stimolato per esempio gli studi di Wheatstone sull'elettromagnetismo e i circuiti. Il telegrafo elettrico fu introdotto da Morse nel 1837. Già alla metà del secolo le linee telegrafiche connettevano Halifax in Canada con

⁽¹⁾ Pochi anni dopo della pubblicazione del suo articolo cruciale sull'elettromagnetismo, che è del 1864.

⁽²⁾ È curioso il destino di questo termine, *governors*, che figura nel titolo del lavoro di Maxwell: il termine fu greccizzato da Wiener, ottenendo il termine *cybernetics* (in italiano *cibernetica*) ormai poco usato. Norbert Wiener si occupava di questi problemi durante la seconda guerra mondiale, nei suoi studi sul progetto di sistemi di regolazione militari. Di fatto il termine *cibernetica* era già stato usato oltre un secolo prima da Ampère nel senso di *buon governo* nel suo *Saggio sulla filosofia delle scienze*, si veda [1, p. 482].

New Orleans e, citando da [13]: «presto arriveranno al Messico, e alla California. All'arrivo de'vapori postali inglesi ad Halifax trasmettessi immediatamente per telegrafo un sunto delle notizie Europee che viene pubblicato lo stesso giorno in tutte le principali città dell'Unione... I servigi che il telegrafo rende al commercio sono incalcolabili, e ne serve di prova lo stesso venir rotti non di rado i fili all'arrivo de'vapori d'Europa, quando vi è cambiamento notevole de'mercati. Fuori di questo caso, rarissimo è che vengano rotti...». La telefonia, diffusa da Bell a partire dal 1876, ebbe velocemente un impatto anche più imponente, ma si scontrò con problemi nuovi e ben più complessi. Tra questi il problema dell'attenuazione e distorsione dei segnali dovuta alla lunghezza delle linee, un problema che era importante anche per le linee telegrafiche; ma in tal caso bastava che arrivasse un segnale, anche deformato, ma sufficiente ad azionare un piccolo interruttore, che a sua volta apriva e chiudeva il circuito di un potente elettromagnete usato per imprimere il segno sulla striscia di carta. Invece, l'attenuazione e le deformazioni del segnale telefonico potevano rendere inutilizzabile la linea. Di conseguenza, alla fine dell'800, le linee telefoniche venivano correntemente usate in città mentre le dispersioni (dovute alla resistenza dei cavi, al cattivo isolamento ecc.) rendevano costosi e difficili i collegamenti fra città diverse. Per diminuire l'effetto di tali dispersioni, il primo cavo telefonico transcontinentale, tra New York e San Francisco, inaugurato nel 1915 (una quarantina di anni dopo il brevetto di Bell) pesava circa 3 quintali al chilometro; nonostante che già impiegasse alcuni amplificatori. Ciascuno di questi però causava distorsioni ed introduceva rumori. Questi disturbi venivano amplificati dagli amplificatori successivi, ciascuno dei quali introduceva ulteriori disturbi, rendendo presto irricognoscibile la voce (si verificavano anche gravi problemi di interferenza e di eco quando più messaggi erano inviati contemporaneamente sulla stessa linea; ma noi non ce ne occuperemo). Per questa ragione solo pochi amplificatori potevano essere usati (ne furono usati 6 nella linea New York-San Francisco). Per confronto, in esperimenti condotti nel 1928, dopo l'introduzione dell'amplificatore reazionato, prestazioni simili a quelle della linea

New York–San Francisco furono ottenute con cavi pesanti circa mezzo quintale al chilometro e 70 amplificatori [3].

Per approfondire le idee relative al concetto di controreazione, si può consultare un buon libro di teoria dei sistemi, come per esempio [6, 12].

2. – Black ed il progetto dell'amplificatore reazionato.

Un amplificatore è un sistema che riceve un segnale (elettrico nel caso che qui interessa) e restituisce *lo stesso segnale* moltiplicato per un coefficiente di amplificazione μ , detto *guadagno*. Per esempio se il segnale di ingresso è $e(t) = \cos \omega t$ allora il segnale di uscita deve essere $\mu \cos \omega t$. Questo obiettivo è irraggiungibile per varie ragioni (diremo qualcosa su questo più avanti) ed il segnale di uscita può essere molto diverso dal segnale desiderato $\mu \cos \omega t$, specialmente se si vuole che il guadagno sia grande. Inoltre, ogni amplificatore aggiunge al segnale un *rumore*.

Non è difficile costruire un singolo stadio amplificatore nel quale questi problemi siano trascurabili; ma, come si è detto, quando si mettono *in cascata*, ossia uno di seguito all'altro, due amplificatori, il secondo amplificherà, oltre al segnale utile, anche tutte le distorsioni, rumori ecc. introdotti dal primo. Ciò impone di limitare il numero degli amplificatori oppure di costruire amplificatori *estremamente buoni*. I primi tentativi di costruire amplificatori estremamente buoni migliorando la qualità dei componenti produssero scarsi risultati. Anche altri trucchi, come quello di far amplificare parte del segnale da amplificatori diversi, non furono soddisfacenti. L'idea giusta venne infine ad Harold Black, un giovane ingegnere dei *Bell Telephone Laboratories* nel 1927. L'idea consisteva nel confrontare istante per istante il segnale in uscita con quello in ingresso (l'analogo del «fare il punto» di una rotta) e usare questa informazione per ridurre i rumori e le distorsioni.

Black aveva studiato questo problema per cinque anni ricavando del tempo dai suoi incarichi principali, dato che in quel periodo non si riteneva produttivo il tempo speso per costruire amplificatori che

annullassero in qualche modo i disturbi da essi stessi prodotti⁽³⁾; e anche quando l'idea fu perfezionata, gli uffici brevetti ebbero enormi difficoltà a capirla: il progetto fu presentato nel 1928 all'ufficio brevetti americano ed a quello inglese, accompagnato da un libretto di 150 tra pagine di testo e figure, che era il primo libro sulla nuova idea di controreazione. Il brevetto americano fu rilasciato solamente nel 1937; e non andò meglio in Inghilterra dove si convinsero della bontà del progetto soltanto dopo aver ricevuto un prototipo funzionante. Ciò nonostante, già nel 1928 l'amplificatore reazionato di Black fu sperimentato con successo.

Diamo ora alcuni dettagli.

2.1. *L'idea di Black.*

Schematicamente un amplificatore si rappresenta come in fig. 1:

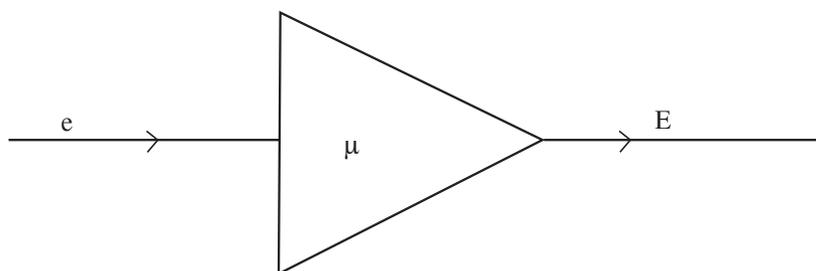


Fig. 1

In questa figura il triangolo rappresenta un'amplificatore di guadagno μ . Questo vuol dire che un *segnale di ingresso* $e(t)$, ossia un segnale applicato all'amplificatore produce, (*vorremmo che producesse!*) in uscita un segnale

$$E(t) = \mu e(t).$$

In pratica μ dovrà essere molto maggiore di 1.

⁽³⁾ Non bisogna pensare che nei *Bell Telephone Laboratories* non ci fosse entusiasmo per le novità. Una bella rievocazione degli entusiasmi di quel periodo si trova in [9].

In realtà le cose non sono quasi mai così semplici per varie ragioni, di cui consideriamo solo le seguenti. Prima di tutto ogni amplificatore genera un rumore e quindi il segnale di uscita è in realtà

$$E(t) = \mu e(t) + n(t);$$

inoltre, ogni amplificatore *distorce* il segnale. Questo vuol dire che se il segnale di ingresso ha forma

$$e(t) = e_0 \cos \omega t$$

allora il segnale di uscita ha forma

$$E(t) = \mu e(t) + n(t) = \mu \cos(\omega t) + n(t).$$

Bisogna però notare che *il guadagno μ dipende dalla frequenza del segnale*, $\mu = \mu(i\omega)$ ⁽⁴⁾.

Se il segnale $e(t)$ non è periodico, si può ancora scrivere la relazione

$$(1) \quad E = \mu e + n$$

tra i segnali di ingresso ed uscita ed il disturbo. Però ora i simboli E , e , n indicano una particolare rappresentazione ⁽⁵⁾ dei segnali su cui non entriamo, usando liberamente la formula (1).

L'idea di Black per ovviare a questi problemi è illustrata dalla

⁽⁴⁾ Convenzionalmente la frequenza si rappresenta come un punto dell'asse immaginario. Per questo scriviamo $\mu(i\omega)$ invece che $\mu(\omega)$. Inoltre, il segnale di uscita potrebbe essere sfasato rispetto a quello di ingresso, ossia potrebbe essere del tipo $\mu \cos(\omega t + \phi) + n(t)$. Lo sfasamento, che non discutiamo per brevità, è altrettanto importante del rumore.

⁽⁵⁾ Questa rappresentazione è uno strumento matematico che serve a rappresentare un segnale come «somma» di segnali periodici.

fig. 2, che è una versione semplificata della fig. 1 dell'articolo [2].

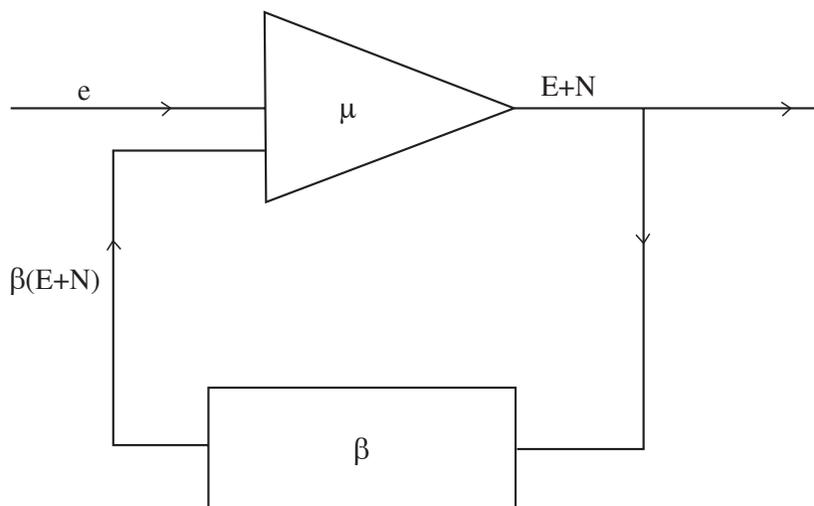


Fig. 2

Il rettangolo rappresenta un circuito, detto di *controreazione*, che preleva una «piccola» parte del segnale di uscita riportandola all'ingresso dell'amplificatore rappresentato dal triangolo. Dunque anche il rettangolo si può considerare come un amplificatore, ma di guadagno β molto minore di 1. Il circuito completo è un nuovo amplificatore la cui uscita sarà somma di un «segnale utile» $E(t)$ e di un rumore $N(t)$ (in generale diverso da $n(t)$, il rumore prodotto in assenza di controreazione, ossia quando $\beta = 0$).

Sia $\beta(E + N)$ la parte del segnale di uscita che la controreazione riporta in ingresso. Allora, il segnale di ingresso all'amplificatore rappresentato dal triangolo è

$$e + \beta(E + N).$$

Il segnale $E + N$ in uscita è dato dalla formula (1), nella quale e va sostituito con $e + \beta(E + N)$. Abbiamo quindi che $E + N$ verifica l'uguaglianza

$$E + N = \mu\{e + \beta[E + N]\} + n.$$

Risolvendo rispetto ad $E + N$ si trova

$$E + N = \frac{\mu}{1 - \mu\beta} e + \frac{1}{1 - \mu\beta} n$$

così che l'uscita del sistema «reazionato», ossia del sistema descritto in figura, si può considerare come somma di

$$E = \frac{\mu}{1 - \mu\beta} e, \quad N = \frac{1}{1 - \mu\beta} n.$$

Si vede quindi che se $|\mu\beta| \gg 1$, il segnale che appare in uscita al sistema reazionato ha per rumore solo una piccola frazione del rumore che avremmo senza controreazione. Inoltre, sempre se $|\mu\beta| \gg 1$, il guadagno $\gamma = \mu/[1 - \mu\beta]$ del sistema reazionato verifica

$$|\gamma| = \left| \frac{1}{\beta} \right| \cdot \left| \frac{1}{(1/\mu\beta) - 1} \right| \sim \left| \frac{1}{\beta} \right|.$$

Come si è detto, il guadagno μ dell'amplificatore dipende dalla frequenza e quindi c'è da aspettarsi che anche β dipenda dalla frequenza. Però in pratica il guadagno μ è molto grande mentre β è piccolo, e non è difficile costruire sistemi che hanno β piccolo e sensibilmente costante. Dunque, mentre il guadagno dell'amplificatore *da solo* risente fortemente della frequenza del segnale, il guadagno dell'amplificatore reazionato è sostanzialmente indipendente dalla frequenza.

Una controreazione β per cui $|1 - \mu\beta| > 1$ si dice *negativa*. Si dice invece *positiva* se $|1 - \mu\beta| < 1$. La funzione S (dipendente dalla frequenza ω , $S = S(i\omega)$) data da

$$S = \frac{1}{1 - \mu\beta}$$

si chiama la *sensibilità* del sistema in ciclo chiuso al disturbo n .

Dunque la controreazione negativa tende ad uniformare le prestazioni dell'amplificatore (e corregge anche ulteriori difetti).

Ci possiamo ora chiedere quale sia il costo di introdurre la con-

troreazione β . La prima osservazione è che l'originale guadagno μ viene sostituito dal guadagno *minore* $1/\beta$. Comunque, dimensionando opportunamente μ e β si può far sì che questo valore sia accettabile.

Un problema ben più importante è il seguente: in principio, chiudendo un circuito di reazione, può verificarsi un fenomeno distruttivo per l'amplificatore: un segnale può circolare infinite volte tra l'amplificatore e il circuito di reazione, amplificandosi via via. In questo caso si dice che l'amplificatore è *instabile*. Se ciò avviene, può accadere che un disturbo iniziale, che agisce appena per un attimo, produca un segnale in uscita che non si estingue, o addirittura che tende ad aumentare (finché il sistema non si rompe, o comunque assume un comportamento inaccettabile). L'esperienza di Black nel progetto di amplificatori suggeriva che questo potesse evitarsi. Lo strumento matematico giusto per controllare il verificarsi di tale fenomeno fu anch'esso ideato nei laboratori della *Bell Telephon Company*, e va sotto il nome di *Metodo di Nyquist* che verrà descritto, per sommi capi, nel paragrafo 3.

Un terzo problema, il problema di «quanto si perde» con l'uso di una controreazione, si illustrerà nel paragrafo 4.

Prima di vedere come si affrontano questi problemi, può essere opportuno soddisfare una legittima curiosità: capire per quale ragione uno dei semplici circuiti amplificatori che si studiano nei corsi di fisica debba dare origine a problemi gravi e come si realizzi e quanto sia complesso un circuito che agisce come controreazione.

2.2. *Un esempio.*

Per soddisfare la curiosità del lettore, mostriamo un semplicissimo amplificatore (estremamente più semplice di quello proposto da Black, si veda per confronto la fig. 2 in [2]). In linea con la strumentazione disponibile al tempo di Black, e inoltre perché il funzionamento di una valvola termoionica è assai più intuitivo, e noto, di quello dei transistori che si usano oggi, descriviamo un amplificatore a triodo.

Il circuito amplificatore che illustriamo è quello nella fig. 3, e dovrebbe essere noto dai corsi di fisica:

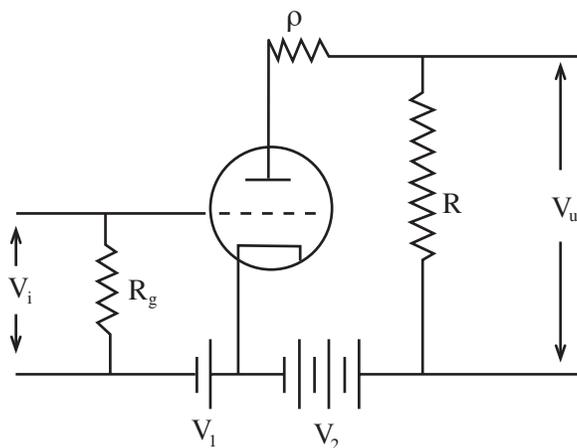


Fig. 3

Supponiamo che il segnale di ingresso V_i sia in modulo minore del potenziale V_1 della batteria che si trova tra catodo e griglia. In questo modo la griglia rimane sempre negativa rispetto al catodo e la sua unica funzione è di ostacolare il passaggio degli elettroni tra catodo ed anodo, tanto più quanto più è negativo il suo potenziale rispetto al catodo. Dunque, nella resistenza R_g non circola corrente. In questo modo il triodo funziona come un generatore di tensione, comandato dalla griglia: la tensione tra catodo e anodo è

$$V_c = V_2 + \mu V_i$$

(ove l'indice «c» sta per «catodo» ed «i» per «ingresso»). Il coefficiente μ è il *guadagno del tubo*.

Niente vieta di assumere come 0 il valore costante V_2 .

Si consideri ora il significato delle due resistenze ρ ed R : con ρ indichiamo la resistenza *interna* del tubo elettronico, mentre R è effettivamente un elemento del circuito. Il potenziale che si preleva in

uscita è il potenziale V_u (u per «uscita») ai morsetti di R e quindi è

$$V_u = \frac{R}{R + \rho} \cdot (\mu V_i) = \mu_0 V_i.$$

Il potenziale V_u è quello effettivamente applicato per esempio ai successivi stadi amplificatori e si vede, dalla formula precedente, che esso, a parità di R , è minore se ρ è grande.

Per avere un guadagno μ_0 circa uguale a quello teorico μ bisogna avere $R \gg \rho$.

In realtà l'impedenza interna del triodo non è indipendente dalla frequenza, e non è nemmeno puramente resistiva. Infatti, la coppia catodo/anodo costituisce un condensatore, se pure di piccola capacità; e quindi $\rho = \rho(i\omega)$. Ciò spiega perché il guadagno μ_0 dell'amplificatore *varia con la frequenza del segnale*.

Si trova sperimentalmente che l'impedenza $\rho(i\omega)$ si può rappresentare connettendo in parallelo le due impedenze che corrispondono alla resistenza interna del tubo (calcolata per segnali costanti) e alla capacità catodo/anodo (e, a frequenze ancora superiori, bisognerà considerare le capacità indotte dalla presenza della griglia). Dunque, il comportamento reale è quello del circuito di fig. 4:

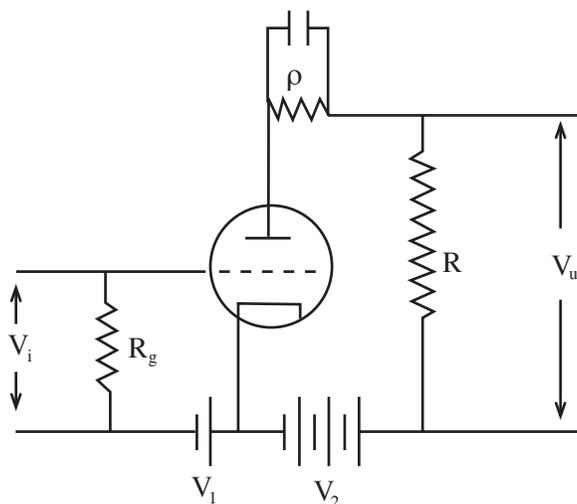


Fig. 4

In conclusione, il guadagno dell'amplificatore che abbiamo descritto è costante solamente su una piccola banda di frequenze. È questa una delle tante ragioni che limitano l'utilizzo di un amplificatore così semplice.

Per costruire un amplificatore il cui guadagno sia costante su una banda di frequenze più ampia, bisogna associare all'amplificatore un ulteriore circuito che aumenti il segnale di ingresso quando il guadagno tende a diminuire. Ciò può ottenersi, per esempio, con un circuito del tipo di quello in fig. 5.

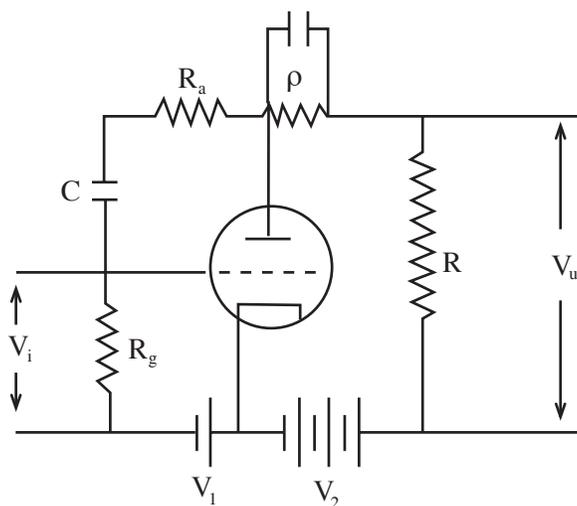


Fig. 5

Il condensatore C serve soltanto ad impedire che la corrente continua prodotta dalle batterie percorra il circuito R_a-C-R_g ; se opportunamente scelto, esso non viene visto dalla corrente del segnale, che può traversarlo liberamente. In questo modo, il triodo funziona anche come un generatore di tensione nel circuito R_a-C-R_g . Si noti che il potenziale che esso genera *si oppone* al potenziale di polarizzazione della griglia. Dunque, l'effetto del circuito R_a-C-R_g è di *diminuire* il potenziale della griglia, *diminuendo quindi il guadagno dell'amplificatore*. Però il guadagno non viene diminuito in modo costante perchè il potenziale che viene sottratto a quello del segnale di

griglia è

$$\frac{R_g}{R_g + R_a} i ;$$

dipende quindi dalla corrente che circola nel circuito R_a -C- R_g e questa a sua volta *dipende dal potenziale generato dal triodo*.

Abbiamo quindi una controreazione che riporta una parte del segnale dall'uscita all'ingresso dell'amplificatore.

Non è affatto ovvio che l'introduzione di una controreazione debba mantenere stabile un sistema. Per esempio, si può pensare che un disturbo imprevisto e «troppo forte» possa rendere la griglia positiva rispetto al catodo, innescando una circolazione di corrente nel circuito di reazione, che non tende ad estinguersi col tempo. Si pone quindi il problema di dare criteri che garantiscano la stabilità del sistema reazionato.

3. – La stabilità dei sistemi reazionati.

Veniamo ora ad un criterio per studiare la stabilità di un sistema reazionato che va sotto il nome di *Criterio di Nyquist*, introdotto in [10]. Si tratta di una semplice applicazione del *principio dell'argomento*, noto dalla teoria delle funzioni di variabile complessa, che ora descriviamo. L'importanza del criterio di Nyquist dipende sia dalla semplicità di applicazione, dato che conduce ad un *metodo grafico*, sia, più ancora, dal fatto che esso può essere applicato a dati ottenuti sperimentalmente su un prototipo, e non soltanto ottenuti nei calcoli teorici nel corso del progetto.

Descriviamo prima di tutto il principio dell'argomento.

3.1. *Il principio dell'argomento.*

Consideriamo le tre funzioni

$$f(z) = (z - z_0)^n, \quad g(z) = \frac{1}{(z - z_0)^n}, \quad h(z) = \frac{1}{(z - z_1)^m}$$

della variabile complessa z . Sia γ la curva di parametrizzazione

$$\gamma: \quad z = z_0 + \rho e^{i\theta}, \quad 0 \leq \theta \leq 2\pi.$$

La curva γ è una circonferenza intorno al punto z_0 , di raggio ρ .

Si noti che al crescere di θ da 0 a 2π , la circonferenza viene descritta in verso *antiorario*, verso che si sceglie convenzionalmente come *verso positivo* sul piano complesso.

Indichiamo con $f(\gamma)$ la curva immagine di γ sotto l'azione di f . Ossia, $f(\gamma)$ è parametrizzata da $f(z_0 + \rho e^{i\theta})$, ancora per $0 \leq \theta \leq 2\pi$.

È immediato vedere che $f(\gamma)$ è ancora una circonferenza però intorno al punto 0, *percorsa n volte ed in verso positivo*; invece, $g(\gamma)$ è ancora una circonferenza intorno al punto 0, *percorsa n volte ed in verso negativo*. Infatti, le parametrizzazioni di $f(\gamma)$ e di $g(\gamma)$ sono rispettivamente

$$\rho^n e^{in\theta}, \quad \rho^{-n} e^{-in\theta};$$

ed è immediatamente evidente che il verso di percorrenza è legato alla presenza di un polo oppure di uno zero della funzione.

Consideriamo ora $h(\gamma)$. Per semplicità consideriamo il caso $m = 1$ e supponiamo $z_0 = 0$ e $z_1 > 0$. In questo caso, si verifica immediatamente che la $h(\gamma)$ ha la seguente parametrizzazione:

$$\frac{1}{\rho e^{i\theta} - z_1} = \frac{\rho e^{-i\theta} - z_1}{\rho^2 - 2\rho z_1 \cos \theta + z_1^2}.$$

Il denominatore è un numero positivo.

Il numeratore è la parametrizzazione di una circonferenza che gira intorno all'origine se e solo se $\rho > z_1$. La positività del denominatore mostra che anche la curva $h(\gamma)$ gira intorno all'origine esattamente se $\rho > z_1$.

Queste osservazioni possono essere generalizzate e si ha il seguente risultato, noto come *principio dell'argomento*: sia $T(z)$ una funzione razionale e sia γ una curva semplice e chiusa percorsa in verso positivo.

Supponiamo che γ non incontri né zeri né poli di $T(z)$.

Il numero dei giri (in senso positivo) di $T(\gamma)$ intorno a 0 è uguale alla differenza tra il numero di zeri ed il numero

dei poli di $T(z)$ racchiusi da γ (zeri e poli vanno contati con la loro molteplicità).

Con un procedimento di approssimazione, quando $T(z)$ è una funzione razionale col grado del denominatore che eccede di due quello del numeratore si può provare un principio analogo con la regione interna a γ sostituita dal semipiano $\{z : \Re z > 0\}$ e con la curva γ sostituita dalla sua frontiera, ossia dall'asse immaginario, percorso dal basso in alto.

Fatte queste premesse, passiamo a considerare il metodo di Nyquist.

3.2. Il metodo di Nyquist.

Come si è detto, il guadagno di un amplificatore dipende dalla frequenza ω del segnale di ingresso e conviene vedere le frequenze come numeri dell'asse immaginario: $\mu = \mu(i\omega)$. È un fatto che il guadagno μ è una funzione razionale *strettamente propria* (ossia col grado del denominatore maggiore di quello del numeratore) della frequenza ω del segnale e quindi che la funzione $\mu(i\omega)$ ha estensione $\mu(z)$ al piano complesso; e che un amplificatore è stabile se e solo se la funzione $\mu(z)$ ha solamente poli con parte reale *negativa*.

Sia ora $\mu(z)$ il guadagno di un amplificatore che è stabile. Il problema consiste nel dire se, riportando in ingresso una frazione β del segnale di uscita, l'amplificatore sia ancora stabile. Ci si riferisca alla fig. 6 per seguire le considerazioni che seguono, notando che β può essere sia positivo che negativo. Il caso $\beta = 0$ corrisponde all'assenza di reazione e non interessa.

Dalla fig. 6 si vedono le seguenti relazioni:

$$y = \mu(z)\{e - \beta y\} \quad \text{da cui } y = \frac{\mu(z)}{1 + \beta\mu(z)} e.$$

Per ipotesi la funzione $\mu(z)$ non ha poli in $\Re z > 0$ e quindi basta dare criteri su β perchè $1 + \beta\mu(z)$ non si annulli in $\Re z > 0$. Se β è tale che $\beta\mu(i\omega) \neq -1$ allora si può fare uso del principio dell'argomento, e si trova la conclusione seguente: l'amplificatore in ciclo chiuso è stabile se la curva di parametrizzazione $1 + \beta\mu(i\omega)$, $-\infty < \omega < +\infty$, non racchiude 0. Questa affermazione discende

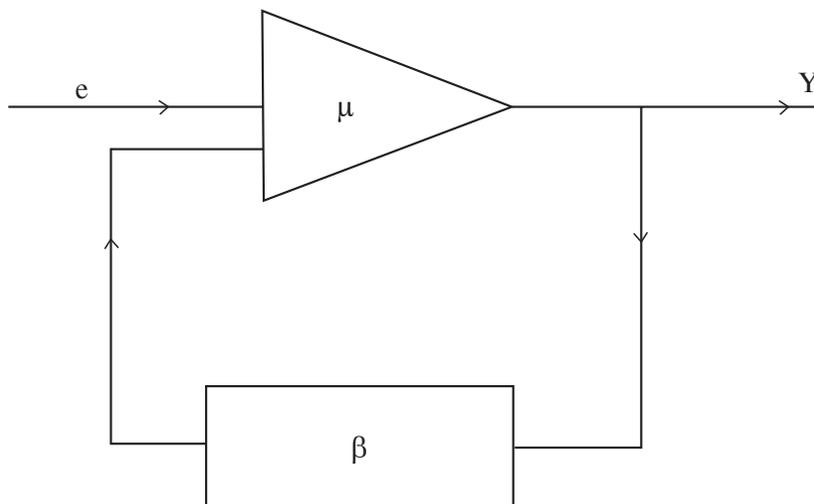


Fig. 6

dal principio dell'argomento e quindi vale quando il denominatore ha grado che eccede di due quello del numeratore.

È questa una forma (assai semplificata) del criterio di Nyquist, che in pratica conviene enunciare traslando verso destra di 1: si ha stabilità quando la curva $\beta\mu(i\omega)$ non racchiude il punto 1. Si noti che $\mu(i\omega)$ tende a zero per $|\omega| \rightarrow +\infty$ e quindi se $\beta\mu(i\omega)$ non gira intorno al punto 1 si ha stabilità del sistema reazionato, nella sola ipotesi che $\mu(i\omega)$ sia strettamente propria.

La forma che abbiamo presentato del criterio di Nyquist fa uso di ipotesi molto restrittive, che possono indebolirsi.

L'interesse del criterio di Nyquist è duplice: prima di tutto se è nota la curva $\mu(i\omega)$, è facile tracciare le curve $\beta\mu(i\omega)$ e quindi non solo si può vedere se il valore «desiderato» β_0 di β mantiene la stabilità del sistema; ma anche vedere quanto esso dista dal primo valore che produce instabilità; ossia valutare il *margin di stabilità*. La seconda ragione è che la curva $\mu(i\omega)$ è facilmente tracciabile in fase di progetto; ma è anche facile ottenerla da dati sperimentali e quindi stabilire se un amplificatore già costruito rimarrà stabile o meno sotto l'azione di una controreazione, senza conoscere esplicitamente la sua struttura interna.

4. – Le limitazioni introdotte dalla controreazione.

Si è detto che navigare correggendo via via la rotta porta facilmente a passare vicino all'obbiettivo, ma con un viaggio più lungo di quello teorico. Di fatto, l'introduzione di una controreazione ha sempre un costo. In particolare si trova sperimentalmente che, in molti casi, se si riduce la sensibilità ai disturbi di certe frequenze, si aumenta la sensibilità a disturbi di frequenza diversa. Questo problema è stato studiato da Bode [11] che ha dedotto una formula, nota come *identità di Bode*, che ora illustriamo. Ricordiamo che la sensibilità è la funzione

$$S = \frac{1}{1 + \mu\beta}$$

con $\mu\beta$ funzione razionale strettamente propria (pensando a μ fissato, la sensibilità dipende dal circuito di reazione, ossia da β). Supponiamo, come frequentemente accade, che il grado del denominatore della funzione $\mu\beta$ ecceda di (almeno) 2 quello del numeratore.

Consideriamo ora due casi:

- nel primo caso la funzione $\mu(z)$ non ha zeri nel semipiano $\Re z \geq 0$. In questo caso il sistema si dice *a minima rotazione di fase*, e vale

$$\int_{-\infty}^{+\infty} \log |S(i\omega)| d\omega = 0 ;$$

- nel secondo caso la funzione $\mu\beta$ ha zeri nel semipiano $\{z : \Re z > 0\}$. In questo caso vale

$$\int_{-\infty}^{+\infty} \log |S(i\omega)| d\omega \geq 0 .$$

Ora, si noti che la sensibilità, per $\omega \rightarrow +\infty$, tende ad 1 e quindi il suo logaritmo tende a zero ed è un infinitesimo di ordine maggiore di 1 nelle ipotesi dette, così che l'integrale è finito. Se per $\omega \in [-\omega_0, +\omega_0]$ la sensibilità è molto minore di 1, e quindi il logaritmo «molto» negativo, deve esserci un insieme di frequenze, al di fuori dell'inter-

vallo $[-\omega_0, +\omega_0]$ in cui il logaritmo deve essere positivo e, o di valore «molto grande» o di valore positivo su un insieme «grande». In ambedue i casi la sensibilità del sistema al di fuori della banda di frequenza $[-\omega_0, +\omega_0]$ non può venir ridotta. E questo effetto è più sensibile (si noti la disuguaglianza invece dell'uguaglianza) se il sistema non è a minima rotazione di fase.

La seconda delle formule precedenti illustra il ruolo «dannoso» che hanno gli zeri di $\mu(z)$ con parte reale positiva. Ciò è reso più evidente dai teoremi seguenti:

TEOREMA 1. — *Supponiamo che $\mu(z)$ non abbia zeri in un semipiano*

$$\Re z > -\sigma, \quad \sigma > 0;$$

ossia in un semipiano anche un po' a sinistra dell'asse immaginario. Per ogni $\varepsilon > 0$ ed ogni $k > 1$ si può trovare una controreazione β tale che la sensibilità sulla banda di frequenza $[-\theta_0, +\theta_0]$ sia minore di ε e in modo che in nessun punto la sensibilità superi k .

Bisogna notare che il circuito di reazione che va usato per applicare il teorema precedente può essere molto complesso. In generale β non sarà costante ma sarà essa stessa una funzione razionale di z , $\beta = \beta(z)$; ad essa si possono però imporre le ulteriori condizioni di essere limitata e priva di poli in $\Re z \geq 0$.

L'identità di Bode mostra che nel teorema precedente, quando $\varepsilon < 1$, non si può scegliere $k \leq 1$.

È chiaro che il teorema precedente non vale, per una generica banda di frequenza, se $\mu(z)$ ha uno zero sull'asse immaginario: in tale punto la sensibilità è uguale ad 1 qualunque sia la scelta della funzione $\beta(z)$, regolare in $\Re z \geq 0$. Più in generale:

TEOREMA 2. — *Supponiamo che z_0 sia uno zero di $\mu(z)$, con $z_0 = x_0 + iy_0$, $x_0 > 0$. Supponiamo inoltre che μ non si annulli sull'asse*

immaginario. Si fissi una qualsiasi banda di frequenza $[\theta_1, \theta_2]$. Esiste un esponente n tale che

$$(3) \quad \max_{[\theta_1, \theta_2]} |S(i\omega)| > \frac{1}{(\|S(i\omega)\|_\infty)^n}$$

dove

$$\|S\|_\infty = \max_{\omega \in (-\infty, +\infty)} |S(i\omega)|.$$

Vediamo quale conseguenza si trae dal Teorema 2. Ricordiamo che il nostro punto di vista è questo: μ è fissata e il circuito di reazione può costruirsi in vari modi, nel tentativo di ridurre la sensibilità sulla banda di frequenza $[\theta_1, \theta_2]$.

Si costruiscono successivamente vari circuiti di reazione che riducono sempre più vicina a zero la sensibilità sulla banda $[\theta_1, \theta_2]$. La (3) dice che, al variare del circuito di reazione, ossia di β ,

$$\frac{1}{\|S(i\omega)\|_\infty} \text{ tende a } 0, \text{ ossia } \|S(i\omega)\|_\infty \text{ tende ad } \infty:$$

riducendo via via la sensibilità sulla banda di frequenza $[\theta_1, \theta_2]$ associando al sistema dei circuiti di reazione via via più efficaci, la sensibilità, in corrispondenza di altre frequenze, deve crescere senza limite. Dunque un buon progettista deve essere capace di scegliere il miglior compromesso, riducendo la sensibilità sulla banda a cui è interessato senza però provocare problemi al di fuori di essa.

La dimostrazione del Teorema 2, da [5], è elementare ed interessante. Concludiamo mostrandola.

Prima di tutto osserviamo che esistono trasformazioni derivabili e con inversa derivabile che trasformano il semipiano $\Re z > 0$ nel disco $|z| < 1$, con l'asse immaginario che viene trasformato nella circonferenza (salvo il punto -1 che proviene dal «punto ∞ » del piano complesso). Una di tali trasformazioni è

$$(4) \quad w = \frac{z_0 - z}{\bar{z}_0 + z}$$

con z_0 un qualsiasi punto di $\Re z > 0$. In questo modo la funzione $w(z)$ viene ad avere un polo in $-\bar{z}_0$, che non appartiene al semipiano destro. L'inversa di questa trasformazione è

$$(5) \quad z = \frac{z_0 - \bar{z}_0 w}{1 + w}$$

il cui unico polo è $w = -1$. Questo punto sta sulla circonferenza ma non all'interno del disco e si può immaginare che il punto -1 venga trasformato nel «punto ∞ » del piano complesso.

Ogni funzione definita sul semipiano destro, in particolare la funzione $S(z)$, si trasforma in una corrispondente funzione definita sul disco: la $S(z)$ si trasforma in $\tilde{S}(w)$,

$$\tilde{S}(w) = S\left(\frac{z_0 - \bar{z}_0 w}{1 + w}\right).$$

Le due funzioni $S(z)$ e $\tilde{S}(w)$ prendono gli stessi valori; e introducendo il simbolo

$$\|\tilde{S}\|_\infty = \max_{|z| < 1} |\tilde{S}(z)|,$$

si ha

$$(6) \quad \|\tilde{S}\|_\infty = \|S\|_\infty.$$

Analogamente, a $\mu(z)$ viene a corrispondere la funzione $\tilde{\mu}(w)$.

Abbiamo detto che il punto z_0 che figura in (4), e quindi in (5) può essere arbitrariamente scelto in $\Re z > 0$. Andiamo a scegliere per z_0 proprio uno dei punti che figurano nell'enunciato del Teorema 2; ossia uno dei punti *con parte reale positiva* nei quali si annulla la funzione μ . Nel suo corrispondente w_0 si annulla $\tilde{\mu}$; e quindi $|\tilde{S}(w_0)| = 1$.

Bisogna sapere che le funzioni razionali di numeri complessi hanno una proprietà importante, che va sotto il nome di *principio del massimo modulo*: il modulo di una funzione razionale letta su un di-

sco chiuso (ossia comprendente anche la circonferenza) prende valore massimo; ma il valore massimo *non* viene assunto in punti interni al disco stesso. In particolare quindi

$$\max_{|w|=1} |\tilde{S}(w)| \geq \max_{|w| \leq 1} |\tilde{S}(w)| \geq |\tilde{S}(w_0)| = 1.$$

Consideriamo ora come si trasforma la banda di frequenza $[-\omega_1, \omega_1]$ sotto l'azione di (4). Ci si riferisca alla fig. 7. L'immagine della banda di frequenza $[-\omega_1, \omega_2]$ è un arco A della circonferenza $|w| = 1$ (che non contiene il punto -1). Indichiamo con ϕ l'ampiezza dell'angolo al centro sotteso da tale arco:

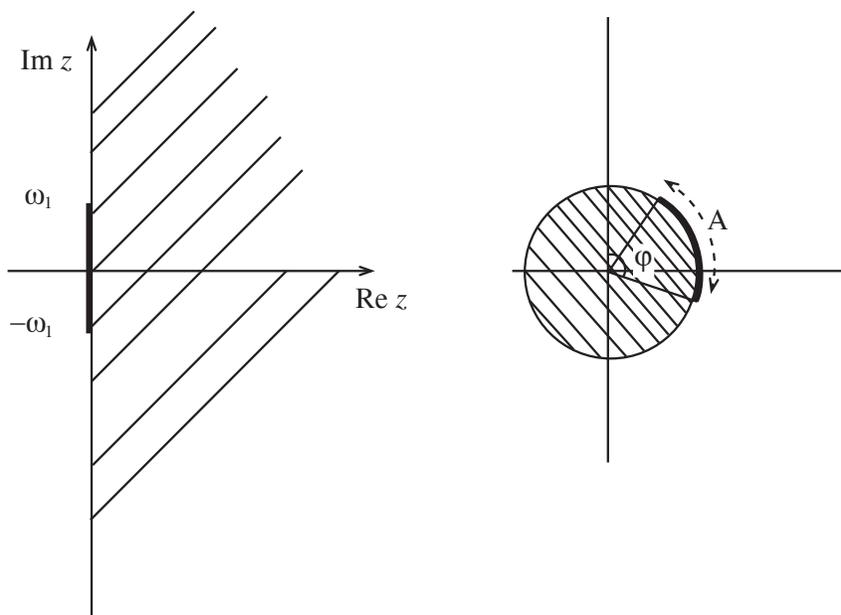


Fig. 7

Introduciamo ora un numero n ,

$$n > \frac{2\pi}{\phi} \quad \text{così che} \quad \psi = \frac{2\pi}{n} < \phi.$$

Fissiamo un punto w di modulo 1 e ruotiamolo degli angoli ψ , 2ψ , 3ψ , ..., $n\psi$, come nella fig. 8:

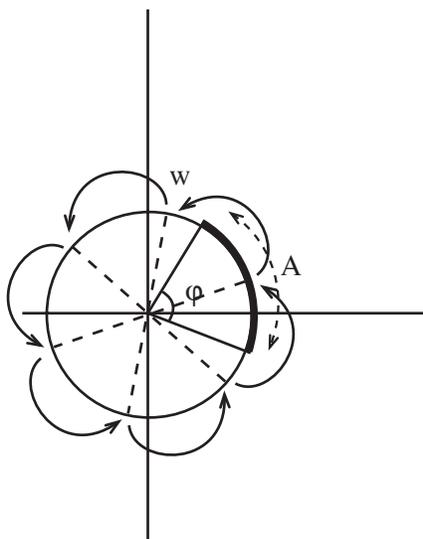


Fig. 8

L'ultima rotazione riporta w su se stesso e una delle intermedie deve portare w in un punto dell'arco A perché $\psi < \phi$.

La rotazione di un punto w di un angolo α sul piano complesso si ottiene moltiplicando w per $e^{i\alpha}$. Dunque la rotazione di un angolo $k\frac{2\pi}{n}$ trasforma il punto w nel punto $we^{k\frac{2\pi}{n}i}$.

Introduciamo ora la funzione

$$T(w) = \tilde{S}(w) \cdot \tilde{S}(we^{\frac{2\pi}{n}i}) \cdot \tilde{S}(we^{\frac{2\pi}{n}i}) \cdot \tilde{S}(we^{\frac{2\pi}{n}i}) \dots \tilde{S}(we^{(n-1)\frac{2\pi}{n}i}).$$

Come si è detto, uno almeno dei punti $we^{k\frac{2\pi}{n}i}$ cade nell'arco A e in tale punto la sensibilità è minore di ε .

D'altra parte, $|T(0)| = 1$ così che

$$\max_{\theta} |T(e^{i\theta})| \geq |T(0)| \geq 1,$$

dal principio del massimo modulo; e, ricordando che il modulo

di un prodotto è il prodotto dei moduli, si trova

$$1 \leq \max_{\theta} |\tilde{S}(e^{i\theta})|^{n-1} \cdot \left[\max_{[\theta_1, \theta_2]} |\tilde{S}(e^{i\theta})| \right] \leq \varepsilon \cdot (\|\tilde{S}\|_{\infty})^{n-1}.$$

Usando (6) si vede che questa è la disuguaglianza (3) cercata.

BIBLIOGRAFIA

- [1] A. M. AMPÈRE, *Opere*, a cura di M. Bertolini, vol.1, UTET, Torino, 1969.
- [2] H. S. BLACK, *Stabilized feed-back amplifiers*, Electrical Engineering, **53** (1934), 114-120 (ristampato in *Proc. of the IEEE*, **72** (1984), 716-722).
- [3] S. H. BLACK, *Inventing the negative feedback amplifier*, IEEE Spectrum, Dic. 1977, 55-60.
- [4] H. W. BODE, *Feedback, the history of an idea*, Proc. of the *Symposium on active networks and feedback systems*, Polytechnic Institut of Brooklyn, Polytechnic Press, 106-123, 1960.
- [5] B. A. FRANCIS - G. ZAMES, *On H^{∞} -optimal sensitivity theory for SISO feed-back systems*, IEEE Trans. Automatic Control, **AC-29** (1984), 9-16.
- [6] A. ISIDORI - A. RUBERTI, *Teoria dei sistemi*, Boringhieri, Torino, 1970.
- [7] O. MAYR, *The origins of feedback control*, The M.I.T. Press, Cambridge, 1970.
- [8] J. C. MAXWELL, *On governors*, Proc. of the Royal Soc., **16** (1867/1868), 270-283 (ristampato in R. Bellman, R. Kalaba Eds., *Mathematical trends in control theory*, Dover Publications, New York, 1964).
- [9] R. E. MCGINN, *Stokowski and the Bell Telephone Laboratories: collaboration in the development of the high-fidelity sound reproduction*, Technology and Culture, **24** (1983), 38-75.
- [10] N. NYQUIST, *Regeneration theory*, Bell Systems Tech. J., **11** (1932), 126-147 (ristampato in R. Bellman, R. Kalaba Eds., *Mathematical trends in control theory*, Dover Publications, New York, 1964).
- [11] H. W. BODE, *Network analysis and feedback amplifier design*, Van Nostrand, New York, 1945.
- [12] S. RINALDI, *Teoria dei sistemi*, CLUP, Milano, 1997.
- [13] A. SECCHI, *Sullo stato attuale della telegrafia*, Annali di Scienze Matematiche e Fisiche, **1** (1850), 23-41.

Luciano Pandolfi, Politecnico di Torino, Dipartimento di Matematica
 Corso Duca degli Abruzzi, 24, 10129 Torino - Italia. E-mail: lucipan@polito.it